External Assessment Group Report commissioned by the NIHR Evidence Synthesis Programme on behalf of NICE

Elacestrant for treating oestrogen receptor-positive, HER2-negative advanced breast cancer with an ESR1 mutation after at least 1 endocrine treatment [ID6225]

Produced by	**Southampton Health Technology Assessments Centre (SHTAC)**
Authors	Emma Maund, Research Fellow, Evidence synthesis Marcia Takahashi, Research Fellow, Health economics Joanne Lord, Professorial Research Fellow, Health economics Jonathan Shepherd, Principal Research Fellow, Evidence synthesis
Correspondence to	Dr Jonathan Shepherd Southampton Health Technology Assessments Centre (SHTAC) School of Healthcare Enterprise and Innovation University of Southampton Alpha House, Enterprise Road, University of Southampton Science Park, Southampton SO16 7NS www.southampton.ac.uk/shtac
Date completed	2nd July 2024; updated 6th August 2024
Source of Funding	This report was commissioned by the National Institute for Health and Care Research (NIHR) Evidence Synthesis Programme as project number 136220.

Acknowledgements

We thank the following for providing expert clinical advice to the project team and for reading and commenting on a draft of the report:

- Dr Peter Simmonds, Consultant Medical Oncologist, University Hospital Southampton NHS Foundation Trust, Southampton General Hospital.

We also thank:

- Lois Woods, Senior Research Assistant, Southampton Health Technology Assessments Centre (SHTAC) for performing literature searching and critically appraising search strategies.

Declared competing interests of the authors and advisors

The authors declare none

Rider on responsibility for report

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

This report should be referenced as follows:

Maund, E; Takahashi, M; Lord, J; Shepherd, J. Elacestrant for treating oestrogen receptor-positive, HER2-negative advanced breast cancer with an ESR1 mutation after at least 1 endocrine treatment [ID6225]. A Single Technology Appraisal. Southampton Health Technology Assessments Centre, 2024.

Contributions of authors

Emma Maund critically appraised the clinical effectiveness systematic review, and drafted the report; Marcia Tomie Takahashi critically appraised the health economic systematic review, critically appraised the economic evaluation, and drafted the report. Joanne Lord critically appraised the health economic systematic review, critically appraised the economic

evaluation, and drafted the report, Jonathan Shepherd critically appraised the clinical effectiveness systematic review, drafted the report and is the project co-ordinator and guarantor.

Confidential (CON) information is highlighted in blue and underlined

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

AE	Adverse event
AIC	Akaike information criteria
ALP + FUL	Alpelisib with fulvestrant
BIC	Bayesian information criteria
BNF	British National Formulary
CI	Confidence interval
CIC	Commercial in confidence
CRD	Centre for Reviews and Dissemination
CS	Company submission
CSR	Clinical study report
CQ	Clarification question
ctDNA	Circulating tumour DNA
DSA	Deterministic sensitivity analysis
DSU	Decision Support Unit
EAG	External Assessment Group
ECOG	Eastern Cooperative Oncology Group
EMC	Electronic Medicines Compendium
EPAR	European Public Assessment Report
ER	Oestrogen receptor
EQ-5D-3L	European Quality of Life Working Group Health Status Measure 3 Dimensions, 3 Levels
EQ-5D-5L	European Quality of Life Working Group Health Status Measure 5 Dimensions, 5 Levels
EQ-VAS	EuroQol Visual Analogue Scale
ESMO	European Society for Medical Oncology
ESR1	Oestrogen receptor 1
ESR1-mut	ESR1 mutation
ET	Endocrine therapy
EVE + EXE	Everolimus with exemestane
GLH	Genomic Laboratory Hub
GnRH	Gonadotropin-releasing hormone
GMS	Genomic Medicine Service
HER2-	Human epidermal factor receptor 2-negative
HR+	Hormone receptor-positive

HR	Hazard Ratio
HRG	Healthcare Resource Group
HRQoL	Health-related quality of life
HTA	Health technology assessment
ICER	Incremental cost-effectiveness ratio
IPD	Individual patient level data
IRC	Imaging review committee
ITT	Intent to treat
MAIC	Matching-adjusted indirect comparison
mBC	Metastatic breast cancer
MHRA	Medicines and Healthcare products Regulatory Agency
mITT	Modified intent to treat
NGS	Next generation sequencing
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health and Care Research
NR	Not reported
OS	Overall survival
PartSA	Partitioned survival analysis
PFS	Progression-free survival
PIK3CA	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha
PIK3CA-mut	PIK3CA mutation
PRO	Patient-reported outcome
PRO-CTCAE	Patient-Reported Outcome Common Terminology Criteria for Adverse Events
PSA	Probabilistic sensitivity analysis
PSM	Parametric survival model
PSS	Personal Social Services
QALY	Quality-adjusted life year
QoL	Quality of life
RCT	Randomised controlled trial
RDI	Relative dose intensity
RR	Relative risk/risk ratio
SAE	Serious adverse event

SD	Standard deviation
SE	Standard error
SERD	Selective oestrogen receptor degrader
SLR	Systematic literature review
SmPC	Summary of product characteristics
TA	Technology appraisal
TEAE	Treatment-emergent adverse event
TSD	Technical Support Document
TTD	Time to treatment discontinuation
UK	United Kingdom
US	United States
VAS	Visual analogue scale
WTP	Willingness to pay

1 EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the external assessment group (EAG) as being potentially important for decision making. It also includes the EAG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs).

Below we provide an overview of the key issues. Section 1.1 provides an overview of key model outcomes and the modelling assumptions that have the greatest effect on the ICER. Sections 1.2 to 1.4 explain the key issues in more detail. Background information on the condition, health technology, evidence and information on the issues are in the main EAG report.

All issues identified represent the EAG's view, not the opinion of the National Institute for Health and Care Excellence (NICE). Overview of the EAG's key issues

Table 1 Overview of EAG key issues

ID	Summary of issue	Report sections
Issue 1	Uncertainty in the clinical effectiveness of elacestrant based on post-hoc trial sub-group analyses	3.2.4
Issue 2	Uncertainty in the results of the matched adjusted indirect comparison (MAIC)	3.3 and 3.4
Issue 3	Uncertain overall survival extrapolations for elacestrant and comparators	4.2.4.2.1 and 4.2.4.3.1
Issue 4	Lack of evidence on comparator treatment duration	4.2.4.2.3 and 4.2.4.3.3
Issue 5	Practical implications and cost of introducing ESR1 mutation testing in the NHS	4.2.6.5

The key differences between the company's preferred assumptions and the EAG's preferred assumptions are:

- Target population (subgroup 1): the overall survival (OS) extrapolation for elacestrant (gamma rather than log-logistic); the price of everolimus (from eMIT rather than BNF).

- Dual mutation subgroup (subgroup 2): the proportion of positive ESR1 mutation tests (20% rather than 50%).

1.1 Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (overall survival) and quality of life in a quality-adjusted life year (QALY). An ICER is the ratio of the extra cost for every QALY gained.

Overall, the technology is modelled to affect QALYs by:

- Increasing overall survival
- Maintaining quality of life for longer due to extended progression-free survival

Overall, the technology is modelled to affect costs by:

- Increasing the cost of treatment in the target population (subgroup 1)
- Reducing the cost of treatment in the subgroup with a dual mutation (subgroup 2)
- Adding costs to introduce ESR1 testing

The modelling assumptions that have the greatest effect on the ICER are:

- The choice of OS extrapolations for elacestrant and the resulting difference in survival relative to comparators
- Differences in treatment duration for elacestrant (based on trial data) and comparators (assumed equal to PFS)
- Use of MAIC hazard ratios to model the comparator survival curves compared with independently fitted curves (using MAIC adjusted data for elacestrant)

1.2 The clinical effectiveness evidence: summary of the EAG's key issues

Issue 1 Uncertainty in the clinical effectiveness of elacestrant based on post-hoc trial sub-group analyses

Report section	3.2.4
Description of issue and why the EAG has identified it as important	Elacestrant is indicated for the treatment of postmenopausal women, and men, with ER+/HER2- locally advanced or metastatic breast cancer with an activating ESR1-mutation who have disease progression following at

	least one line of endocrine therapy (ET) including a CDK4/6 inhibitor. The company proposes that treatment with elacestrant should be targeted at two sub-groups of people eligible according to the marketing authorisation: • Subgroup 1 is people with an ESR1-mutation who have disease progression following ≥*12 months prior treatment* with endocrine therapy in combination with CDK4/6 inhibitor. • Subgroup 2, nested within subgroup 1, comprises people with an ESR1-mutation *and a PIK3CA-mutation (dual mutation*) who have disease progression following ≥*12 months prior treatment* with endocrine therapy in combination with CDK4/6inhibitor. The clinical effectiveness evidence for elacestrant in these subgroups is based on post hoc analyses of patients from the ongoing pivotal phase III, multicentre, randomised, open-label, active controlled trial comparing the efficacy and safety of elacestrant to endocrine monotherapy treatment (the EMERALD trial). The EAG urges caution in the interpretation of these results due to: • Small sample sizes, notably for post hoc subgroup 2 (13% of randomised patients). • Some evidence of selection bias due to imbalances in baseline characteristics between trial arms, affecting post hoc subgroup 2. In this subgroup there was a higher percentage of patients in the elacestrant arm with certain adverse prognostic factors, suggesting slightly more advanced cancer than the comparator arm. The impact of these imbalances is unclear. • The trial was not statistically powered for subgroups, thus statistical significance cannot be inferred from the

	results. The findings should be considered as exploratory, hypothesis-generating, rather than confirmatory.
What alternative approach has the EAG suggested?	None at present
What is the expected effect on the cost-effectiveness estimates?	This is uncertain currently.
What additional evidence or analyses might help to resolve this key issue?	Ideally a follow-up RCT in which patients in subgroups 1 and 2 are randomised to elacestrant and SOC, based on an appropriate sample size calculation. However, it is not feasible to design and complete such a trial within the timeframe of this NICE technology appraisal.

Issue 2 Uncertainty in the results of the matched adjusted indirect comparison (MAIC)

Report section	3.3 and 3.4
Description of issue and why the EAG has identified it as important	None of the treatments in the standard of care comparator arm of the EMERALD trial match the company's chosen comparators in the decision problem. Furthermore, none of the trials of the company's chosen comparator treatments tested patients for the ESR1 mutation. This limited the ability to do an indirect treatment comparison of elacestrant in patients with the ESR1 mutation in the EMERALD trial versus comparator treatments in similar patients in comparator trials. Due to the scarcity of ESR1 mutation testing in the UK and Europe the company did a targeted search for sources of real-world evidence in the US. They selected a registry of patient health records (the Flatiron database) to obtain data on patients with the ESR1 mutation treated with the relevant comparators.

	The company constructed an unanchored matched adjusted indirect treatment comparison (MAIC) using individual patient data from patients treated with elacestrant in the EMERALD trial, matched to aggregate data from patients treated with everolimus and exemestane or alpelisib and fulvestrant in Flatiron. The EAG notes some uncertainties in the methods used to construct the MAIC: • A set of 14 prognostic factors/effect modifiers were identified by key opinion leaders, but little information is given on the process and methodology. Sufficient information was available for just 3 of the 14 factors to allow their inclusion in the MAIC for the purpose of matching patients from EMERALD to Flatiron. Some widely accepted prognostic factors were not included such as bone metastases; number of metastatic sites and de novo vs. recurrent/progressed disease. This is a key limitation of the MAIC. • Other limitations include small effective sample sizes after weighting, particularly for post hoc subgroup 2 (dual mutation), and imbalances in weighted prognostic factors between elacestrant and comparator, again, notably in post hoc subgroup 2. • It is not explicitly stated how data on duration of previous endocrine therapy was identified in Flatiron. Exposure time for previous CDK6/4 inhibitor treatment was available and the EAG assumes that exposure time for previous CDK6/4 inhibitor treatment = exposure time for previous endocrine therapy, since in practice CDK6/4 inhibitor is usually given in combination with endocrine therapy. • Limited detail is provided on the methods of searching for relevant sources of real-world evidence. The Flatiron database was selected based on a “targeted” search in the US, rather than a systematic global search.

What alternative approach has the EAG suggested?	The Flatiron database could be replaced in the MAIC with the alternative real-world evidence source considered by the company - Patient360 Breast (ConcertAI). This appears to have a smaller sample of relevant patients than Flatiron, but it may potentially provide more comprehensive data on prognostic factors. Though uncertainty would likely remain, it could nonetheless be informative for decision making (e.g. as a scenario analysis).
What is the expected effect on the cost-effectiveness estimates?	The impact on the ICER is uncertain
What additional evidence or analyses might help to resolve this key issue?	In the shorter term, additional real-world evidence with greater coverage of prognostic factors relevant to this patient population. If this is not available from Flatiron a systematic search might identify other relevant patient registries. In the longer-term, clinical trial data comparing elacestrant head-to-head with other available treatments (e.g. everolimus + exemestane or alpelisib + fulvestrant) in patients with ESR1 mutation and PIK3CA-mutations.

1.3 The cost-effectiveness evidence: summary of the EAG's key issues

Issue 3 Uncertain overall survival extrapolations

Report section	4.2.4.2.1 and 4.2.4.3.1
Description of issue and why the EAG has identified it as important	There is high uncertainty over the OS extrapolations in the economic model due to the use of an unanchored MAIC, and the limited sample sizes for the subgroups from the EMERALD trial and the Flatiron comparator cohorts. We agree with the use of the gamma distribution for the everolimus + exemestane comparator in subgroup 1, as this is closest to current survival expectations. However, we consider that the company's choice of a log-logistic extrapolation for elacestrant that gives a long projected

	survival benefit is overly optimistic given the current evidence base. The company base case OS extrapolations for post hoc subgroup 2 are also uncertain, but do not give such an extended projection of survival benefit (survival estimates are similar between arms after 6 years).
What alternative approach has the EAG suggested?	For EAG analysis, we prefer a gamma OS extrapolation for elacestrant as well as for the comparator in subgroup 1. This gives a good statistical and visual fit in both arms and similar survival projections after 5 years.
What is the expected effect on the cost-effectiveness estimates?	The company's base case ICER increases from £24,893 to £43,793 pprer QALY gained in subgroup 1 (including the 1.2 QALY severity modifier weight) when a gamma distribution is used to extrapolate elacestrant OS (see 6.1.1.1).
What additional evidence or analyses might help to resolve this key issue?	Additional clinical expert opinion to assess the plausibility of the survival extrapolations. However, uncertainty over this issue cannot be resolved without more robust comparative evidence and longer follow-up.

Issue 4 Lack of evidence on comparator treatment duration

Report section	4.2.4.2.3 and 4.2.4.3.3
Description of issue and why the EAG has identified it as important	Mature data on treatment duration is available for elacestrant from the EMERALD trial. However, data on treatment duration is not available for comparators from the Flatiron cohorts. The company assume that time to treatment discontinuation (TTD) for the comparators is equal to PFS in the economic model. We are concerned about the potential for bias due to the use of different modelling assumptions for TTD in the elacestrant and comparator arms. This will result in over-estimation of treatment costs for the comparator relative to elacestrant if, in practice, a proportion of patients discontinue the comparator treatments before progression, as was observed for elacestrant. The difference between the

	company's TTD estimates for elacestrant and those for alpelisib + fulvestrant in subgroup 2 are particularly marked.
What alternative approach has the EAG suggested?	We report exploratory scenario analysis using an option included in the company's model to adjust the TTD curves for the comparators using an assumed hazard ratio relative to the comparator PFS.
What is the expected effect on the cost-effectiveness estimates?	The EAG scenario with ALP+FUL TTD estimated assuming a 0.5 hazard ratio relative to the ALP+FUL PFS curve in subgroup 2 changed the results of the company's base case from elacestrant being dominant to an ICER of £4,362 per QALY (see 6.1.1.2).
What additional evidence or analyses might help to resolve this key issue?	Additional evidence on the duration of treatment for alpelisib + fulvestrant in a population similar to subgroup 2 (dual mutation with at least 12 months of prior ET+CDK4/6i). Clinical expert opinion on expected treatment duration.

1.4 Other key issues: summary of the EAG's view

Issue 5 Introduction of ESR1 mutation testing

Report section	4.2.6.5
Description of issue and why the EAG has identified it as important	A test for ESR1 mutation would be necessary to assess patients' suitability for treatment with elacestrant, but this is not currently provided in the NHS. Genetic testing for breast cancer is routine prior to treatment, using a tissue sample and digital PCR assay. However, as ESR1 is an acquired mutation, analysis of the primary tumour sample may not be accurate. Digital PCR could be used to test for the ESR1 mutation when treatment with elacestrant is being considered. However, this would require a repeat biopsy, which may not reflect disease status due to tumour heterogeneity, and there is potential for delay to the start of treatment. In the EMERALD study, ESR1 testing was conducted using a blood sample and circulating tumour DNA (ctDNA) test. The company state that they would expect such a test to be

	introduced if elacestrant were to be recommended by NICE, as the PIK3CA test was introduced when alpelisib was recommended (TA816). North Thames NHS Genomic Laboratory Hub (GLH) currently provide a ctDNA test that can identify the ESR1 mutation (Marsden360 assay), and we understand that other NHS GLHs are exploring this or a similar approach. This test is relatively expensive and not routinely available. However, the cost would be likely to fall if testing for the ESR1 mutation and other potential treatment targets were to become routine, with next generation sequencing panel testing of ctDNA samples.
What alternative approach has the EAG suggested?	For their base case, the company assumed a cost of £300 per test (based on digital PCR) and 50% prevalence of ESR1 mutation: or £600 per case identified for treatment. We conducted exploratory scenario analysis assuming a higher cost for ctDNA ([redacted] or [redacted]) with and without adjustment for prevalence.
What is the expected effect on the cost-effectiveness estimates?	Assuming a cost of [redacted] per test and 50% prevalence ([redacted] per case identified), the company's base case ICER for subgroup 1 increases from £24,893 per QALY to £28,858 per QALY (QALY weight of 1.2 applied) The long-term impact on the ICER is lower if we assume that the cost of the ctDNA test would fall with routine use (e.g. £26,343 per QALY at [redacted] per test). It is also arguable that the test cost should not be adjusted for prevalence, or even that the test cost should not be included in ICER calculations, as and when NGS ctDNA testing were to become routine for multiple treatment targets at this point in the care pathway.
What additional evidence or analyses might help to resolve this key issue?	Further information on the expected cost of ESR1 mutation testing if implemented in the NHS.

1.5 Summary of EAG's preferred assumptions and resulting ICER

The cumulative effects of EAG preferred assumptions on the company's base case analysis are shown in Table 2 (subgroup 1 - *ESR1-mut* + ≥12 months prior ET with CDK4/6i population) and Table 3 (subgroup 2 - *ESR1-mut+PIK3CA-mut* + ≥12 months ET with CDK4/6i population). These results include a confidential patient access scheme (PAS) discount for elacestrant, but other drugs are costed at non-confidential NHS prices. We report results, including all confidential discounts for comparators and subsequent treatments in a confidential 'cPAS' addendum to this report.

Table 2 Cumulative effect of EAG changes to the company's base case analysis for subgroup 1 – patients with an activating ESR1-mutation with disease progression following ≥12 months prior treatment with ET + CDK4/6i

Scenario	Incremental cost	Incremental QALYs	ICER (£/QALY) No QALY weight	ICER (£/QALY) With 1.2 QALY weight
Company's base case	£18,883	0.632	£29,872	£24,893
+ Mean age from Flatiron ([redacted] years)	£18,872	0.630	£29,942	£24,952
+ Everolimus price from eMIT 2023	£30,080	0.630	£47,723	£39,769
+independent PSM extrapolation: Gamma for both arms	£27,898	0.317	£87,869	£73,224
EAG's base case	£27,898	0.317	£87,869	£73,224

Table 3 Cumulative effect of EAG changes to the company's base case analysis for *ESR1-mut+PIK3CA-mut* + ≥12 months ET with CDK4/6i population (subgroup 2)

Scenario	Incremental cost	Incremental QALYs	ICER (£/QALY)
Company's base case	-£12,269	0.277	Dominant
+ Mean age from Flatiron ([redacted] years)	-£12,269	0.277	Dominant
+ Proportion of positive cases after ESR1-mut testing (20%)	-£11,369	0.277	Dominant

Scenario	Incremental cost	Incremental QALYs	ICER (£/QALY)
EAG's preferred base case	-£11,369	0.277	Dominant

Modelling errors identified and corrected by the EAG are described in section 5.3. For further details of the exploratory and sensitivity analyses done by the EAG, see section 6.1.4.2.6.5

2 INTRODUCTION AND BACKGROUND

2.1.1 Introduction

This report is a critique of the company's submission (CS) to NICE from Menarini Stemline UK Ltd on the clinical effectiveness and cost effectiveness of elacestrant for treating oestrogen receptor-positive, HER2-negative advanced breast cancer with an ESR1-mutation after at least one endocrine treatment. It identifies the strengths and weakness of the CS. Clinical experts were consulted to advise the external assessment group (EAG) and to help inform this report.

Clarification on some aspects of the CS was requested from the company by the EAG via NICE on 13th May 2024. A response from the company via NICE was received by the EAG on 4th June 2024 and this can be seen in the NICE committee papers for this appraisal.

2.2 Background

2.2.1 Background information on ER+/HER2- advanced breast cancer with an ESR1 mutation

The CS considers advanced / metastatic breast cancer to encompass people with unresectable (inoperable) Stage III locally advanced breast cancer and Stage IV metastatic breast cancer (mBC). Approximately 35% of people with early or locally advanced resectable breast cancer will progress to mBC within 10 years of diagnosis and approximately 13% of people with breast cancer will have advanced/mBC at diagnosis. Of the various histopathological subtypes of breast cancer (determined by oestrogen receptor (ER) and/or progesterone receptor (PR) and human epidermal factor receptor (HER2) status) the most common is ER+/HER2-, accounting for approximately 70% of cases. Survival rates at 5 years are 36%, reducing with each successive line of therapy.

The CS mentions that patients with ER+/HER2 breast cancer receiving endocrine therapy (ET) over time are at risk of acquired resistance, including acquired mutations in the ESR1 (Oestrogen receptor 1) gene, known as the ESR1 mutation or *ESR1-mut*. Acquisition of this mutation happens almost exclusively after treatment with an aromatase inhibitor (AI) and is more common with longer exposure to ET. It is stated that the prevalence of the ESR1-mutation is higher in those treated with an AI plus a CDK4/6 inhibitor compared to AI alone. The CS estimates that up to 50% of patients who have received an AI will develop the ESR1-mutation on disease progression, thus creating a "novel population" of ER+/HER2- ESR1-mutated advanced/metastatic breast cancer. Importantly, this population experiences faster disease progression and poorer survival than those without

an ESR1-mutation. The evidence cited in support of this claim comes from the company's analysis of studies of endocrine therapy in advanced hormone receptive breast cancer, including the BOLERO-2 trial,[1] the BYLieve trial,[2] Clatot et al (2016),[3] the MAINTAIN trial,[4] and pooled analysis of the SoFEA and EFFECT trials.[5]

Some patients with ER+/HER2- develop the PIK3CA mutation (Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha) and some have both the PIK3CA mutation and the ESR1-mutation. The latter group are referred to in the CS as the "dual mutation" group and are eligible for elacestrant according to the marketing authorisation.

2.2.2 Background information on elacestrant

The CS describes elacestrant as a next-generation, nonsteroidal, orally bioavailable SERD (selective oestrogen receptor degrader). It received its marketing authorisation in the UK in December 2023 from the Medicines and Healthcare products Regulatory Agency (MHRA), and is indicated for *"the treatment of postmenopausal women, and men, with ER+/HER2-, locally advanced or mBC with an activating ESR1-mutation who have disease progression following at least one line of ET including a CDK4/6i."* (CS page 14, reproduced from the Summary of Product Characteristics).

Elacestrant is administered as an oral tablet (345 mg) once daily for as long as clinical benefit is observed or until unacceptable toxicity occurs. Dose modifications are permitted depending on adverse reactions, as detailed in the Summary of product characteristics (SmPC).

Elacestrant is described as the first targeted treatment option specifically indicated for patients with ER+/HER2- ESR1-mutated advanced/mBC. The CS states that patients with ER+/HER2- advanced breast cancer should be selected for treatment with elacestrant based on the presence of an activating ESR1-mutation in plasma specimens, using a CE-marked in vitro diagnostic (IVD) with the corresponding intended purpose. However, the company notes that that genomic testing for the ESR1-mutation is not currently funded as standard practice in the UK. They anticipate that testing will be funded in the future with the introduction of elacestrant treatment. For the purposes of this NICE appraisal the company has included ESR1-mutation testing using liquid biopsy, based on polymerase chain reaction (PCR) testing (see section 4.2.6.5 of this report for a discussion of how testing is modelled in the economic evaluation).

Expert clinical advice to the EAG suggests that ESR1 testing is currently not widely available in the NHS, and that the introduction of testing would not likely introduce delays

to the clinical management of patients being considered for elacestrant therapy. Test turnaround times would likely be in-keeping with current commercial testing timelines. See section 4.2.6.5 for further discussion.

2.2.3 The current care pathway for advanced/metastatic ER+/HER2- breast cancer

The CS describes the current treatment pathway and where in the pathway the company suggests elacestrant would be of most benefit. They draw on recommendations from relevant clinical guidelines, notably the European Society for Medical Oncology (ESMO) Guideline for mBC and the ESMO mBC Living Guideline for patients with ER+/HER2- mBC. Recommendations from previous NICE appraisals of treatments for ER+/HER2- advanced/mBC are also mentioned, as well as NICE clinical guideline CG81 'Advanced breast cancer: diagnosis and treatment' and NG101 'Early and locally advanced breast cancer: diagnosis and management'.

2.2.3.1 First line therapy for advanced/metastatic breast cancer

Figure 1 reproduces the company's illustration of the treatment pathway (CS Figure 6) in the advanced/mBC setting. As can be seen, patients can receive successive lines of therapy as their cancer progresses. First line treatment is endocrine therapy (e.g. an aromatase inhibitor such as **anastrozole or letrozole**) combined with a CDK4/6 inhibitor (e.g. **palbociclib, ribociclib, abemaciclib**). Chemotherapy may be given if imminent organ failure is suspected.

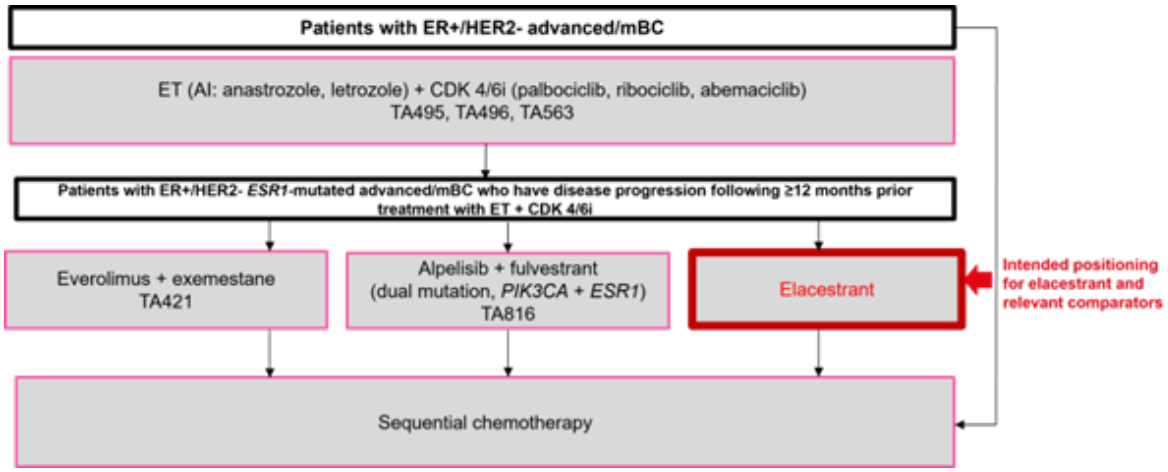

Figure 1 Current treatment pathway in England and Wales for patients with ER+/HER2- advanced/mBC

Source: Reproduced from CS Figure 6

The EAG notes that the pathway in Figure 1 doesn't distinguish between previously treated (adjuvant relapsed) patients and untreated patients with de novo advanced//metastatic disease.

Expert clinical advice to the EAG is that If relapse occurs whilst on an aromatase inhibitor, or less than 12 months after stopping, this is likely to indicate intrinsic resistance. Therefore, re-treatment with a drug sharing the same mechanism of action would be ineffective. Relapse more than 12 months after finishing treatment is more likely to be due to acquired resistance resulting in upregulation of the CDK pathways, which can be overcome by combining an aromatase inhibitor with a CDK4/6 inhibitor.

Patients with de novo advanced/metastatic breast cancer who have not been exposed to any previous hormonal therapy generally would be treated with a combination of an **aromatase inhibitor and a CDK4/6 inhibitor** (if premenopausal, they would also need to have ovarian suppression, usually with goserelin or a similar drug).

Our expert also commented that clinicians will soon start to see patients who are relapsing having already had a CDK4/6 inhibitor in the adjuvant setting. These patients would switch to an alternative hormone therapy (**aromatase inhibitor) or tamoxifen**. Some patients might also receive fulvestrant (depending on local funding agreements), or fulvestrant in combination with alpelisib (if PIK3CA mutated tumour, or exemestane + everolimus.

2.2.3.2 Second line therapy for advanced/metastatic breast cancer

Until elacestrant was licensed there were no available ESR1 mutation-targeted treatments and, hence, genomic testing for this mutation is not included in the current pathway. Instead, the CS states that advanced/metastatic breast cancer patients with an ESR1-mutation progressing from first line treatment are currently "managed empirically", with non-targeted medicines. It is unclear to the EAG what the company means by managed empirically, but we assume the choice of second line treatment is based on an assessment of signs, symptoms and prognostic factors (e.g. performance status) collectively indicating the aggressiveness of the tumour, the likely rate of progression and the fitness of the patient to undergo further treatment.

The CS identifies a subgroup of patients with ER+/HER2- ESR1-mutated advanced/ metastatic breast cancer who have disease progression following ≥12 months prior treatment with ET + CDK4/6i. This is the group the company propose should be offered elacestrant, as reflected in their decision problem and submission to NICE (see section 2.3 below for a discussion of the decision problem). The EAG notes that this is a narrower

population than that covered by the marketing authorisation - the latter does not stipulate a minimum duration of prior treatment with ET + CDK4/6i (≥12 months) before elacestrant can be given. We discuss the clinical rationale for this subgroup in section 2.3 below.

The EMSO metastatic breast cancer living guideline[6] for patients with ER+/HER2- metastatic breast cancer lists a number of treatment options for patients with ER+/HER2- advanced/metastatic breast cancer (CS Table 4). The guideline states that the optimal sequence of endocrine therapy after progression with an ET + CDK4/6i depends on factors such as which hormonal treatments the patient used previously, the duration of their response to prior treatment, tumour mutational status, disease burden and patient preference. Of the treatment options listed (excluding elacestrant itself which the EMSO guideline recommends for patients with an ESR1-mutation) the company considers two existing treatments as relevant for patients with ER+/HER2- ESR1-mutated advanced/ metastatic breast cancer who have disease progression following ≥12 months prior treatment with ET + CDK4/6i. These are:

- **everolimus and exemestane** (as recommended in NICE TA421) [7] and
- **alpelisib and fulvestrant** (as recommended for patients with the PIK3CA mutation in NICE TA816). [8]

As elacestrant is intended for use as a second line therapy these two dual therapies are relevant comparators for this appraisal (see section 2.3 for further detail on comparators).

The remaining second line treatments listed in the EMSO guideline are: **everolimus + fulvestrant** (preferred over everolimus + exemestane if the patient is ESR1-mutation positive); **switching ET ± CDK4/6i** or **fulvestrant monotherapy**; and **chemotherapy** for patients at imminent risk of organ failure. According to expert clinical opinion sought by the company, endocrine monotherapy, and endocrine therapy with chemotherapy, are rarely used in practice in the patient population under consideration in the CS (i.e. people with ER+/HER2- ESR1-mutated advanced/metastatic breast cancer who have disease progression following ≥12 months prior treatment with ET + CDK4/6i). The EAG's expert clinical adviser agrees.

The EAG notes that the CS does not comment on the EMSO guideline recommendation (CS Table 4) that **everolimus and fulvestrant** is preferred over **everolimus and exemestane** for treating ESR1 mutated tumours. However, expert clinical advice to the EAG is that everolimus and fulvestrant are not funded by the NHS.

Expert clinical advice to the EAG is that patients previously treated in the adjuvant setting who progress after first line treatment in the advanced/metastatic breast cancer setting would switch to

- A different aromatase inhibitor (usually from non-steroidal to steroidal) with or without everolimus,
- Or switch to tamoxifen.
- Or switch to fulvestrant + alpelisib if they have a PIK3CA mutated tumour (provided that they have not already received fulvestrant in combination with a CDK4/6 inhibitor).

Patients with de novo advanced/metastatic breast cancer who progress after first line treatment in the advanced/metastatic breast cancer setting would also switch to a different aromatase inhibitor or to tamoxifen. Patients with the PIK3CA mutation would switch to **alpelisib and fulvestrant** in combination. The expert commented that, contra to the EMSO guideline, fulvestrant monotherapy would not be used as it is not recommended by NICE (TA239).

2.2.3.3 Third line treatment for advanced/metastatic breast cancer

The CS does not comment on treatment options for patients who progress from second line treatment, other than noting that sequential chemotherapy is recommended by the EMSO guideline (CS Figure 4). The EAG's clinical expert advisor commented that factors taken into account when considering third line therapy include the patient's clinical condition, the extent of metastases, which sites are affected, the rate of disease progression and also their treatment history. Patients with hormone responsive cancer who progress on second line therapy might switch to third line hormone therapy, with whichever drugs they haven't already received. The expert also noted that many patients have slow progressing disease and are candidates for third line treatment.

2.2.4 Justification for the position of elacestrant in the treatment pathway

As described above, the company proposes elacestrant as a treatment for ER+/HER2- ESR1-mutated advanced/metastatic breast cancer who have disease progression following ≥12 months prior treatment with ET + CDK4/6i. The CS notes that since the introduction of ET+ CDK4/6i, there has been a rise in the prevalence of ESR1 mutations associated with prolonged duration of treatment. The CS notes that current standard treatments, such as the combination of everolimus and exemestane or alpelisib and fulvestrant, have not been evaluated in patients with ER+/HER2- ESR1-mutated advanced/metastatic breast cancer who have disease progression following ≥12 months of prior treatment with ET + CDK4/6i. Furthermore, the CS points out some of the limitations of current standard treatments,

citing significant toxicity (everolimus, alpelisib) and the pain and inconvenience of attending clinic to receive fulvestrant injections. The CS contends that there is increasing unmet need for a treatment specifically tailored for patients with the ESR1-mutation, with an acceptable safety profile and which can be taken orally rather than injected intramuscularly. This would be more convenient for patients and their carers and would require fewer healthcare resources to manage.

The EAG's expert clinical advisor commented that clinicians would view elacestrant as an oral drug that works in a similar way to fulvestrant, which has to be given by intramuscular injection. In the longer term it would be preferable for patients to have an oral alternative to fulvestrant. Fulvestrant is mostly used in combination with other drugs, however, there is currently no available evidence on the efficacy and safety of elacestrant in combination therapy.

EAG comment on the background information

The background section of the CS provides detailed information about the epidemiology of breast cancer, the course of disease and its subtypes, and the impact on morbidity and mortality. The anticipated place of elacestrant in the current treatment pathway is clearly defined, though the overall pathway depicted doesn't explicitly acknowledge that the choice of treatments for advanced/metastatic breast cancer will depend on the patient's previous treatment history, and may require switching to different hormone treatments at each successive line.

2.3 Critique of the company's decision problem

Table 4 summarises the NICE scope for this appraisal, the company's decision problem, and the EAG's critique of the company's approach. As the table shows, the decision problem adheres to the NICE scope, albeit with two notable exceptions: the patient population and the choice of comparator treatments.

Table 4 Summary of the decision problem

	Final scope issued by NICE	Company's decision problem	Rationale if different from the final NICE scope	EAG comments
Population	People who have been through menopause and men with ER+/HER2- locally advanced or mBC with an activating ESR1-mut after at least 1 line of ET including a CDK4/6i.	Postmenopausal women, and men, with ER+/HER2-, locally advanced/mBC with an activating ESR1-mut who have disease progression following ≥12 months prior treatment with ET + CDK4/6i	This is the population of patients where clinicians perceive the most value for elacestrant to be in UK clinical practice. In a post hoc subgroup analysis of the pivotal phase III study (EMERALD), patients treated with elacestrant had a greater improvement in PFS with longer exposure (≥12 months) to prior ET +	The company clarified the rationale for ≥12 months of prior ET + CDK4/6i (as opposed to other potential thresholds for prior treatment). They presented a post-hoc subgroup analysis of the EMERALD trial at an international cancer conference in 2022.[9 10] Longer duration on CDK4/6i was associated with improvement in PFS for patients treated with elacestrant, and this was more pronounced in

	Final scope issued by NICE	Company's decision problem	Rationale if different from the final NICE scope	EAG comments
			CDK4/6i vs. ET monotherapy. The results of this post hoc subgroup analysis support the beneficial activity of elacestrant in patients with longer exposure (i.e. ≥12 months) to prior ET + CDK4/6i.	patients with at least 12 months of prior CDK4/6i duration. The EAG notes that these subgroups (i.e. <6 months, 6-12 months, 12-18 months, ≥18 months) were selected post hoc after examination of the data. Whilst the results indicate greater PFS according to length of previous treatment, these findings are exploratory, and not confirmatory. The EAG also notes there is a similar pattern in the results of the All-patient population.
Intervention	Elacestrant	Elacestrant	Not applicable	No comment
Comparators	Everolimus + exemestane; ET with or without chemotherapy; the	Everolimus + exemestane;	UK clinical expert opinion suggests that:	Expert advice to the EAG confirms that endocrine

	Final scope issued by NICE	Company's decision problem	Rationale if different from the final NICE scope	EAG comments
	Chemotherapy; Alpelisib + fulvestrant (for people whose BC is PIK3CA-mutated)	Alpelisib + fulvestrant (for people whose BC is PIK3CA-mutated)	ET monotherapy or ET + chemotherapy is rarely used in clinical practice in England and Wales in the patient population under consideration in this submission. Chemotherapy in the UK is reserved predominantly for patients with imminent risk of organ failure	monotherapy is not standard practice in the NHS.
Outcomes	OS PFS Response rate Adverse effects of treatment HRQoL	OS PFS Response rate Adverse effects of treatment HRQoL	Not applicable	No comment
Economic analysis	The reference case should be followed. The economic modelling should include the costs	Not stated	Not stated	The company do not refer to the economic analysis in the decision problem. However, as discussed in section 4.2.1 of this

	Final scope issued by NICE	Company's decision problem	Rationale if different from the final NICE scope	EAG comments
	associated with diagnostic testing for ESR1 and where relevant, PIK3CA mutations in people with oestrogen receptor-positive HER2 negative locally advanced or metastatic breast cancer who would not otherwise have been tested. A sensitivity analysis should be provided without the cost of the diagnostic test.			report, the economic model complies with the reference case, and the cost of ESR1 testing is included in the model, (and removed in a sensitivity analysis).
Subgroups	Mutations in both ESR1 and PIK3CA	Mutations in both ESR1 and PIK3CA	For the dual mutated population only those patients progressing following ≥12 months prior treatment with ET + CDK4/6i are considered.	See comment above in Population

Source: Reproduced in part from CS Table 1

3 CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

In CS Appendix D the company describe their systematic literature review (SLR) to identify clinical evidence (RCT and non-RCT) for elacestrant and comparators (everolimus + exemestane and alpelisib + fulvestrant) for ER+/HER2- ESR1-mutated advanced/metastatic breast cancer. The EAG ‘s appraisal of the company’s systematic review methods is summarised in Appendix 1. Briefly, the company carried out an initial SLR, referred to in the CS as “the global clinical SLR”, which had broader eligibility criteria for interventions and comparators than the NICE final scope (CS Appendix D Table 4). To identify relevant evidence for the appraisal, the company then used narrower eligibility criteria aligned with the NICE final scope (CS Appendix D Table 5), to rescreen included studies identified from the initial SLR. The EAG considers these narrower eligibility criteria appropriate in terms of the appraisal.

The EAG did, however, note two potential issues with the company’s searches which may result in relevant evidence being missed. First, the searches were approximately eight months old when the CS was received by the EAG. Second, the RCT filter used in the searches excluded conference abstracts. The EAG therefore reran the company’s searches for the last 8 months and, separately, the Embase search for the past three years using terms that would include conference abstracts. After deduplication, these EAG searches yielded a total of 217 records. The EAG screened all 217 titles and abstracts, and subsequent eight full papers, against the eligibility criteria aligned to the NICE final scope (CS Appendix D Table 5). None of these full papers were relevant to the NICE final scope. Overall, the EAG believe the company’s review is comprehensive and matches the decision problem.

3.2 Critique of studies of the technology of interest, the company’s analysis and interpretation (and any standard meta-analyses of these)

3.2.1 Included studies

The initial broader SLR identified 23 publications (CS Appendix D Figure 1). On rescreening these 23 publications against the narrower SLR eligibility criteria, which was aligned with the NICE final scope (CS Appendix D Table 5), 13 publications were subsequently excluded because the intervention was not relevant to the scope of this technology appraisal (CS Appendix D Figure 1). The company reports 10 publications were therefore relevant to the

NICE final scope (CS Appendix D.2, CS Appendix D Figure 1, CS Appendix D Table 6). Of these 10 publications:

- Seven publications concerned one RCT, the **EMERALD** trial, of the efficacy and safety of elacestrant versus clinician's choice of fulvestrant, anastrozole, letrozole, or exemestane monotherapy in postmenopausal women and men with ER+/HER2-, advanced or metastatic breast cancer, whose disease has relapsed or progressed on at least one and no more than two lines of prior ET for advanced or metastatic breast cancer, which must have included a CDK4/6i in combination with fulvestrant or an AI. A subgroup of these patients had an activating ESR1-mutation (ESR1-mut). Key results from the trial are presented in an article in the Journal of Clinical Oncology. [11]

- Three publications concerned two studies of alpelisib in combination with fulvestrant.
  - One non-RCT (BYLieve; NCT03056755; 2 publications)[2 12]
  - One retrospective real world cohort study (one publication)[13]
- The company reports that no evidence was identified for everolimus in combination with exemestane in the population defined in the company decision problem i.e. ESR1-mut and ≥12 months' prior ET including a CDK4/6i (CS section B.2.1).

CS section B.2.2 only lists the EMERALD RCT as the relevant clinical effectiveness evidence for the appraisal and CS document B section B.2.11 states that there are no other ongoing studies of elacestrant. At the EAG's request the company provided a detailed list of all elacestrant phase 1, phase 2 and phase 3 clinical trials (Company clarification response A3). After assessing this list, the EAG agree that the EMERALD trial is the only relevant trial of elacestrant for this appraisal.

3.2.1.1 Study characteristics

The **EMERALD** study (study RAD1901-308; ClinicalTrials.gov number NCT03778931)[11] is an ongoing phase III, multicentre, randomised, open-label, active controlled trial comparing the efficacy and safety of elacestrant to endocrine monotherapy treatment (investigator's choice of fulvestrant or an aromatase inhibitor) in postmenopausal women, or men, with ER-positive/HER2-negative advanced/metastatic breast cancer. The primary outcome of the trial was progression free survival (PFS) based on blinded imaging review committee (IRC)-assessment in either all patients (i.e. with ESR1 mutations (*ESR1-mut*) or without detectable ESR1 mutations *(ESR1-mut*-nd)) or in patients with ESR1 mutations only (CS B.2.3.1, B.2.11). Patients were enrolled from 17 countries, including the UK. Fifty four percent of patients were enrolled from Europe and 29.5% from North America. The trial results support

the company's regulatory marketing authorisation for elacestrant. Evidence from the trial also inform the assessments of cost-effectiveness in the company's economic model (CS B.2.2; see sections 4.2.4, 4.2.5.2 and 4.2.5.3 of this this report). The EAG note that the populations addressed in the company's submission, i.e. ESR1-mut only, or dual mutated (mutations in ESR1 and PIK3C), who have disease progression following ≥12 months prior treatment with ET + CDK4/6 inhibitors, are post-hoc specified subgroups (henceforth referred to in this report as "post-hoc subgroup 1 (ESR1 mutation)" and "post-hoc subgroup 2 (dual mutation)" respectively). Post-hoc subgroup 2 (dual mutation) itself is a subgroup nested within post-hoc subgroup 1 (ESR1 mutation). Table 5, below, summarises the EMERALD trial methodology.

Table 5 Summary of EMERALD trial methodology

Study characteristics	
Trial design	RCT Open label 2 arm - elacestrant versus standard of care (SOC) (investigator's choice of fulvestrant, anastrozole, letrozole, or exemestane monotherapy)
Randomisation	1:1 Stratified by ESR1-mut status (ESR1-mut vs. ESR1-mut not detected), prior treatment with fulvestrant (yes or no) or presence of visceral metastasis (yes or no) n=478 patients enrolled (including 12 from UK), of which 228 were ESR1-mut (including 9 from UK)
Evaluation of ESR1-mutational status	Evaluated in cell-free circulating DNA at a central laboratory; blood samples were analysed using the Guardant360 CDx (GuardantHealth, RedwoodCity, CA). ESR1 mutations defined as any missense mutation in codons 310 - 547. ESR1 mutation status was not provided to study sites during treatment.
Study duration	10/05/2019 – 08/2024 (estimated); no further data cuts expected. The company provided a CSR, along with its associated protocol, SAP and addendum. CSR v.2 reports trial results from a data cut of **06 September 2021** for the whole trial population and ESR1-mut population. This data cut includes the primary analysis of the primary outcome (blinded-IRC assessed PFS) and interim results

Study characteristics	
	of OS. The main findings of the trial, with the same data cut, were published in the Journal of Clinical Oncology (Bidard et al, 2022).[11] An overall survival addendum to CSR v.2, with a data cut of **02 September 2022**, reports the final OS analyses. For *post-hoc* subgroups [subgroups 1 and 2] the data cut was **02 September 2022** for PFS and OS, and **8th July 2022** for patient-reported outcome (PRO) data.
Location	Europe (Austria, Belgium, Denmark, France, Greece, Hungary, Ireland, Italy, Portugal, Spain, UK), Asia (Israel, South Korea), North America (Canada, United States), Other (Argentina, Australia).
Included population	Postmenopausal women, or men, aged ≥ 18 years with ER-positive/HER2-negative advanced or metastatic breast cancer who have progressed or relapsed following one to two prior lines of ET for advanced or metastatic disease, one of which was given in combination with a CDK4/6i. Patients must have received no more than one line of cytotoxic chemotherapy for metastatic breast cancer and had an ECOG PS of 0 or 1.
Excluded population	Patients with symptomatic metastatic visceral disease or any of the following cardiovascular events within 6 months of enrolment: severe/unstable angina, myocardial infarction, coronary/peripheral artery bypass graft, prolonged corrected QT interval grade ≥ 2, uncontrolled atrial fibrillation, ongoing grade ≥ 2 cardiac dysrhythmias, New York Heart Association Class II or greater heart failure, coagulopathy (thrombosis), cerebrovascular accident and in the UK patients were excluded if they had a QTcF of ≥450 msec.
***Post-hoc* specified subgroups of relevance to the submission**	**Post-hoc subgroup 1 (ESR1 mutation):** ESR1-mut who have received ≥12 months of prior ET + CDK4/6i **Post-hoc subgroup 2 (dual mutation):** Mutations in both ESR1 *and* PIK3CA (dual mutated) who have received ≥12 months of prior ET + CDK4/6i
Intervention	Elacestrant dihydrochloride 400 mg/day (equivalent to elacestrant 345 mg), once-daily orally. Protocol-defined dose reductions permitted to 300 mg or 200 mg daily.

Study characteristics	
Comparator	Investigator's choice of one of the following monotherapies [a]: **Fulvestrant:** 500 mg intramuscularly on cycle 1[b] day 1, cycle 1 day 15, cycle 2 day 1 and day 1 of every subsequent 28-day cycle **Anastrozole:** 1 mg/day orally on a continuous dosing schedule **Letrozole:** 2.5 mg/day orally on a continuous dosing schedule **Exemestane**: 25 mg/day orally on a continuous dosing schedule
Primary outcome	PFS based on blinded -IRC-assessment in i) all patients (i.e. with or without detectable ESR1 mutations) or ii) in patients with ESR1 mutations only.
Secondary outcomes informing the economic model	Overall survival, EQ-5D-5L, adverse events
Other secondary outcomes	**Efficacy:** Response rate (Blinded IRC assessed ORR, DOR and CBR) **HRQoL:** EQ-VAS score, EORTC QLQ-C30, PRO-CTCAE **Other:** time to chemotherapy **Safety:** treatment compliance and exposure, treatment emergent adverse events, deaths and serious adverse events.

Source: Partly reproduced from CS document B Table 6 and Table 7
Abbreviations: CBR, clinical benefit rate; CDK4/6i, cyclin-dependent kinase 4/6 inhibitor; CSR, clinical study report; DOR, duration of response; ECOG, Eastern Cooperative Oncology Group; EORTC QLQ-C30, European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; EQ-5D-5L, EuroQoL Five-dimension Five-level; ER, oestrogen receptor; ESR1, oestrogen receptor 1 gene; ET, endocrine therapy; HER2, human epidermal growth factor receptor 2; HRQoL, health-related quality of life; IRC, imaging review committee; mut, mutation; ORR, objective response rate; OS, overall survival; PFS, progression-free survival; PIK3CA, phosphatidylinositol 3 kinase; PRO, patient-reported outcome; PRO-CTCAE, Patient-Reported Outcome Common Terminology Criteria for Adverse Events; RCT, Randomised Controlled Trial; RECIST, response evaluation criteria in solid tumours; SAP, statistical analysis plan; SOC, standard of care; UK, United Kingdom
[a] No other anti-cancer agents were allowed
[b] 28 day cycle
[c] Common Terminology Criteria for Adverse Events criteria

The EAG considers there are two issues regarding the design of the EMERALD trial in relation to this appraisal:

1. the choice of comparators and
2. the type of test used to assess ESR1 mutational status.

These are discussed in further detail below.

Comparators

As shown in Table 5 above, comparators used in the EMERALD trial were investigator's choice of one of the following monotherapies: fulvestrant, anastrozole, letrozole or exemestane. Clinical expert advice to the company were that the use of monotherapy after progression on CDK4/6i is not representative of standard clinical practice.[14] The EAG clinical expert agreed. Additional issues regarding comparators the EAG clinical expert highlighted were:

- Fulvestrant is not allowed to be used as a single agent in clinical practice due to NICE guidelines (TA239).[15]
- Some patients in the EMERALD trial comparator arm had prior exposure to a non-steroidal aromatase inhibitor and were assigned to receive another in the trial. Switching from one drug to another that works in the same way is rarely done in clinical practice as the likelihood of overcoming resistance would be expected to be very low. The company state that while patients in the EMERALD trial could also receive a steroidal aromatase inhibitor following a non-steroidal one and vice versa, this was not the preferred option. A few patients received several lines of therapy and may have received a similar AI in one of these prior lines, but not in the line directly prior to starting the trial.
- The lack of tamoxifen as a comparator choice is perplexing given that most patients in the EMERALD trial had no prior exposure to tamoxifen (approximately 8% in each arm of the ESR1-mut subgroup received tamoxifen as prior therapy; CS document B Table 9).

Test to evaluate ESR1-mutational status

The EAG clinical expert believed that the proposed test to identify ESR1-mutation status in the NHS is not the same, and has disadvantages, compared to the test used in the EMERALD trial.

The proposed test for the NHS would utilise a tissue sample, either a primary tumour sample, which is limited due to being a historic sample, or a single site repeat biopsy, which is limited by the potential to not fully reflect disease status due to within tumour heterogeneity. Conversely, the ESR-1 mutation status testing in the EMERALD trial is tissue free, using a current blood sample for circulating tumour DNA analysis (Emerald protocol section 7.6.2). It is therefore an assessment of the current tumour *and* is more likely to assess the totality of the tumour rather than that of an individual sample site.

3.2.1.2 Patients' baseline characteristics

The CS presents baseline characteristics for the following EMERALD trial populations only: all patients with ESR1-mut (CS B.2.3.1.2 and CS document B Table 9), and the post-hoc-subgroups 1 (ESR1 mutation) and 2 (dual mutation; CS B.2.7.1 and CS document B Table 20).

The CS states baseline characteristics for both post-hoc subgroups were similar to those of all patients with ESR1-mut (CS section B.2.7.1). Briefly, the median age of participants was approximately 63 years and all were female. In terms of race/ethnicity, most participants (approximately 75%) identified themselves as White. Approximately half of patients had ECOG performance 0 (indicating the participant is fully active with no performance restrictions) and the other half ECOG performance 1 (cannot do strenuous physical activity but is fully ambulatory and can do light work). The proportion of patients with visceral metastases (including lung, liver, brain, pleural, and peritoneal involvement) was approximately 75%. Over half of participants had received prior adjuvant therapy. In terms of prior treatment for advanced or metastatic disease, all participants had received prior CDK4/6i therapy and over 96% received prior ET with the remaining patients progressing during or within 12 months of adjuvant endocrine therapy. In the advanced or metastatic setting, approximately two-thirds of participants had one prior line of ET and one-third had two lines of prior endocrine therapy. In terms of experience with chemotherapy, approximately three-quarters of patients had no prior lines of chemotherapy and one-quarter had one-line of prior chemotherapy.

The CS states that baseline characteristics for all patients with ESR1-mut, and for both post-hoc subgroups, were well balanced between the two study arms (CS B.2.3.1.2, CS B.2.7.1). While the EAG in general agree with the company's statement, we note the following imbalances/differences with respect to the post-hoc subgroups (CS document B Table 20):

- Post-hoc subgroup 1 (ESR1 mutation):
  - A [redacted] proportion of the elacestrant arm received fulvestrant as prior therapy for advanced or metastatic disease compared to the SOC arm ([redacted])
  - A [redacted] proportion of the elacestrant arm received mammalian target of rapamycin (mTOR) inhibitor as prior therapy for advanced or metastatic disease compared to the SOC arm ([redacted])

- Post-hoc subgroup 2 (dual mutation):

- Median age was slightly [redacted] in the elacestrant arm than in the SOC arm ([redacted]).
- A [redacted] proportion of participants in the elacestrant arm has visceral metastasis (including lung, liver, brain, pleural, and peritoneal involvement) compared to the SOC arm ([redacted])
- A [redacted] proportion of the elacestrant arm received mTOR inhibitor as prior therapy for advanced or metastatic disease compared to the SOC arm ([redacted])
- In the advanced or metastatic setting a [redacted] proportion of the elacestrant arm received one prior line of endocrine therapy compared to the SOC arm ([redacted]), and a [redacted] proportion of the elacestrant arm received two prior lines of endocrine therapy compared to the SOC arm ([redacted]).

The above baseline characteristics indicate that patients in the elacestrant arm of post-hoc subgroup 2 (dual mutation) were [redacted] compared to patients in the SOC arm. The impact of these imbalances is unclear.

EAG comment on included studies

The EMERALD trial is a large ongoing phase III, multicentre, randomised, open-label, active controlled trial of the safety and efficacy of elacestrant. It was used as the source of evidence in the granting of the marketing authorisation and is the sole source of evidence on elacestrant to inform this NICE appraisal. The trial included a pre-specified subgroup of participants with the ESR1 mutation, comprising almost half of the randomised trial population (228/478 participants, 48%). One of the main limitations of the EMERALD trial is that the comparator arm (investigators choice of standard of care endocrine monotherapies), and therefore the elacestrant treatment comparison, is of limited relevance to the scope and the decision problem for this NICE appraisal.

3.2.2 Risk of bias assessment

The company's methodological quality assessment (also referred to as risk of bias assessment) of the EMERALD trial was conducted using the Centre for Reviews and Dissemination (CRD) guidance for undertaking reviews in healthcare.[16] An overview of the company's assessment is presented in CS document B Table 12 and their full assessment, which includes justification for their judgements, is presented in CS Appendix D Table 7. The EAG independently critically appraised the trial using the same criteria, and an overview of

our judgements, alongside those of the company, are presented below in Table 6 (disagreements between the company and EAG judgements are in bold and are discussed the text below the table).

Table 6 Overview of company and EAG risk of bias judgements

Criterion	Company judgement	EAG judgement
Was randomisation carried out appropriately?	Yes	Yes
Was the concealment of treatment allocation adequate?	No	Yes
Were the groups similar at the outset of the trial in terms of prognostic factors?	Yes	Yes
Were the care providers, patients and outcome assessors blind to treatment allocation?	No	No, with exception of blinded-IRC assessments, which includes primary analysis of PFS
Were there any unexpected imbalances in dropouts between groups?	No	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No
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 for all outcomes except for missing data for EQ-5D-5L presented in the CS (Note, the economic model uses all the EQ-5D data collected, as per preferred NICE methodology)[a]

Source: Partly reproduced from CS document B Table 12 and CS Appendix D Table 7. Additional sources: CS B 2.3.1, CS document B figure 3, CS Appendix D figure 2, CSR sections 9.4.4 and 9.4.6, CSR Tables 14.1.4.1 and 14.1.5.1
Abbreviations: EQ-5D-5L, EuroQoL Five-dimension Five-level; IRC, imaging review committee; PFS, progression-free survival

[a] Company clarification response A5

The EAG agreed with the company's judgements for all criteria except the following:

Concealment of allocation

The company judged the concealment of allocation was inadequate due to the trial being open-label and therefore patients and investigators were not blind to treatment assignment. The EAG suggest that the company is confusing allocation concealment with blinding. Allocation concealment is performed when the treatment allocation system is set up so that the person enrolling participants does not know in advance which treatment the next person will get. CS Appendix D Table 7 and CSR section 9.4.4 describe randomisation being conducted by Interactive Randomization Technology (IRT), which provided the randomisation number and treatment assignment.[17] The EAG therefore consider that allocation concealment was adequate.

Blinding of care providers, patients and outcome assessors to treatment allocation

The company judged that as the trial was open-label, patients and investigators were not blind to treatment assignment. The EAG agree that patients and caregivers were not blind, therefore patient reported outcomes and safety-related outcomes could be subject to bias. However, response and progression, including the primary analysis of PFS included in the CS, were assessed by a blinded IRC. The risk of outcome assessment related bias for these outcomes is therefore unlikely. Furthermore, the key secondary outcome of overall survival was an objective outcome and therefore unlikely to be influenced by knowledge of the treatment received.

Missing data

There is considerable missing data for EQ-5D-5L index scores for the ESR1-mut subgroup (CS B.2.6.4). First, the company's decision to obtain EQ-5D-5L index scores only for countries in which the validated tool was available (5 out of 17 countries enrolled in the trial; see company clarification response A5) resulted in large differences in the number of patients in each arm of the ESR1-mut subgroup with an EQ-5D-5L index score versus an EQ-VAS score (50 (43%) versus 108 (94%) in the elacestrant arm and 50 (44%) versus 98 (87%) in the SOC arm). The company clarified that this issue is in relation to EQ-5D-5L index scores presented in the CS but that the economic model uses all the EQ-5D data collected.

Second, there is a difference in the total number of patients with ESR1-mut enrolled in [redacted]; CSR Table 14.1.1.2) of the [redacted] countries and those that had a baseline EQ-5D-5L score (CSR Table 14.2.6.4.1). In total [redacted] ESR1-mut patients were enrolled from these [redacted] countries, with [redacted] assigned to elacestrant and [redacted] to SOC, yet baseline EQ-5D-5L index scores are only available for [redacted] patients in each arm (CSR Table 14.1.1.2 and CSR Table 14.2.6.4.1). It is unclear to the EAG why there is this discrepancy.

3.2.3 Outcomes assessment

All outcomes included in the NICE scope (OS, PFS, response rate, adverse effects of treatment and HRQoL) were measured in the EMERALD trial.[18] CS document B, CS Appendix E, and company clarification response A9 present results of these outcomes for all patients with ESR1-mut, and for the two post-hoc subgroups. Results for the whole EMERALD trial population i.e. with or without ESR1-mut, were reported in the main trial publication (Bidard et al., 2022) [11] and in the CSR provided by the company.[11 17] Table 7 provides a summary of the NICE scope and decision problem related outcomes reported in the EMERALD trial.

Table 7 List of NICE scope and decision problem related outcomes reported in the EMERALD trial

Endpoint	Outcome	Definition
Primary	Blinded IRC-assessed progression free survival (PFS)	Length of time from randomisation until the date of objective disease progression per RECIST version 1.1 or death from any cause
Key secondary	Overall survival (OS)	Length of time from randomisation until the date of death from any cause
Other secondary	Blinded IRC-assessed objective response rate (ORR)	Percentage of patients with measurable disease who had achieved either a confirmed CR or PR per RECIST v1.1
	Blinded IRC-assessed clinical benefit rate (CBR)	Percentage of patients who had achieved either a confirmed CR or PR or stable disease at ≥24

Endpoint	Outcome	Definition
		weeks from randomisation per RECIST v 1.1
	Blinded IRC-assessed duration of response (DOR)	Duration of time from the date when criteria are met for either a CR or PR (whichever is first recorded) per RECIST v1.1 until the first date that recurrent or PD is objectively documented, or death from any cause
	Safety and tolerability	AEs: deemed treatment related if they occurred after the first dose of study drug and ≤30 days after the last dose of study drug SAEs led to death, hospitalisation, or prolonged hospitalisation, persistent or significant incapacity or disruption to normal daily life, congenital anomaly/birth defect, were life-threatening or required intervention to avoid one of the above Dose modifications Clinical laboratory parameters, ECGs, ECOG performance status, and vital signs
	Patient reported outcomes (PROs) and health related quality of life (HRQoL)	EQ-5D-5L, EORTC QLQ-C30 and PRO-CTCAE

Source: Partly reproduced from CS document B Table 8
AE, adverse event; CBR, clinical benefit rate; CR, complete response; DOR, duration of response; ECG, electrocardiogram; ECOG, Eastern Cooperative Oncology Group; EORTC QLQ-C30, European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; EOT, end of treatment; EQ-5D-5L, EuroQoL Five-dimension Five-level; HRQoL, health-related quality of life; IRC, imaging review committee; ORR, objective response rate; OS, overall survival; PD, progressive disease; PFS, progression-free survival; PR, partial response; PRO, patient-reported outcomes; PRO-CTCAE, Patient-Reported Outcome Common Terminology Criteria for Adverse Events; RECIST, response evaluation criteria in solid tumours; SAE, serious adverse event

For the whole ESR1-mut population, the CS reports the final OS from a data cut of 02 September 2022 and for the remaining efficacy and safety results from a data cut of 06 September 2021. For both post-hoc subgroups the data cut off was 02 September 2022 for PFS and OS and response rates, and 8th July 2022 for patient-reported outcomes (PRO) and adverse events data.

Outcomes informing the economic model were:

- Progression free survival (for post-hoc subgroups 1 (ESR1 mutation) and 2 (dual mutation); CS B.3.3.4)
- Overall survival (for post-hoc subgroups 1 (ESR1 mutation) and 2 (dual mutation); CS B.3.3.4)
- Time to treatment discontinuation
- HRQoL via the EQ-5D-5L (for subgroup 1 (ESR1 mutation) mapped to the EQ-5D-3L). Company clarification response A5 stated that the overall EQ-5D scores reported in clinical sections of the CS (B.2) are based on a subset of the EQ-5D data collected in EMERALD, but the economic model uses all the EQ-5D data collected, as per preferred NICE methodology. The EAG discuss this further in section 3.2.2, and in the cost-effectiveness section 4.2.5.2 below.
- Adverse events for elacestrant (Grade ≥ 3 occurring in ≥2% of patients receiving elacestrant in the ESR1-mut subgroup; CS B.3.4.4)

Appendix 2 of the trial protocol and CSR Table 6 show the methods, frequency and timing of all outcome assessments were identical between trial arms, reducing the risk of evaluation time bias.[17 19]

EAG comment on outcomes assessment

Overall, we consider the efficacy, HRQoL and safety outcomes to be appropriate to the decision problem and scope.

3.2.4 Statistical methods of the included studies

The CS provided details of the statistical methods used in the EMERALD trial in the CS, with additional detail to be found in the study protocol, SAP, CSR, and in company clarification response A5. A summary and EAG critique of the statistical methods used in the EMERALD trial are presented below in Table 8.

Table 8 Summary and critique of the statistical methods used in the EMERALD trial

Analysis populations
Intention-to treat (ITT) population: defined as all randomised subjects, with patients analysed according to their randomized treatment assignments. This is the primary analysis population for **PFS, OS** and **PROs**, including **HRQoL** (All ITT patients: N=478; ESR1-mut N=228) Per protocol (PP) and modified per protocol (mPP): defined as all randomised patients except those who had a major protocol deviation. This population was used for **sensitivity analyses for PFS** if the primary endpoint was statistically significant. (All PP patients: N=464; ESR1-mut PP: N=221; all mPP patients: N=461; ESR1-mut N=219) Response Evaluable (RE) population: defined as all ITT subjects who had measurable disease (i.e. at least 1 target lesion) at baseline and at least 1 postbaseline RECIST assessment on any (target or non-target) lesions and/or had a new lesion. This is the analysis population for **ORR** and **DoR.** (IRC assessed RE population: All patients: N=361; ESR1-mut N=171) Clinical Benefit Evaluable (CBE) population: defined as all ITT subjects who had measurable and/or evaluable disease (i.e. target and/or non-target lesions) at baseline and at least 1 post-baseline RECIST assessment on any (target or non-target) lesions and/or had a new lesion. This is the analysis population for **CBR.** (IRC assessed CBE population: All CBE patients: N=443; ESR1-mut N=212) Safety population: defined as all patients who received at least 1 dose of study medication. Patients were analysed according to the treatments they actually received in Cycle 1 [CSR section 9.7.1.2 p64]. This is the analysis population for **all safety outcomes** (All safety patients: N=467; ESR1-mut N=221)
EAG comment: The analysis populations are appropriate. As a proportion of all randomised patients, the safety population included 97.7% and the ESR1-mut safety population subgroup included 96.9%, thus minimal attrition bias.
Sample size calculations
The power calculation was based on the primary outcome, PFS. It was planned that 200 patients with ESR1-mut would need to be randomised to obtain **160** PFS events to provide 80% power to detect an HR of 0.610 at the two-sided alpha level of 2.5%. (CS Table 11). For all patients (ESR1-mut and ESR1-mut not detectable), 466 patients would

need to be randomised to obtain approximately **340** PFS events to have 92% power to detect a HR of 0.667 at the 2-sided alpha level of 2.5% (SAP 4.1)

EAG comment: CS B.2.12.2.1 states the final PFS analysis was conducted after **140** events due to an additional year needed to observe the pre-specified 160 number of events for the ESR1-mut subgroup. There were **300** events for the whole EMERALD trial population at this timepoint (CSR section 11.6.2.11). The EAG therefore considers the study to have reduced power and therefore uncertainty in the results of PFS for all patients and for the ESR1-mut subgroup.

Methods to account for multiplicity

The truncated Hochberg procedure was used to adjust for multiple statistical testing of the primary endpoints PFS for all patients and for patients with ESR1-mut only, and OS for all patients and for patients with ESR1-mut only (CS document B Table 11, CSR section 9.6.2)

EAG comment: The company's approach to handling multiple testing of outcomes is appropriate.

Analysis of outcomes

Primary analysis

Blind-IRC assessed PFS was performed on the ITT population incorporating randomisation stratification factors (for all patients these include ESR1- mutational status (ESR1-mut vs ESR1-mut-nd), prior treatment with fulvestrant (yes vs no), and presence of visceral metastases (yes vs no); for ESR1-mut subjects only, this includes prior treatment with fulvestrant (yes vs no) and presence of visceral metastases (yes vs no)). The Kaplan-Meier (KM) method was used to summarise time-to event outcomes. The Cox-proportional hazards model was used to estimate hazard ratios with 95% CI. The difference between treatment groups was analysed using the stratified log-rank test with the randomisation stratification factors for generation of p-value.

Key secondary outcome

OS was analysed using the same methods for PFS. (SAP 4.7.1, 4.7.2.1, 4.7.3.3). An interim OS analysis was performed at the primary PFS analysis, with a pre-specified adjusted 2-sided alpha level of 0.0001. The final analysis of OS was performed after the pre-specified 50% of patients had died, with a 2-sided alpha level of 0.0499 (SAP 4.7.2.1)

Secondary outcomes

ORR was compared between treatment groups using the Cochran-Mantel-Haenszel tests adjusting for randomisation stratification factors. The same methods were used for **CR**.

DoR was analysed using the KM method.

For **PROs** (**EQ-5D-5L**, **EORTC QLQ-C30** and the **PRO-CTCAE**) changes from baseline by study visit (with 95% CI) for each treatment group were used. In addition, for EORTC QLQ-C30, mixed model repeated measures (MMRM) were used to analyse change from baseline over study visits through to cycle 6.

For **safety outcomes,** only descriptive statistics (e.g., frequency, counts) were used.

EAG comment: Appropriate analytical methods were used for primary and secondary outcomes.

Handling of missing data

PFS (Primary analysis)

Censoring rules for the primary analysis of blinded IRC assessed PFS in the CS (CS document B Table 8) specified date of progression or censoring relating to missing assessments in the primary analysis:

- No baseline measurable or evaluable lesion: from date of randomisation
- No post-baseline assessments and no death: from date of randomisation
- Censored progression or death after missing ≥2 consecutive post-baseline tumour assessments: on date of last tumour assessment before missed assessments or date of randomisation, whichever is later.

The SAP (Table 2) additionally specified the date of progression for documented progression or death after missing 1 post-baseline tumour assessment should be the date of documented progression or death.

EQ-5D-5L

The company only had EQ-5D-5L index scores for countries in which the validated tool was available (5 countries: Denmark, France, Spain, UK and USA). For all other patients in the other countries the overall score was set to missing (Company clarification response A5). This missing data issue is in relation to EQ-5D-5L index scores presented in the clinical effectiveness section of the CS (B.2 and Appendix E) only - it does not apply to the EQ-5D analysis used to inform the economic model (CS B.3.4.1 and B.3.4.2, and clarification response B4 and Table 6).

EAG comment:

Primary analysis

Censoring relating to missing assessments in the primary analysis for PFS was similar between treatment groups for both ESR1-mut group and for all patients (CSR section .4.1.1).

EQ-5D-5L
The company's decision to obtain EQ-5D-5L index scores only for countries in which the validated tool was available (5 out of 17 countries enrolled in the trial) resulted in large differences in the number of patients in each arm of the ESR1-mut subgroup with an EQ-5D-5L index score versus an EQ-VAS score (50 (43%) versus 108 (94%) in the elacestrant arm and 50 (44%) versus 98 (87%) in the SOC arm). EQ-5D-5L index score data for the ESR1-mut subgroup presented in the clinical effectiveness section CS B.2 and for post-hoc subgroups 1 (ESR1 mutation) and 2 (dual mutation) in CS Appendix E should be interpreted with due caution given this small, unrepresentative sample.

Sensitivity analyses

PFS
For events that were recorded after missing 2 or more consecutive tumour assessments: '**actual event PFS analysis**' that defined the event date as the actual event date after the 2 missed tumour assessments.

For events that were recorded after missing 2 or more consecutive tumour assessments a '**backdating PFS analysis**' which defined the event date as the date of the next scheduled tumour assessment after the last adequate tumour assessment.

Assessing the impact of stratification and compared the two treatment groups using an **unstratified log-rank test**.

Using **Per Protocol population** in the same manner as the primary efficacy analysis if the primary endpoints were statistically significant.

Patient reported outcomes (PROs)
Excluding patients who had **at least 1 missing visit due to COVID-19**. Performed for all PRO outcomes in the same manner as the primary PRO analyses.

EAG comment: The sensitivity analyses are comprehensive.

Subgroup and post-hoc analyses

Pre-specified subgroup analyses (in addition to ESR1-mut) included:

- Prior treatment with fulvestrant; presence of visceral metastasis; age (<65 years, ≥65 years, <75 years. ≥75 years); race (Caucasian, Asian, other); region (Europe, North America, Asia); baseline ECOG Performance Status (0,1); measurable disease at baseline (yes, no); number of prior lines of endocrine therapy in the advanced/metastatic setting (1,2); number of lines of chemotherapy in the advanced/metastatic setting (0,1).

These subgroup analyses were performed for PFS, OS, ORR, DoR and CBR outcomes. CS document B section 2.7 specified that subgroup analyses were not performed if the number of patients in the subgroup of each treatment group was <5% however, company clarification response A6 confirmed these analyses were performed regardless of this threshold.

Post-hoc subgroup analyses reported in the CS included patients with:

- ESR1-mut who had received ≥12 months of prior ET + CDK4/6i (referred to as "post-hoc subgroup 1 (ESR1-mutation)" in this report), and
- ESR1-mut and PIK3CA mutations (dual mutated) who had received ≥12 months of prior ET + CDK4/6i (referred to as "post-hoc subgroup 2 (dual mutation)" in this report).

Company clarification response A7 provides a list of post-hoc analyses from the EMERALD trial in the public domain as conference abstracts.

EAG comment:

- The chosen pre-specified subgroups are appropriate to this condition. However, clinical expert advice to the EAG is that bone metastases is a very important prognostic factor and should have considered for inclusion as a subgroup.
- The CS presents results of pre-specified subgroup analyses only for blinded IRC-assessed PFS (as opposed to other outcomes), and for the ESR1-mut population (not the whole trial population) (CS Appendix E.1).
- As the CS itself notes, the trial was not statistically powered for subgroups, therefore statistical significance cannot be inferred from the results of any subgroup analyses.

Additional caution is needed in the interpretation of the two post hoc subgroup analyses:

- The sample sizes are small, notably in subgroup 2 (dual mutation group). Subgroup 1 included 33% of the randomised trial population (n=159/478); Subgroup 2 included 13% of the randomised trial population (n=62/478).

- In subgroup 2, the distribution of patients between the elacestrant and SOC trial arms is slightly uneven (11% vs 15%, respectively).
- Baseline characteristics (demographic, treatment history and performance status) were generally balanced across the trial arms, but with some notable differences in the percentage of patients in each arm (10% to 20% of patients) mainly affecting subgroup 2 (dual mutation patients). In this subgroup there was a [redacted] percentage of patients in the elacestrant arm with visceral metastases. Likewise, a [redacted] proportion of elacestrant patients previously had two lines of endocrine therapy in the advanced/metastatic setting, and [redacted] had received prior adjuvant therapy. This suggests that patients treated with elacestrant were in a [redacted] than was the case for patients receiving standard of care endocrine monotherapy.
- The post hoc status of the subgroup analysis means the results are at increased risk of bias, although the impact of these imbalances is unclear.Post hoc subgroup analyses in clinical trials should be considered as exploratory, hypothesis generating, rather than being confirmatory.

The list of post-hoc analyses provided by the company is limited to those in the public domain. It is unclear whether additional post-hoc analyses were performed that are not in the public domain.

Source: Partly reproduced from CS Table 11. Additional sources: CS B. 2.12.2.1; CS document B Table 10, CS Appendix E.1; Protocol section 11.2; SAP sections 3.1, 4.1, 4.7.1, 4.7.2.1, 4.7.3.3 and 4.8.4; CSR sections 4.1.1, 9.6.2, 9.7.1.2, 11.4.1.1 and 11.6.2.11; CSR Tables 11, 14.2.1.1.1 and 14.2.1.1.2; Company clarification response A5
Abbreviations: CBR, clinical benefit rate; CR, complete response; DOR, duration of response; EORTC QLQ-C30, European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; EQ-5D-5L, EuroQoL Five-dimension Five-level; ESR1, oestrogen receptor 1 gene; HRQoL, health-related quality of life; IRC, imaging review committee; mut, mutation; ORR, objective response rate; OS, overall survival; PFS, progression-free survival; PR, partial response; PRO, patient-reported outcomes; PRO-CTCAE, Patient-Reported Outcome Common Terminology Criteria for Adverse Events; SOC, standard of care.

EAG comment on study statistical methods

The main limitation of the statistical analysis of the EMERALD trial was that the study was not adequately powered for the analysis of the primary efficacy outcome (PFS) for all patients and for the ESR1-mut subgroup. The EAG therefore considers there is uncertainty in the results of PFS for all patients and for the ESR1-mut subgroup. Furthermore, results for the two post-hoc subgroups, 1 (ESR1 mutation) and 2 (dual mutation) should also be interpreted with caution given they were not powered to detect statistical significance, are relatively small in sample size and were selected for analysis

based on knowledge of their results, rather than being pre-specified before data collection.

3.2.5 Efficacy results of the intervention studies

Below we summarise results from the EMERALD trial for outcomes used in the economic model, namely progression free survival, overall survival, HRQoL via the EQ-5D-5L, and adverse events. Results for other outcomes (e.g. tumour response) are available in the CS and/or the trial CSR.[17]

3.2.5.1 Progression-free survival (PFS)

Blinded-IRC assessed PFS was the primary endpoint of the EMERALD trial. The company submission reported results for blinded-IRC assessed PFS for the ESR1- mut subgroup (CS document B section 2.6.1), post-hoc subgroup 1 (ESR1-mutation; CS document B section 2.7.2.1) and post-hoc subgroup 2 (dual mutation; CS document B section 2.7.3.1). Results for blinded-IRC assessed PFS for all patients were reported in Bidard et al., 2022 and the CSR.[11 17]

ESR1-mut subgroup

Table 9 summarises the primary analysis of blinded IRC-assessed PFS for the ESR1-mut subgroup in the ITT population. At the 6 September 2021 data cut a total of 140 PFS events had been recorded which was less than the 160 PFS events planned for the primary analysis (see Table 8). The EAG therefore considers the study to have reduced power and therefore uncertainty in the results of PFS for the ESR1 mut subgroup presented.

Fewer patients in the elacestrant arm progressed or died compared to the SOC arm (n=62 [53.9%] vs. 78 [69.0%], a difference of 15.1%). An absolute increase of 1.9 months in median PFS was observed with elacestrant (3.8 months; 95% CI 2.17 to 7.26) versus SOC (1.9 months (95% CI 1.87 to 2.14). The stratified HR was 0.55 (95% CI 0.39 to 0.77) signifying a 45% reduction in the risk of disease progression or death in patients with the ESR1 mutation receiving elacestrant.

Table 9 Primary analysis of blinded IRC-assessed PFS in the ESR1-mut subgroup in the EMERALD trial

	Elacestrant **N=115**	**SOC** **N=113**
HR (95% CI) P-value	0.55 (0.39 to 0.77) 0.0005	

	Elacestrant N=115	SOC N=113
Median PFS months (95% CI)	3.8 (2.17 to 7.26)	1.9 (1.87 to 2.14)
Events, n (%) Death Progression	62 (53.9) 3 (2.6) 59 (51.3)	78 (69.0) 1 (0.9) 77 (68.1)
3-month PFS rate (95% CI)	55.93 % (45.80 to 66.05)	39.55% (29.44 to 49.65)
6-month PFS rate (95% CI)	40.8% (30.1 to 51.4)	19.1% (10.5 to 27.8)
12-month PFS rate (95% CI)	26.8% (16.2 to 37.4)	8.2% (1.3 to 15.1)
18-month PFS rate (95% CI)	24.33% (13.68 to 34.98)	-

Source: Reproduced from CS Table 13
Abbreviations: CI, confidence interval; ESR1, oestrogen receptor 1 gene; HR, hazard ratio; IRC, imaging review committee; mut, mutation; PFS, progression-free survival; SOC, standard of care

The Kaplan Meier plot of blinded IRC assessment of PFS (CS figure 9, not reproduced here) shows a separation of the survival curves after 2 months. A consistently higher proportion of patients remained alive and progression free in the elacestrant arms compared to SOC at 2 months, 6 month, 12 months and 18 months.

Sensitivity analyses were consistent with results of the primary study in the ITT population (see Table 10). Results for pre-specified subgroup analyses are reported in section 3.2.5.4.

Table 10 Sensitivity analyses of blinded IRC-assessed PFS in the ESR1-mut subgroup in the EMERALD trial

Sensitivity analysis[a]	Hazard ratio	95% CI	P-value
Actual event PFS	0.542	0.385 to 0.759	0.0004
Back dating PFS	0.542	0.385 to 0.759	0.0004
Unstratified	0.531	0.378 to 0.743	0.0002
Per protocol population	0.543	0.385 to 0.764	0.0005

Source: Partly reproduced from CSR Tables 14.2.1.2.1, 14.2.1.3.1, 14.2.1.4.1 and 14.2.1.6.1
Abbreviations: CI, confidence interval; ESR1, oestrogen receptor 1 gene; IRC, imaging review committee; mut, mutation; PFS, progression-free survival
[a] See Table 8 of this report for definitions of these sensitivity analyses

All patients

Overall, the results for blinded IRC-assessed PFS for all patients were consistent with those for the ESR1-mut subgroup, albeit the reduction in the risk of disease progression or death with elacestrant compared to SOC was less (30%; HR 0.70; 95% CI 0.55 to 0.88; Bidard et

al., 2022; CSR Tables 17, 14.2.1.2.2, 14.2.1.3.2, 14.2.1.4.2 and 14.2.1.6.2).[11 17]. It should be noted that at the 6 September 2021 data cut a total of 300 PFS events had been recorded which was less than the 340 PFS events planned for the primary analysis (see Table 8). The EAG therefore considers the study to have reduced power and therefore uncertainty in the results of PFS for all patients.

Post-hoc subgroups 1 (ESR1 mutation and ≥12 months prior ET + CDK4/6i) and 2 (dual mutation)

Table 11 summarises analyses of blinded IRC-assessed PFS, with a data cut of 2 September 2022, for post-hoc subgroups 1 (ESR1 mutation) and 2 (dual mutation). Interpretation of the following results of these post-hoc analyses should be made with caution given they were not powered to detect statistical significance.

Table 11 Blinded IRC-assessed PFS in post-hoc subgroups 1 (ESR1 mutation) and 2 (dual mutation) in the EMERALD trial

	Post-hoc subgroup 1 (ESR1 mutation)		**Post-hoc subgroup 2 (dual mutation)**	
	Elacestrant N=78	SOC N=81	Elacestrant N=27	SOC N=35
HR (95% CI) p-value	0.410 (0.262 to 0.634) <0.0001		0.423 (0.176 to 0.941) -	
Median PFS months (95% CI)	8.61 (4.14 to 10.84)	1.91 (1.87 to 3.68)	5.45 (2.14 to 10.84)	1.94 (1.84 to 3.94)
Events, n (%) Death Progression	39 (50) 1 (1.3) 38 (48.7)	53 (65.4) 1 (1.2) 52 (64.2)	[redacted]	[redacted]
3-month PFS rate (95% CI)	68.30 (56.67 to 79.93)	41.55 (29.19 to 53.90)	[redacted]	[redacted]
6-month PFS rate (95% CI)	55.81 (42.69 to 68.94)	22.66 (11.63 to 33.69)	[redacted]	[redacted]
12-month PFS rate (95% CI)	35.81 (21.84 to 49.78)	8.39 (0.00 to 17.66)	[redacted]	[redacted]
18-month PFS rate (95% CI)	28.49 (14.08 to 42.89)	0.00 (-)	[redacted]	[redacted]

Source: Reproduced from CS Tables 21 and 23
Abbreviations: CI, confidence interval; ESR1, oestrogen receptor 1 gene; HR, hazard ratio; IRC, imaging review committee; n, number of patients with the observed characteristic; N, total number in group; PFS, progression-free survival; SOC, standard of care

Overall, the results for both post-hoc subgroups 1 (ESR1 mutation) and 2 (dual mutation) were consistent to those for the ESR1-mut subgroup, albeit:

- The reduction in the risk of disease progression or death with elacestrant compared to SOC was greater (post-hoc subgroup 1 (ESR1 mutation): 59%; HR 0.41 95% CI 0.26 to 0.63; post-hoc subgroup 2 (dual mutation): 58%; HR 0.42 95% CI 0.18 to 0.94; ESR1-mut: 45%; HR 0.55 (95% CI 0.39 to 0.77).
- The absolute increase in median PFS observed with elacestrant versus SOC was greater (post-hoc subgroup 1 (ESR1 mutation): 6.7 months; post-hoc subgroup 2 (dual mutation): 3.51 months; ESR1-mut: 1.9 months).

Clinical expert advice to the EAG were that the absolute median increase in PFS observed with elacestrant versus SOC in post-hoc subgroup 1 (ESR1-mut) would provide a meaningful benefit to most patients, while that observed in post-hoc subgroup (dual mutation) was less so.

As with ESR1-mut subgroup, Kaplan Meier plots for post-hoc subgroups 1 (ESR1 mutation) and 2 (dual mutation) (CS Figures 12 and 14 respectively; not reproduced here) show a separation of the survival curves after 2 months.

3.2.5.2 Overall Survival (OS)

Overall survival (OS) was the key secondary endpoint of the EMERALD trial. The company submission reported results for an interim analysis (data cut 6 September 2021) and final analysis (data cut 2 September 2022) for the ESR1-mut subgroup (CS document B section 2.6.2); and results of the final analysis (data cut 2 September 2022) for post-hoc subgroups 1 and 2 (CS document B section 2.7.2.2 and 2.7.3.2 respectively). For all patients, results for an interim analysis (data cut 6 September 2021) were reported in Bidard et al., 2022 and the CSR (section 11.4.1.2) and, for the final analysis (data cut 02 September 2022), in an Overall Survival Addendum provided by the company.[11 17 20]

Interim analysis

An interim analysis of OS was performed on the same data cut (6 September 2021) as the final analysis for PFS. At this time, in the ESR1-mut subgroup, 24.3% of patients in the elacestrant arm had died and 35.4% in the SOC arm. The stratified HR was 0.59 (95% CI 0.36 to 0.96). The stratified log rank test p-value was 0.0325. At a pre-specified adjusted alpha level of 0.0001 (Table 8), the difference in OS between elacestrant and SOC was not statistically significant. Results for the interim analysis for all patients were similar (HR 0.75, 95% CI 0.54 to 1.04; p=0.0821; Bidard et al., 2022).[11]

Final Analysis

The data cut for the final OS analysis for ESR1-mut subgroup, all patients, and post-hoc subgroups 1 (ESR1 mutation) and 2 (dual mutation) was 02 September 2022.

There was no statistically significant difference in the hazard rate of death for elacestrant compared to SOC for the ESR1-mut subgroup (stratified HR 0.903, 95% CI 0.629 to 1.298; p-value =0.5823). Results were similar for all patients (stratified HR 0.912, 95% CI 0.708 to 1.175; p=0.476; Table 1 Overall Survival Addendum).[20]

Results for post-hoc analyses need to be interpreted with caution given they were not powered to detect statistical significance. There was no difference in the hazard rate of death for elacestrant compared to SOC for either post-hoc subgroup 1 (stratified HR [redacted], 95% CI [redacted] to [redacted]; p=[redacted]) or subgroup 2 (stratified HR [redacted], 95% CI [redacted] to [redacted]; p=[redacted]).

3.2.5.3 HRQoL outcomes

Data on EQ-5D-5L were reported in CS document B. section 2.6.4 (patients with ESR1-mut), CS document B section 2.7.2.3 and Appendix E .2.1.1 (post-hoc subgroup 1 (ESR1 mutation)) and CS document B section 2.7.3.3 and Appendix E.3.1.1 (post-hoc subgroup 2 (dual mutation)).

There are two main issues concerning missing data for the EQ-5D-5L index score for the ESR1-mut subgroup presented in the sections of the CS listed above (i.e. they do **not** apply to the EQ-5D analysis that was used to inform the economic model), which impact on their relevance for this appraisal. First, the company decided to obtain EQ-5D-5L index scores only for countries in which the validated tool was available (5 out of 17 countries enrolled in the trial; see company clarification response A5). For the ESR1-mut subgroup this resulted in just under half of patients in each arm having an EQ-5D-5L index score. Second, there is a difference in the total number of patients with ESR-mut 1 enrolled in four (France, Spain, UK and USA; CSR Table 14.1.1.2) of the five countries and those that had a baseline EQ-5D-5L score (CSR Table 14.2.6.4.1). These issues are described in more detail in section 3.2.2 of this report.

For completeness, the EAG report the company's findings for EQ-5D-5L index score. Namely, the CS (document B section 2.6.4) reports that EQ-5D-5L index scores for ESR-mut subgroup were similar between elacestrant and SOC at end of treatment, with no changes within groups over time. Results were similar for all patients (CS document B section 2.6.4) and for post-hoc subgroups 1 (ESR1 mutation) and 2 (dual mutation) (CS Appendix E

section 2.1.1 and section 3.1.1 respectively). However, given the issues with missing data for this outcome, the EAG considers these findings irrelevant for decision making purpose*s.* See section 4.2.5.2 below for discussion of the utility analysis of EQ-5D-5L index scores that informed the company's economic model, which used a more complete data set.

3.2.5.4 Subgroup analyses

CS Appendix E Figure 3 reports a forest plot of pre-specified subgroup analyses for the primary outcome of blinded IRC-assessed PFS for the ESR1 mut subgroup only at the 6th September 2021 data cut.

Subgroups included:

- **baseline demographic characteristics** (age (<65 years, ≥65 years, <75 years, ≥75 years), race, region),
- **measures of base disease status** (presence of visceral metastasis, baseline ECOG Performance Status, measurable disease at baseline) and
- **prior treatment** (prior treatment with fulvestrant, number of prior lines of endocrine therapy in the advanced/metastatic setting, number of lines of chemotherapy in the advanced/metastatic setting).

In CS Appendix E.1 the company state hazard ratios in patients with ESR1-mut across all pre-specified subgroups numerically favoured elacestrant and demonstrated consistency with the primary endpoint PFS (HR 0.531, 95% CI 0.378 to 0.743). The EAG agree that the point estimates for the hazard ratios were less than one, signifying a reduction in risk of disease progression or death, however, 95% confidence intervals for the following subgroups crossed 1:

Table 12 Pre-specified subgroup analyses of blinded IRC-assessed PFS in all patients with ESR1 mut where 95% CI crossed 1

Pre-specified subgroup	Hazard Ratio (95% CI)
Demographics	
Age: ≥75 years	0.514 (0.193 to 1.273)
Race: Asian	0.891 (0.122 to 4.652)
Race: other	0.289 (0.040 to 1.503)
Region: Europe	0.624 (0.386 to 1.011)
Region: Asia	0.552 (0.149 to 1.678)
Measures of base disease status	
Measurable disease at baseline: no	0.834 (0.333 to 2.178)

Pre-specified subgroup	Hazard Ratio (95% CI)
Presence of visceral metastasis: no	0.736 (0.381 to 1.443)
Prior treatment	
Prior treatment with fulvestrant: yes	0.621 (0.297 to 1.257)
Number of lines of chemotherapy in advanced or metastatic setting: 1	0.696 (0.358 to 1.308)

Source: Partly reproduced from CS Appendix E Figure 3
Abbreviations: CI, confidence interval; ESR1, oestrogen receptor 1 gene; IRC, imaging review committee; mut, mutation; PFS, progression-free survival

Caution however, is required in the interpretation of the results of these subgroup analyses given that the trial was not powered to demonstrate statistically significant treatment differences according to subgroups. Furthermore, some HRs, and their 95% confidence intervals, are calculated based on low numbers of events.

3.2.5.5 Safety outcomes

Data on adverse were reported in CS document B section 2.10 (both for all patients and for patients with ESR1-mut), CS document B section 2.7.2.3 and Appendix E .2.2 (post-hoc subgroup 1 (ESR1 mutation)) and CS document B section 2.7.3.3 and Appendix E.3.2 (post-hoc subgroup 2 (dual mutation)).

The majority of patients (>84%) in both the elacestrant and SOC arms in all patients, ESR1-mut subgroup, and post-hoc subgroups 1 and 2 experienced treatment emergent adverse events (see Table 13 and Table 14). The most common adverse event for patients receiving elacestrant was nausea, which was consistent for all patients, ESR1-mut subgroup, and post-hoc subgroups 1 and 2 (35.0%, 34.8%, 38.5% and [redacted] respectively; see Table 13 and Table 14). The most common adverse event for the SOC group differed between patient populations: for all patients nausea and fatigue (both 19.1%), for ESR1-mut subgroup fatigue (19.8%), for post-hoc subgroup 1 [redacted] and post-hoc subgroup 2 [redacted]. The proportion of patients experiencing adverse events with a severity grade ≥ 3 was similar between elacestrant and SOC for all patients, ESR1-mut subgroup, [redacted] (see Table 13 and Table 14).

The proportion of patients who experienced adverse events leading to dose interruption was greater in the elacestrant group compared to the SOC groups for all patients, ESR1-mut subgroup, [redacted] (see Table 13 and Table 14).

Treatment-related adverse events, serious adverse events, fatal events and adverse events leading to discontinuation were reported for all patients and for the ESR1-mut subgroup only (see Table 13). The findings for these adverse events were consistent between all patients and the ESR1- mut subgroup. Briefly,

- A similarly higher proportion of events were considered treatment related in the elacestrant group (63.3% and 61.7%) compared to the SOC group (43.5% and 46.2%).
- A similar proportion of patients experienced serious adverse events in the elacestrant group (12.2% and 12.2%) compared to the SOC group (10.9% and 11.3%).
- There were a small number of fatal events in the elacestrant group (1.7% and 2.6%) and SOC group (2.6% and 0.9) with none of the deaths considered treatment related.
- A similar proportion of patients experienced adverse events that led to discontinuation in the elacestrant group (6.3% and 5.2%) compared to the SOC group (4.3% and 3.8%)

Table 13 Summary of adverse events for the All patients and for ESR1-mut subgroup

Adverse event (AE)	**All Patients**		**ESR1-mut**	
	Elacestrant N=237 n (%)	**SOC N=230 n (%)**	**Elacestrant N=115 n (%)**	**SOC N=106 n (%)**
Any TEAE	218 (92.0)	198 (86.1)	105 (91.3)	92 (86.8)
Treatment related AE	150 (63.3)	100 (43.5)	71 (61.7)	49 (46.2)
Grade ≥3	64 (27.0)	48 (20.9)	32 (27.8)	23 (21.7)
Serious AE	29 (12.2)	25 (10.9)	14 (12.2)	12 (11.3)
Fatal events	4 (1.7)	6 (2.6)	3 (2.6)	1 (0.9)
AE leading to discontinuation	15 (6.3)	10 (4.3)	6 (5.2)	4 (3.8)
AE leading dose interruption	36 (15.2)	12 (5.2)	25 (21.7)	7 (6.6)
AE reported in ≥ 10% of patients in either trial arm				
Nausea	83 (35.0)	44 (19.1)	40 (34.8)	19 (17.9)
Arthralgia	34 (14.3)	37 (16.1)	23 (20.0)	19 (17.9)
Vomiting	45 (19.0)	20 (8.7)	21 (18.3)	10 (9.4)
Fatigue	45 (19.0)	44 (19.1)	20 (17.4)	21 (19.8)
Decreased appetite	35 (14.8)	22 (9.6)	19 (16.5)	8 (7.5)
Diarrhoea	33 (13.9)	23 (10.0)	17 (14.8)	13 (12.3)

Adverse event (AE)	All Patients		ESR1-mut	
	Elacestrant N=237 n (%)	SOC N=230 n (%)	Elacestrant N=115 n (%)	SOC N=106 n (%)
Back pain	33 (13.9)	22 (9.6)	16 (13.9)	9 (8.5)
Headache	29 (12.2)	26 (11.3)	15 (13.0)	11 (10.4)
Dyspepsia	24 (10.1)	6 (2.6)	13 (11.3)	3 (2.8)
Insomnia	18 (7.6)	11 (4.8)	13 (11.3)	7 (6.6)
Constipation	29 (12.2)	15 (6.5)	12 (10.4)	8 (7.5)
Aspartate aminotransferase increased	31 (13.1)	29 (12.6)	12 (10.4)	15 (14.2)
Anaemia	22 (9.3)	17 (7.4)	11 (9.6)	11 (10.4)
Hot flush	27 (11.4)	19 (8.3)	11 (9.6)	8 (7.5)
Alanine aminotransferase increased	22 (9.3)	24 (10.4)	6 (5.2)	13 (12.3)

Source: Partly reproduced from CS document B Table 31
Abbreviations: AE, adverse event; ESR1, oestrogen receptor 1 gene; mut, mutation; n, number of patients with the observed characteristic; N, total number in group; PFS, progression-free survival; SOC, standard of SOC, standard of care; TEAE, treatment-emergent adverse event

Table 14 Summary of adverse events for post-hoc subgroups 1 (ESR1 mutation) and 2 (dual mutation)

Adverse event (AE)	Post-hoc subgroup 1 (ESR1 mutation)		Post-hoc subgroup 2 (dual mutation)	
	Elacestrant N=78 n (%)	SOC N=75 n (%)	Elacestrant N=27 n (%)	SOC N=32 n (%)
Any TEAE	[redacted]	[redacted]	[redacted]	[redacted]
Grade ≥3 in ≥ 2% of patients	[redacted]	[redacted]	[redacted]	[redacted]
AE leading dose interruption	[redacted]	[redacted]	[redacted]	[redacted]
AE reported in ≥ 10% of patients in either trial arm				
Nausea	30 (38.5)	11 (14.7)	[redacted]	[redacted]
Arthralgia	[redacted]	[redacted]	[redacted]	[redacted]
Vomiting	16 (20.5)	6 (8)	[redacted]	[redacted]
Diarrhoea	16 (20.5)	9 (12)	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]	[redacted]	[redacted]
Back pain	[redacted]	[redacted]	[redacted]	[redacted]

Adverse event (AE)	Post-hoc subgroup 1 (ESR1 mutation)		Post-hoc subgroup 2 (dual mutation)	
	Elacestrant N=78 n (%)	SOC N=75 n (%)	Elacestrant N=27 n (%)	SOC N=32 n (%)
Headache	13 (16.7)	9 (12)	[redacted]	[redacted]
Decreased appetite	12 (15.4)	5 (6.7)	[redacted]	[redacted]
Dyspepsia	10 (12.8)	3 (4)	[redacted]	[redacted]
Hot flush	9 (11.5)	7 (9.3)	[redacted]	[redacted]
Pain in extremity	[redacted]	[redacted]	[redacted]	[redacted]
Asthenia	[redacted]	[redacted]	[redacted]	[redacted]
Aspartate aminotransferase increased	[redacted]	[redacted]	[redacted]	[redacted]
Blood cholesterol increased	[redacted]	[redacted]	[redacted]	[redacted]
Urinary tract infection	[redacted]	[redacted]	[redacted]	[redacted]
Insomnia	[redacted]	[redacted]	[redacted]	[redacted]
Dyspnoea	[redacted]	[redacted]	[redacted]	[redacted]
Anaemia	[redacted]	[redacted]	[redacted]	[redacted]
Blood glucose increased	[redacted]	[redacted]	[redacted]	[redacted]
Stomatitis	[redacted]	[redacted]	[redacted]	[redacted]
Musculoskeletal pain	[redacted]	[redacted]	[redacted]	[redacted]
Alanine aminotransferase increased	[redacted]	[redacted]	[redacted]	[redacted]

Source: Partly reproduced from CS Appendix E Table 11, Table 12, Table 15 and Table 16
Abbreviations: AE, adverse event; ESR1, oestrogen receptor 1 gene; mut, mutation; n, number of patients with the observed characteristic; N, total number in group; PFS, progression-free survival; SOC, standard of SOC, standard of care; TEAE, treatment-emergent adverse event

3.2.6 Pairwise meta-analysis of intervention studies

CS section B.2.8. states that since only one trial of elacestrant relevant to this NICE appraisal is available (i.e. the EMERALD trial) it is therefore not possible to conduct meta-analysis currently. The EAG concurs with this assertion.

3.3 Critique of studies included in the indirect treatment comparison

3.3.1 Rationale for the indirect treatment comparison

As mentioned earlier (section 3.2.1), the pivotal EMERALD trial compared elacestrant against standard of care endocrine monotherapy, comprising either fulvestrant or an aromatase inhibitor (anastrozole, letrozole, or exemestane) chosen by investigators at each study centre. None of the treatments in the comparator arm of the trial match the company's

chosen comparators in the decision problem (i.e. everolimus plus exemestane, or alpelisib plus fulvestrant). For this reason, an indirect treatment comparison was required to provide comparative efficacy estimates for elacestrant in the company's proposed subgroup patients with an ESR1-mutation who have disease progression following ≥12 months prior treatment with ET + CDK4/6i.

3.3.2 Identification, selection and feasibility assessment of studies for the indirect treatment comparison

In addition to studies of the efficacy and safety elacestrant, the company's "global clinical SLR" was designed to identify any treatments relevant to the decision problem. These included endocrine therapy, CDK4/6 inhibitors, and chemotherapy.

Neither everolimus plus exemestane, or alpelisib plus fulvestrant are indicated for patients with the ESR1 mutation and, unsurprisingly, the company's SLR didn't identify any trials of these treatments in patients relevant to the decision problem (i.e. ESR1-mutation patients treated with ≥12 months of prior ET + CDK4/6i) which could be included in an indirect treatment comparison. For this reason the company decided to use matching-adjusted indirect comparison (MAIC) methodology, informed by the individual patient data from the EMERALD trial and aggregated data from a source of real-world evidence. The CS states their approach is aligned with the core principles outlined in the NICE real-world evidence framework, though no further detail is given specifically on how the framework was applied, nor is a definition of real-world evidence given.

Few details of the search for real-world evidence are provided in the CS. The CS states that due to the absence of ESR1 mutation testing in the UK they searched for real-world evidence sources "outside the UK and Europe" (CS page 82). It is not stated whether ESR1 mutation testing is done elsewhere in Europe and whether (non-UK) European sources were searched. In response to an EAG clarification question the company stated that no European datasets were found which reported the ESR1-mutation status of patients (clarification question A11). Consequently a "targeted literature review" was performed for electronic health record real-world data sources in the United States (US). They do not state whether searches were done for real-world evidence elsewhere other than Europe and the US.

The EAG has summarised the company's criteria for selecting a real-world evidence source – specifically a registry of patient health records - in Table 15 below. As we comment, some of the criteria are not fully defined and the process by which these were assessed is not

specified. However, the EAG recognises that a pragmatic approach may be needed when there is limited choice of evidence available.

Table 15 The company's criteria for selecting real-world evidence

Criterion	EAG comment
"The primary criterion was the detailed and accurate documentation of ESR1-mutations."	This is appropriate to the elacestrant marketing authorisation, i.e. treatment of patients with the ESR1-mutation.
"A sufficiently large sample size to ensure statistical validity and robustness".	There is no indication of how many patients would be needed to fulfil this criterion.
"Accuracy of mutation documentation and treatment records"	It is not stated how accuracy was demonstrated. For example, whether based on standard database quality assurance procedures, or whether the company performed checks of their own.
"Compliance with all relevant data protection regulations and ethical standards"	The regulations and standards are not specified, but we presume the company checked these with the database owners.

Source: Partly reproduced from company's response to EAG clarification question A11.

Two US databases were considered by the company as potential evidence sources for the ITC: **Patient360 Breast** (ConcertAI) and the **Flatiron Health Clinico-Genomic Database** (FLATIRON HEALTH). The CS does not mention if any other US databases were considered. Of the two options, the company chose the Flatiron database to inform their analysis. The CS describes Flatiron as "*a real-world database which gathers clinical data from electronic health records filled by cancer care providers across the US*" (page 82). In response to clarification question A11 the company state they chose Flatiron due to its:

- Larger sample of patients meeting the inclusion criteria for this study (the EAG presumes they mean the decision problem for this NICE appraisal),
- Greater number of patients who received everolimus and exemestane as second or third-line therapy;
- Robustness and its "*regulatory-grade quality and proven acceptability"*

EAG comment

The company's justification for an indirect treatment comparison is appropriate. The EAG agrees that a matched adjusted indirect treatment comparison (MAIC) is appropriate given the specific patient population in the decision problem. The EAG recognises the necessity to use real-world evidence for the comparator treatments (due to a lack of suitable clinical trial data), however, this introduces an additional level of uncertainty to the indirect treatment comparison. Limited detail is given about the company's search for a suitable patient health record database for the comparator treatments. The database selected by the company was one of two sources identified by a targeted search in the US. It is unclear whether any other potentially relevant sources are available, hence a more systematic search on a global scale would have been preferred.

3.4 Critique of the methods and procedures for conducting the MAIC

The process followed by the company to construct and implement the MAIC involved a series of steps. We discuss and critique these in the sub-sections below.

3.4.1 Application of the inclusion criteria for the Flatiron database to the EMERALD trial

The company selected patients from Flatiron according to criteria aligned to the EMERALD trial including: confirmed diagnosis of breast cancer; evidence of ER+/HER2; tested positive for ESR1-mutation any time before or within 28 days after the start date of index line; diagnosis at stage III unresectable/stage IV (or earlier diagnosis); evidence of treatment with endocrine therapy or a CDK6/4 inhibitor in first line and/or second line.

In addition to the above, patients had to have received everolimus and exemestane or alpelisib and fulvestrant in second line and/or third line in the advanced/metastatic setting. It is not explicitly stated how patients who had disease progression following ≥12 months prior treatment with endocrine therapy and CDK6/4 inhibitor were identified in Flatiron, but the EAG notes that outcome data (OS and PFS) are stratified by CDK4/6 inhibitor exposure time. In the absence of information on how duration of previous endocrine therapy was identified the EAG assumes that exposure time for previous CDK6/4 inhibitor treatment = exposure time for previous endocrine therapy since, in practice, CDK6/4 inhibitor is usually given in combination with endocrine therapy. Importantly, disease progression on previous CDK4/6i treatment in combination with fulvestrant or an aromatase inhibitor was an inclusion criterion for the EMERALD trial. Hence, reassurance is needed that the relevant patients were accurately identified from Flatiron.

3.4.2 Identification of prognostic factors and treatment effect modifiers to be included in the MAIC

The CS presents a list of 14 prognostic factors and treatment effect modifiers (with no distinction between the two) identified by "key opinion leaders" (CS Table 5). There is no further detail given on the key opinion leaders (e.g. how many were consulted; their professional background/speciality/position; their geographical location) or the process by which they identified the prognostic factors and treatment effect modifiers (e.g. based on clinical experience and/or empirical evidence; Delphi-consensus setting exercise). It is not clear whether the key opinion leaders is the same group of UK expert clinicians who the company consulted regarding the position of elacestrant in the care pathway.

The factors identified as prognostic included patient characteristics (namely, age and menopausal status); ECOG performance status; metastases (e.g. bone, visceral); previous treatment history (e.g. number of treatment lines in the metastatic setting, prior chemotherapy); cancer diagnosis (e.g. de novo advanced/metastatic vs. recurrent disease (adjuvant)). Of the 14 prognostic factors and treatment effect modifiers identified (CS Table 5), only three had data available to enable them to be included in the MAIC for the purpose of matching patients from EMERALD to Flatiron. These were:

- **Age** (50 years and older),
- **Prior endocrine therapy** (number of lines), and
- **Prior chemotherapy** status.

Additionally, three further prognostic factors were "partially" included in the MAIC:

- **Menopausal status** –assumed based on age restriction to patients 50 years old or greater from Flatiron (proxy measure).
- **Length of time on prior CDK4/6i** –"implicitly through population restriction (prior CDK4/6i ≥12 months)".
- **Oestrogen receptor expression** – "implicitly" included through focus on the ESR1 mutation.

The CS also comments that approximately 25% of patients in the Flatiron MAIC populations were missing ECOG performance status data. To address this the company did a sensitivity analysis redistributing patients without an ECOG performance status to the known categories (i.e. ECOG performance status of 0, 1, 2 etc). It is not stated what proportions of these patients were assigned to the ECOG categories, for example, whether weighting was

proportional to the relative size of each existing category. The company state that the sensitivity analysis showed similar results observed to the base case, though no data are provided to substantiate this.

The EAG is aware of at least one published systematic review of prognostic factors in with ER+/HER2-, locally advanced/metastatic breast cancer (Cuyún Carter et al., 2021)[21]. This review (which is not cited in the CS) included 79 studies and identified a set of prognostic factors associated with worse OS and worse PFS, based on the strongest evidence from their review. Table 48 in Appendix 2 of this EAG report lists the adverse prognostic factors identified by Cuyún Carter et al (2021)[21] alongside those proposed by key opinion leaders consulted by the company, in the style of a matrix. It can be seen that there is reasonable agreement between the Cuyún Carter review and the key opinion leaders in choice of factors, but there was also a handful of prognostic factors unique to each respective source. It is noticeable that only a minority of all these prognostic factors were included in the MAIC. Amongst the factors which were not matched due to lack of data were some of notable importance such as bone metastases / bone metastases only; number of metastatic sites and de novo vs. recurrent/progressed disease. Their omission is a key limitation of the MAIC.

3.4.3 Estimation of the weights for EMERALD patients

The CS reports brief details of the weighting process. A logistic regression model was used based "*on a similar approach to propensity score weighting*" (CS page 83).

3.4.4 Comparison of weighted-elacestrant and comparator patient characteristics

CS Table 26 gives the characteristics of elacestrant-treated patients before and after weighting compared to the characteristics of patients receiving everolimus + exemestane in Flatiron (subgroup 1). The characteristics listed are the prognostic factors identified by key opinion leaders, as discussed above (e.g. age/menopausal status; number of lines of previous endocrine therapy; prior chemotherapy)(section 3.4.2). After weighting, the effective sample size for elacestrant was reduced from 78 [redacted] patients ([redacted] of the initial sample size), compared to 32 comparator patients. Importantly, however, there are some imbalances in characteristics between the elacestrant and the everolimus + exemestane arms. For example, the percentage of elacestrant patients with ECOG 0 was twice that of comparator patients, though this is explained by missing data on ECOG status for 25% of comparator patients in Flatiron. The company adjusted for the missing data in a sensitivity analysis but did not report the adjusted distribution of patients across the known ECOG categories or the results of the sensitivity analysis, other than commenting that it had *"similar*

results observed to the base case" (CS page 85). This remains as an uncertainty in the EAG's view.

CS Table 28 gives the characteristics of elacestrant-treated patients before and after weighting compared to the characteristics of patients receiving alpelisib + fulvestrant in Flatiron (subgroup 2). After weighting, the effective sample size for elacestrant reduced from 27 [redacted] patients [redacted] of the initial sample size), compared to 33 comparator patients. Again, the missing ECOG performance status score data for 25% of patients from Flatiron meant that there were imbalances between elacestrant and comparator arms. There was also disparity between the arms for the percentage of patients who had previously received chemotherapy in the advanced/metastatic setting (higher in the comparator arm).

3.4.5 Statistical methods for the MAIC

The company reported that the MAIC was constructed following methodological guidance regarding population-adjusted indirect comparisons set out in the NICE Decision Support Unit (DSU) Technical Support Document (TSD) 14 which deals with survival analysis and extrapolation from patient level data (company response to clarification question A12). They comment that although the guidance is applicable to data from randomised trials, they applied the same principles to the observational real-world evidence. For example, they sought real-world data for patients who most closely matched the population covered by the marketing authorisation for elacestrant.

The MAIC was built using R software, and the programming code was supplied to the EAG (company response to clarification question A13).

No further detail on the statistical methods is given, aside from that mentioned above (section 3.3 and sections 3.4.1 to 3.4.5).

EAG comment on the methods for the MAIC

The MAIC was produced according to methodological guidance from the NICE Decision Support Unit (DSU) Technical Support Document (TSD) on methods for population-adjusted indirect comparisons in submission to NICE. As far as the EAG can tell from the company's description of the MAIC, the methods were implemented appropriately. However, the MAIC suffers from some key limitations. For example, the selection of prognostic factors was poorly described and many of the factors identified could not be included in the matching of EMERALD trial patients to Flatiron database patients due to lack of available data. Furthermore, following weighting, the number of patients in the

analyses was reduced, with some imbalances in weighted prognostic factors between elacestrant and the comparator, particularly evident in post hoc subgroup 2.

3.5 Results of the MAIC

3.5.1 Progression free survival (PFS)

The CS provides Kaplan Meier PFS curves from the MAIC for elacestrant (weighted and unweighted) compared to everolimus + exemestane for subgroup 1 (patients with ESR1-mutation who have disease progression following ≥12 months prior treatment with ET + CDK4/6i) (CS Figure 17). The analyses indicate [redacted] PFS for elacestrant compared to everolimus + exemestane, with separation of the survival curves evident after the first few months and remaining so for the rest of the follow-up period (approximately 30 months). Table 16 below gives the median PFS (in months) and HR from the MAIC. The HR of 0.59 (0.36 to 0.96) indicates increased PFS associated with elacestrant, and the confidence intervals do not cross 1. However, due to the methodological limitations in the MAIC, as discussed above, inferences of statistical significance should not be made.

Table 16 MAIC PFS, elacestrant versus everolimus + exemestane (subgroup 1)

Outcome	Median (95% CI)		HR[a]
	Elacestrant weighted	Everolimus + exemestane	
PFS	[redacted]	[redacted]	0.59 (0.36, 0.96)

Source: Reproduced from CS Table 27.
[a] HR elacestrant vs everolimus + exemestane
[b] Months

CS Figure 19 provides Kaplan Meier PFS curves from the MAIC for elacestrant (weighted and unweighted) compared to alpelisib + fulvestrant for subgroup 2 (dual ESR1 and PIK3CA mutation). Initially, PFS is [redacted] for alpelisib + fulvestrant until around month 6, when the [redacted] and the [redacted]. For much of the remaining follow-up period (approximately 30 months) the curves [redacted] several times. Table 17 below gives the median PFS and HR from the MAIC. The confidence intervals are wide, notably so for the HR of 1.05 (0.50, 2.20) suggesting much uncertainty in the treatment effect. The CS describes the PFS results as [redacted] between elacestrant and alpelisib + fulvestrant. The EAG notes that they do appear [redacted], but there is insufficient evidence to conclude [redacted] between the treatments. Caution is advised in the interpretation of the results due to the methodological limitations of this analysis, as we have discussed above (section 3.3 and section 3.4).

Table 17 MAIC PFS, elacestrant versus alpelisib + fulvestrant (subgroup 2)

Outcome	Median (95% CI)		HR[a]
	Elacestrant weighted	Alpelisib + fulvestrant	
PFS	[redacted]	[redacted]	1.05 (0.50, 2.20)

Source: Reproduced from CS Table 29.
[a] HR elacestrant vs everolimus + exemestane
[b] Months

3.5.2 Overall survival (OS)

CS Figure 16 provides Kaplan Meier OS curves from the MAIC for elacestrant (weighted and unweighted) compared to everolimus + exemestane for subgroup 1 (patients with ESR1-mutation who have disease progression following ≥12 months prior treatment with ET + CDK4/6i). The curves indicate [redacted] OS for elacestrant until around month 34 when the curves cross, indicating violation of the proportional hazards assumption. Table 18 below gives median OS (in months) and HR from the MAIC, which indicate increased OS associated with elacestrant. However, due to the methodological limitations in the MAIC, as discussed above, inferences of statistical significance should not be made.

Table 18 MAIC OS, elacestrant versus everolimus + exemestane (subgroup 1)

	Median (95% CI)		HR[a]
	Elacestrant weighted	Everolimus + exemestane	
OS	[redacted]	[redacted]	0.64 (0.35, 1.16)

Source: Reproduced from CS Table 27.
[a] HR elacestrant vs everolimus + exemestane
[b] Months
[c] The CS defines NR as “not reported”, but the EAG suggests this is an error and that NR in this context should mean “not reached”

CS Figure 18 provides Kaplan Meier OS curves from the MAIC for elacestrant (weighted and unweighted) compared to alpelisib + fulvestrant for subgroup 2 (dual ESR1 and PIK3CA mutation). After around 12 months the curves separate, indicating greater OS for elacestrant, before overlapping again after month 30. Due to the overlapping curves the proportional hazards assumption cannot be supported. Table 18 below gives median OS (in months) and HR from the MAIC, which indicate a small increase in OS associated with elacestrant. However, due to the methodological limitations in the MAIC, as discussed above, inferences of statistical significance should not be made.

Table 19 MAIC OS, elacestrant versus alpelisib + fulvestrant (subgroup 2)

	Median (95% CI)		HR[a]
	Elacestrant weighted	Alpelisib + fulvestrant	
OS	[redacted]	[redacted]	0.80 (0.33, 1.92)

Source: Reproduced from CS Table 27.
[a] HR elacestrant vs everolimus + exemestane
[b] Months
NR, Not reached

4 COST EFFECTIVENESS

4.1 EAG comment on company's review of cost-effectiveness evidence

The company conducted a combined search for health economic literature, including cost-effectiveness studies and estimates of health-related quality of life (HRQoL), resource use and costs. We consider that the search strategy was appropriate but note that the searches are out of date as the latest update search was conducted in April 2023 (CS B.3.1 and Appendix G). One cost-effectiveness study was included in the company's review; the analysis conducted for the NICE technology appraisal of alpelisib with fulvestrant for HR+, HER2-negative, PIK3CA-mutated advanced breast cancer (TA816, 2022).[8] The company argues that TA816 is the most relevant previous NICE appraisal and outlines key features of the TA816 economic analysis in CS Table 38 and Appendix Table 26.

The EAG conducted targeted searches in PubMed and Google scholar and identified two recent economic studies that included elacestrant:

- **Vidal et al. 2023** estimated the number of clinical and resource use events associated with treating patients with elacestrant rather than standard care over a three-year time horizon.[22] We do not consider this study further as it is not an economic evaluation, and it is only reported as a conference poster with limited detail.

- **Zeng et al. 2023** reported a cost-effectiveness analysis of elacestrant versus standard endocrine therapy for second and third-line treatment of patients with advanced HR+/HER2- advanced breast cancer from a US payer perspective.[23] This used a partitioned-survival model with survival curves fitted to digitised Kaplan-Meier data from the EMERALD trial, similar to the company's approach. However, the results are not comparable due to differences in the study populations and comparators. Zeng et al. estimated cost-effectiveness for the whole EMERALD trial population and the subgroup with ESR1 mutation and used the 'investigator's choice' control arm from EMERALD and fulvestrant alone as comparators.

4.2 Summary and critique of the company's submitted economic evaluation

4.2.1 NICE reference case checklist

The company list key features of their analysis in CS B.3.2.2.3. The EAG considers that the company's analysis is consistent with the NICE reference case (see Table 20 below).[24] We note two potential areas of confusion in the company's reporting of base case cost-effectiveness results (CS Tables 81 and 82):

- The standard discount rate of 3.5% is applied to costs and QALYs, but not to life years gained (LYG), which is not stated in the tables or footnotes. We report discounted LYG for the company's base case in section 5.1 below.
- The company apply a decision modifier severity weight of 1.2 to the incremental QALYs and ICERs for Subgroup 1. We consider it more appropriate to first report results without the QALY weight, and then show how these results change with the weight, as it is a matter for the committee to consider whether the QALY weighting should be used. In the results sections 5 and 6 below, we report total and incremental QALYs without the severity weight, and we report ICERs both without and with the severity weight applied. We critique the company's absolute and proportional QALY shortfall calculations in section 7.

Table 20 NICE reference case checklist

Element of HTA	Reference case	Is the company analysis consistent with reference case criteria?
Perspective on outcomes	All health effects, whether for patients or, when relevant, carers	Yes (no direct health effects assumed for carers)
Perspective on costs	NHS and personal social services (PSS)	Yes
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	Yes
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Yes (lifetime horizon)
Synthesis of evidence on health effects	Based on systematic review	Yes

Element of HTA	Reference case	Is the company analysis consistent with reference case criteria?
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults.	Yes
Source of data for measurement of health-related quality of life	Reported directly by patients or carers, or both	Yes (EQ-5D-5L data from EMERALD trial). See section 4.2.5.2 below.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	Yes (UK tariff, Hernández-Alava formula)[25]
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit, except in specific circumstances	Yes (QALY weight of 1.2 applied for Subgroup 1. No QALY weight applied to the dual mutated subgroup) See Section 7 below.
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	Yes
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	Yes, for costs and QALYs (no discounting applied to LYs reported in the CS)

Source: Produced by the EAG based on information in CS section B.3 and Table 38

4.2.2 Model decision problem

4.2.2.1 Population

The company reports cost-effectiveness results for two subgroups:

- Subgroup 1 (target population): *ESR1-mut* + ≥12 months prior ET with CDK4/6i
- Subgroup 2 (dual mutated): *ESR1-mut+PIK3CA-mut* + ≥12 months ET with CDK4/6i

The company's target population for elacestrant is restricted to the subgroup of the licensed population with disease progression after at least 12 months of endocrine therapy with CDK4/6 inhibitors. They state that this will provide best value in UK clinical practice (CS B.3.2.1), based on clinical feedback informed by the post hoc subgroup analyses of EMERALD trial data by duration of prior treatment (Bardia 2023, Menarini 2024).[9] [14] See section 3.2.5.4 above for the EAG description and critique of this subgroup analysis.

Baseline characteristics for these subgroups in the EMERALD trial and Flatiron cohorts are reported in CS Tables 20 and 26, respectively. For the base case economic analysis, the company used patient characteristics from EMERALD for both subgroups: [redacted] with mean ages [redacted] years for subgroup 1 and [redacted] years for subgroup 2 (CS Table 39). In response to clarification question B2, the company added a scenario with baseline patient characteristics from the Flatiron cohorts: [redacted] years for subgroup 1 and [redacted] for subgroup 2 (CQ response Table 5). This gave a small increase in the ICER for subgroup 1 and had a negligible impact on cost-effectiveness for subgroup 2.

4.2.2.2 Intervention and comparators

The modelled intervention is elacestrant at 345 mg orally, once daily (CS B.3.2.3.1). To account for dose interruptions and modifications in the economic model, elacestrant costs are adjusted with a relative dose intensity (RDI) estimated from the EMERALD trial (see CS B.3.5.1.1 and section 4.2.6.1 below).

The company include one comparator for each subgroup in their economic model, based on clinical advice (CS section B.3.2.3.2) that these are the most relevant current treatments in the subgroups of interest: everolimus + exemestane for the target population (subgroup 1); and alpelisib + fulvestrant for the dual mutated subgroup (subgroup 2). Other comparators specified in the NICE scope (endocrine therapy with or without chemotherapy, and chemotherapy alone) are excluded on the basis that these are rarely used in practice for the target population. Data from the control arm of the EMERALD trial (investigator's choice of fulvestrant, anastrozole, letrozole, or exemestane monotherapy) is therefore not used in the economic model. As there is no direct evidence for the effectiveness of elacestrant versus everolimus + exemestane or alpelisib + fulvestrant, and the pivotal trials for these treatments did not include the subgroups of interest (so a network meta-analysis is not feasible), the company rely on data from the Flatiron cohorts and the unanchored MAIC (CS B.2.9) to estimate survival outcomes for the economic model.

EAG conclusions on the modelled decision problem

The economic model reflects the company's target population for elacestrant, and the subgroup with dual mutations as requested in the NICE scope. As the model relies on MAIC-adjusted survival outcomes, with trial data weighted to reflect baseline prognostic factors in the Flatiron cohorts, the EAG prefers the analysis with mean ages at baseline from the Flatiron cohorts (CS Tables 20 and 26).

The EAG agrees that the focus on the comparators everolimus + exemestane for subgroup 1 and alpelisib + fulvestrant for subgroup 2 is reasonable, although endocrine therapy with or without chemotherapy, or chemotherapy alone may be used for some patients (see discussion in sections 2.2.3.2 and 2.3 above).

4.2.2.3 Perspective, time horizon and discounting

The analysis is in line with the NICE Reference Case with respect to the perspective (NHS and PSS); time horizon (lifetime); and discounting (3.5% applied to costs and QALYs).

4.2.3 Model structure and assumptions

4.2.3.1 Overview of the model structure

The company describe the structure of their economic model in CS section B.3.2.2. They use a cohort-level partitioned survival analysis (PartSA), implemented in Microsoft Excel (see CS Figures 20 and 21). The model has a one week cycle length and a lifetime horizon. A summary of model assumptions is provided in CS Table 80, and a list of the base case model parameters and probabilistic distributions in CS Table 79.

The distribution of the modelled cohort between health states is determined by survival curves fitted to time to treatment discontinuation (TTD), progression-free survival (PFS) and overall survival (OS) data from the EMERALD trial for elacestrant, and to KM curves from the Flatiron dataset for the comparators. The MAIC approach described in section 3.4 above is used to weight the data for the elacestrant arm of the EMERALD trial to improve alignment with baseline prognostic characteristics in the Flatiron cohorts (see CS Tables 26 and 28 for subgroup 1 and 2, respectively).

The model includes constraints to ensure that:

- The proportion of patients on treatment cannot exceed progression-free survival;
- The proportion who are progression-free cannot exceed overall survival; and
- The risk of death is no lower than for people of the same age and sex in the general population.

We critique the model structure and key assumptions in the following section. See section 4.2.4 below for EAG critique of the fitted TTD, PFS and OS extrapolations. Other model parameters include health-related quality of life for the progression-free and progressed disease states (section 4.2.5), and resource use and costs (section 4.2.6).

4.2.3.2 EAG critique of model structure and assumptions

The partitioned survival analysis (PartSA) modelling approach is common in cancer appraisals and provides a practical alternative to a health-state transition model when data to estimate transition probabilities is sparse. However, as described in NICE Decision Support Unit Technical Support document 19, PartSA requires two key assumptions: that the survival endpoints (TTD, PFS and OS) can be modelled and extrapolated independently; and that trends in the hazards of these endpoints from the study period persist over the time horizon.[26] The risk of bias due to these assumptions is mitigated to some extent in the company's model by the constraints applied to ensure that TTD ≤ PFS, PFS ≤ OS and the risk of mortality is no less than for people of the same age in the general population. However, careful consideration of the clinical plausibility of the survival curve extrapolations is still essential. See section 4.2.4 below for discussion on the methods used to fit TTD, PFS and OS curves for elacestrant and comparators, and the plausibility of the extrapolations.

As there is no direct evidence to compare elacestrant with everolimus + exemestane or alpelisib + fulvestrant in the company's target population and the dual-mutated subgroup, the model relies on an unanchored MAIC for estimation of survival outcomes. The economic model results are therefore vulnerable to bias from the MAIC due to the lack of data on identified prognostic factors and effect modifiers (CS Table 25 and EAG discussion in 3.4). There is also considerable uncertainty around the survival curves due to the small sample sizes for both subgroups of interest in the Flatiron datasets, and also from the elacestrant arm of the EMERALD trial (particularly for the dual mutated subgroup).

The lack of data on treatment duration in the Flatiron datasets for the comparator arms is also problematic. The company use observed data from the EMERALD trial for elacestrant but assume that TTD is equal to PFS for the comparators. It is quite common in cancer appraisals to assume that treatment continues until disease progression, and this is often reasonable. However, the use of different assumptions for the intervention and comparator is a potential source of bias, that would have a direct impact on costs and hence on the ICER. The elacestrant trial data used in the model also shows a difference between TTD and PFS, with a proportion of patients in the subgroups of interest stopping treatment before

progression. Consideration of alternative sources of data or assumptions regarding the duration of treatment for the comparators is therefore important.

Other model assumptions that are potentially important are the cost and practical impact of introducing ESR1 testing the NHS to assess suitability for elacestrant, and the mix of subsequent treatments that are used in NHS practice after disease progression.

EAG conclusions on the model structure and assumptions

- We consider that the use of a partitioned survival model is appropriate, and that the implemented model is of a high standard.
- However, we do have concerns about the robustness and plausibility of the PFS and OS extrapolations due to the reliance on an unanchored MAIC and the sparsity of data for the company's target population and the dual mutated subgroup from the EMERALD trial and the Flatiron cohorts.
- We are also concerned over the lack of data on treatment duration for the comparators, and the potential for bias from the company's assumption that treatment will always continue until disease progression in the comparator arms, whereas treatment with elacestrant can stop prior to progression (as observed in the EMERALD trial).
- We conduct additional scenario analyses to explore alternative assumptions regarding these concerns, as well as other uncertainties, including the cost of introducing ESR1 testing and NHS practice regarding subsequent treatment.

4.2.4 Clinical effectiveness and extrapolation

4.2.4.1 Overview of methods for extrapolation of survival outcomes

The economic model uses parametric survival curves for PFS, OS and TTD in the two subgroups, which are fitted to patient-level data from the EMERALD trial for elacestrant and to pseudo patient-level data derived from KM curves for the Flatiron comparator cohorts (CS B.3.3.4).[27] MAIC weights are applied to the elacestrant patient-level data to better align prognostic characteristics with those in the Flatiron cohorts (CS B.2.9.1).

The company report results for six standard parametric survival distributions (exponential, generalised gamma, Gompertz, log-logistic, log-normal and Weibull and gamma). Alternative flexible survival models are not explored. The base case distribution in each case was chosen on the basis of fit to the KM estimates, using visual inspection and Akaike and Bayesian information criteria (AIC, BIC) statistics, and consideration of the clinical plausibility

of the long-term extrapolations. The company do not report formal elicitation of survival expectations from clinical experts.

In the base case, OS and PFS curves are fitted to each dataset independently, on the grounds that the proportional hazards assumption 'may not hold' due to crossover of the elacestrant and comparator KM curves: see CS Figures 16 and 17 for subgroup 1, and Figures 18 and 19 for subgroup 2. Formal tests of proportional hazards are not reported. The company report scenario analysis with parametric OS and PFS curves fitted to the MAIC-weighted trial data for elacestrant, which are then adjusted for the comparator arms using MAIC hazard ratios (CS Tables 27 and 29).

TTD data from the EMERALD trial is mature (CS Figure 26). So for elacestrant, the company use the KM curves directly in the base case, and parametric curves fitted to the MAIC-weighted EMERALD data in scenario analysis. However, data was not available to estimate TTD for the comparator arms, as the Flatiron datasets do not include treatment duration. The company considered estimating comparator TTD from median treatment duration but could not find this reported in the literature for the particular subgroups of interest. The company therefore made an assumption, setting TTD equal to PFS for the comparator arms. The model includes an option to estimate comparator TTD by applying an assumed hazard ratio to the PFS but did not report scenario analysis using this option.

We discuss the company's assumptions and selection of survival extrapolations for their base case and scenarios below.

4.2.4.2 Survival curves for subgroup 1

CS Figures 16 and 17 show the unweighted and MAIC-weighted OS and PFS KM plots for elacestrant and everolimus + exemestane in subgroup 1. The sample size for this subgroup is moderate for elacestrant (n=78; effective sample size after MAIC adjustment n=[redacted]) and very low for everolimus + exemestane (n=32) (CS Table 26). There is therefore high uncertainty over the KM estimates, particularly for the comparator and in the later sections of follow up, as the numbers of patients at risk and the number of events are low.

The company discuss their choice of OS, PFS and TTD distributions for subgroup 1 in CS section B.3.3.4.1. We show survival extrapolation graphs for this subgroup in Appendix 3: see Figure 17 and Figure 16 for the company's base case extrapolations for elacestrant and everolimus + exemestane respectively.

4.2.4.2.1 Overall survival

Overall survival estimates and model fit statistics for the six parametric distributions in subgroup 1 are summarised in Table 21 below.

Table 21 OS extrapolations: subgroup 1

Distribution	Model fit			Survival estimates (year)				
	AIC	BIC	Rank	1	2	3	5	10
Everolimus + exemestane								
Kaplan-Meier	-	-	-	62.3%	37.5%	28.1%	14.1%	-
Exponential	173.17	174.63	1	63.7%	40.3%	25.7%	10.4%	1.1%
Gen. gamma	176.57	180.97	7	63.4%	40.3%	27.0%	13.3%	3.1%
Gompertz	175.10	178.03	5	62.7%	40.2%	26.7%	12.7%	2.9%
Log-logistic	174.32	177.25	2	62.3%	38.6%	26.6%	15.3%	6.5%
Log-normal	175.23	178.16	6	61.2%	40.2%	29.0%	17.4%	7.1%
Weibull	175.10	178.03	4	64.6%	40.1%	24.7%	9.2%	0.7%
Gamma	175.01	177.94	3	64.8%	39.8%	24.4%	9.0%	0.7%
Elacestrant (weighted to everolimus + exemestane)								
Kaplan-Meier	-	-	-	86.6%	51.6%	14.7%	-	-
Exponential	342.10	344.45	7	74.3%	54.8%	40.7%	22.5%	5.0%
Gen. gamma	334.16	341.23	5	83.8%	55.3%	26.8%	1.3%	0.0%
Gompertz	332.93	337.64	2	83.9%	56.6%	24.3%	0.1%	0.0%
Log-logistic	334.04	338.75	4	83.5%	54.6%	34.5%	15.7%	4.3%
Log-normal	337.04	341.75	6	80.5%	54.3%	37.4%	19.3%	5.4%
Weibull	332.50	337.21	1	83.8%	54.7%	29.6%	5.2%	0.0%
Gamma	333.35	338.06	3	82.8%	54.4%	32.4%	9.8%	0.3%

Source: Table collated by the EAG from CS Tables 40, 41, 46 and 47 and the company's model
Company base case distributions in bold

For everolimus + exemestane, the parametric distributions have a similar visual and statistical fit to the Flatiron KM data. Survival estimates are similar over the first 2 years, but there is then some divergence (see Figure 11 below). The distribution with the best statistical fit is the exponential (constant hazard), but the company select the gamma for their base case (the third best statistical fit), on the basis that this has the lowest 5-year survival (9%), which is closest to clinical expectations. This assessment is based on a clinical estimate of 5% five-year survival for patients with HR+, HER2- PIK3CA-mutated advanced breast cancer treated with everolimus + exemestane, as reported in the alpelisib company submission for NICE appraisal TA816.[8] We note that the alpelisib company also reported

clinical estimates of 50% and 33.3% survival 1 and 2 years, respectively (TA816 EAG report). Survival estimates from the Flatiron KM and parametric distributions in Table 21 all exceed these expectations. It is not clear if this relates to differences in the populations under consideration, and/or to other differences between the data sources.

For elacestrant, with the exception of the exponential, the visual and statistical fits and survival estimates at 1 and 2 years are similar for the different parametric distributions. However, there is a wide range of survival projections at 3 and 5 years. The Weibull gives the best statistical fit, but the company conclude that the log-normal and log-logistic curves have a good visual fit to the KM data (see Figure 12). They also argue that they expect the superiority of elacestrant over everolimus + exemestane over the first 2.5 years of follow-up to persist at 5 years. On this basis, they select the log-logistic distribution for their base case.

Figure 2below shows the OS KM estimates and fitted distributions used in the company's base case: log-logistic for elacestrant and gamma for the comparator arm. The company also report results for scenario analyses with Weibull and exponential OS extrapolations for everolimus + exemestane and gamma and log-normal extrapolations for elacestrant in subgroup 1.

EAG conclusions on OS extrapolations for subgroup 1:

- There is high uncertainty over the OS extrapolations due to the limited sample sizes (particularly for the comparator arm) and the use of an unanchored MAIC.
- We agree with the use of the gamma distribution for the comparator arm based on clinical advice on current survival expectations in this subgroup.
- Expert advice to the EAG is that 5-year survival with current treatment in this population is likely to be around 5%, and that although there may well be a small proportion of patients who gain a long-term benefit with elacestrant, this is as yet untested. We therefore conclude that the company's base case log-logistic OS extrapolation for elacestrant is overly optimistic given the current evidence base.
- For EAG analysis, we prefer to use an independent gamma OS extrapolation for elacestrant as well as for the comparator arm (Figure 3, below). The gamma has a good statistical and visual fit in both arms and similar survival projections after year 5.
- To test the impact of a wider range of OS extrapolations, we also report additional EAG scenarios using the MAIC HR option in the company's model (see 6.1.1).

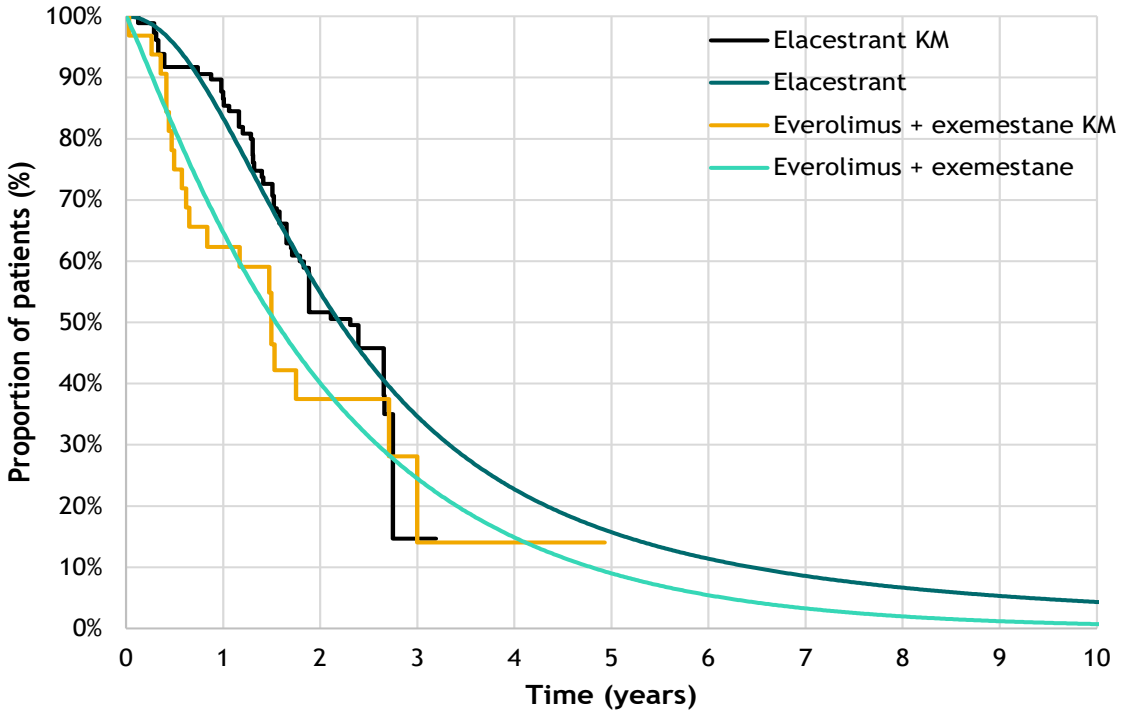

Figure 2 OS extrapolations for the company's base case: subgroup 1

Source: produced by the EAG from the company's model

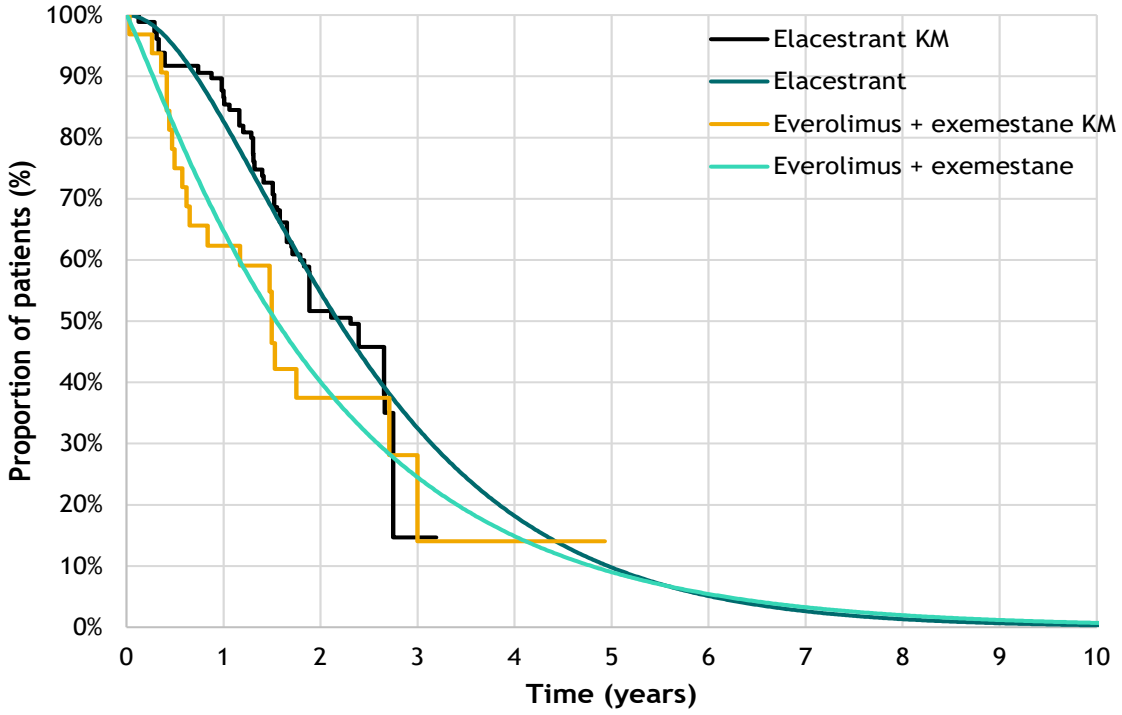

Figure 3 OS extrapolations, independent gamma for both arms: subgroup 1

Source: produced by the EAG from the company's model
Gamma extrapolation for elacestrant and everolimus + exemestane

4.2.4.2.2 Progression free survival

Statistical measures of fit and survival estimates for PFS extrapolations for subgroup 1 are summarised in Table 22.

Table 22 PFS extrapolations: subgroup 1

Distribution	Model fit			Survival estimates (year)				
	AIC	BIC	Rank	1	2	3	5	10
Everolimus + exemestane								
Kaplan-Meier	-	-	-	14.6%	-	-	-	-
Exponential	150.53	151.99	7	12.5%	1.5%	0.2%	0.0%	0.0%
Gen. gamma	146.20	150.60	5	7.2%	0.4%	0.0%	0.0%	0.0%
Gompertz	148.74	151.67	6	5.9%	0.0%	0.0%	0.0%	0.0%
Log-logistic	144.20	147.14	1	8.4%	1.8%	0.7%	0.2%	0.0%
Log-normal	144.84	147.77	3	9.0%	1.2%	0.3%	0.0%	0.0%
Weibull	145.69	148.62	4	5.5%	0.0%	0.0%	0.0%	0.0%
Gamma	144.62	147.55	2	5.9%	0.1%	0.0%	0.0%	0.0%
Elacestrant (weighted to everolimus + exemestane)								
Kaplan-Meier	-	-	-	34.3%	29.3%	-	-	-
Exponential	250.31	252.67	4	37.0%	13.4%	5.0%	0.7%	0.0%
Gen. gamma	212.37	219.44	1	31.1%	20.7%	16.5%	12.3%	8.3%
Gompertz	250.63	255.34	5	36.2%	18.4%	11.9%	7.4%	5.4%
Log-logistic	245.92	250.64	3	30.8%	14.2%	8.6%	4.4%	1.7%
Log-normal	242.02	246.73	2	32.2%	14.2%	7.8%	3.1%	0.7%
Weibull	252.31	257.02	7	37.1%	13.6%	5.1%	0.7%	0.0%
Gamma	252.13	256.84	6	36.4%	12.3%	4.2%	0.5%	0.0%

Source: Table collated by the EAG from CS Tables 42, 43, 48 and 49 and the company's model
Company base case distributions in bold

For everolimus + exemestane, all distributions give a similar fit to the KM, with the exception of exponential and Gompertz. Projected progression free survival is similar for the remaining distributions, with some patients remaining progression free at 3 years with the log-logistic and log-normal. The company select the log-normal distribution for their base case, and use log-logistic and gamma for scenario analysis.

For elacestrant, the best statistical fit is the generalised gamma, although this has a poor visual fit after the first few months and a very optimistic long-term projection (over 12% still progression free at 5 years). The log-normal and log-logistic have similar statistical and

visual fit and give similar long-term projections. The company use the log-normal for their base case, and log-logistic in a scenario.

Figure 4 PFS extrapolations for the company's base case: subgroup 1

Source: produced by the EAG from the company's model
Log-logistic extrapolation for elacestrant and log-normal for everolimus + exemestane

EAG conclusions on PFS extrapolations for subgroup 1:

- The company's base case PFS extrapolations for subgroup 1 are reasonable. We also test scenarios with Weibull for everolimus + exemestane, and exponential for elacestrant (see 6.1.1).

4.2.4.2.3 Time to treatment discontinuation

As data on time to discontinuation of elacestrant in the EMERALD trial is mature, the company used the KM curve directly in the economic model. Fitted parametric curves were included in the economic model for use in scenario analysis. See Table 23 for a summary of fit statistics and treatment continuation rates for subgroup 1. The company report results for scenarios using log-normal and log-logistic distributions for elacestrant TTD. We note that, compared with the KM estimates, all of the parametric extrapolations underestimate the proportion of patients still on elacestrant at 2 years.

Table 23 Elacestrant TTD: subgroup 1

Distribution	Model fit			Survival estimates (year)				
	AIC	BIC	Rank	1	2	3	5	10
Kaplan-Meier	-	-	-	[redacted]	[redacted]	-	-	-
Exponential	455.13	457.48	5	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gen. gamma	431.34	438.41	1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	453.91	458.62	4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	442.37	447.08	3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	438.63	443.34	2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	456.89	461.61	6	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gamma	457.03	461.74	7	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: Table collated by the EAG from CS Tables 44, 45 and the company's model
Company base case distributions in bold

Treatment discontinuation data is not available for the comparator everolimus + exemestane arm. The company make an assumption that for everolimus + exemestane TTD is equal to PFS. This results in broadly similar TTD curves for the two arms in the company's base case in this subgroup (see Figure 5).

EAG conclusion for TTD extrapolations in subgroup 1

We agree with the use of KM data from the EMERALD trial rather than a fitted extrapolation to estimate time to treatment duration for elacestrant. As the data is mature, this will provide the best available estimate. To further explore sensitivity to treatment duration for elacestrant, we report an additional scenario using the best-fit extrapolation (generalised gamma) for elacestrant TTD (see section 6.1.1).

We are concerned about the potential for bias due to the use of different modelling assumptions for TTD in the elacestrant and comparator arms. In practice, it is likely that some patients in the comparator arm may discontinue treatment prior to progression, as was observed for elacestrant. If so, this will result in over-estimation of treatment costs for the comparator relative to elacestrant. We explore the impact of such an effect using the option provided in the company's model to apply a hazard ratio to reduce TTD relative to PFS in the comparator arm.

Figure 5 TTD extrapolations for the company's base case: subgroup 1

Source: produced by the EAG from the company's model
KM and fitted generalised gamma distribution for elacestrant; and base case fitted distribution for PFS (log-normal) assumed for everolimus + exemestane

4.2.4.3 Survival curves for subgroup 2

CS Figures 18 and 19 show the unweighted and MAIC-weighted OS and PFS KM plots for elacestrant and alpelisib + fulvestrant in subgroup 2. The sample size for this subgroup is very low for both elacestrant (n=27; effective sample size n=[redacted]) and alpelisib + fulvestrant (n=33) (CS Table 26), so there is very high uncertainty over the KM estimates.

The company discuss their choice of OS, PFS and TTD distributions for subgroup 2 in CS section B.3.3.4.2. We show survival extrapolation graphs for this subgroup in Appendix 4. Figure 24 and Figure 23 show the company's base case extrapolations for alpelisib + fulvestrant and elacestrant respectively.

4.2.4.3.1 Overall survival

Table 24 summarises statistical measures of fit and survival estimates for OS in subgroup 2. The company report that the generalised gamma distribution did not converge. The exponential distribution has the worst statistical fit and poor visual fit to the KM in both arms (Figure 18 and Figure 19).

For alpelisib + fulvestrant, the Gompertz also has a relatively poor statistical and visual fit. The gamma distribution has the best statistical fit and a good visual fit to the KM. The other distributions all have a similar statistical and visual fit. The company chose the gamma distribution for their base case, and the Weibull and log-normal for scenario analysis.

For elacestrant, the best statistical fit is the Gompertz, but the company conclude that this is unrealistic, as it predicts [redacted]. The other distributions have a similar statistical fit, and the company chose the Weibull for their base case, and the gamma and log-normal for scenario analysis.

Table 24 OS extrapolations: subgroup 2

Distribution	Model fit			Survival estimates (year)				
	AIC	BIC	Rank	1	2	3	5	10
Alpelisib + fulvestrant								
Kaplan-Meier	-	-	-	84.7%	55.1%	34.4%	-	-
Exponential	126.69	128.18	6	76.4%	58.1%	44.4%	26.0%	6.7%
Gen. gamma	-	-	-	-	-	-	-	-
Gompertz	123.71	126.71	5	85.3%	61.4%	31.8%	0.6%	0.0%
Log-logistic	122.44	125.43	4	86.1%	56.0%	34.0%	14.1%	3.3%
Log-normal	122.33	125.32	2	84.8%	55.3%	35.3%	15.4%	2.9%
Weibull	122.33	125.32	3	86.5%	58.3%	31.8%	5.2%	0.0%
Gamma	122.14	125.13	1	86.1%	56.8%	32.7%	8.6%	0.2%
Elacestrant (weighted to everolimus + exemestane)								
Kaplan-Meier	-	-	-	88.8%	73.6%		-	-
Exponential	90.62	91.92	7	83.2%	68.9%	57.3%	39.7%	15.7%
Gen. gamma	-	-	-	-	-	-	-	-
Gompertz	88.00	90.59	1	92.5%	73.4%	37.7%	0.0%	0.0%
Log-logistic	89.17	91.76	4	91.9%	70.9%	50.3%	24.9%	6.8%
Log-normal	89.65	92.24	5	90.4%	69.7%	52.5%	30.6%	9.9%
Weibull	88.61	91.20	2	92.1%	71.0%	46.1%	11.4%	0.0%
Gamma	88.96	91.56	3	91.4%	70.4%	49.1%	20.0%	1.3%

Source: Table collated by the EAG from CS Tables 50, 51, 56 and 57 and the company's model
Company base case distributions in bold

Figure 6 OS extrapolations for the company's base case: subgroup 2

Source: Source: produced by the EAG from the company's model
Weibull extrapolation for elacestrant and Gamma for everolimus + exemestane

EAG conclusions on OS extrapolations for subgroup 2:

- We agree with the company's base case OS extrapolations of gamma for alpelisib + fulvestrant and Weibull for elacestrant in subgroup 2 (Figure 6).
- We report additional EAG scenario analyses with Gompertz, Weibull and Gamma distributions and the MAIC HR option (see 6.1.1).

4.2.4.3.2 Progression free survival

See Table 25 for a summary of model fit statistics and survival estimates for PFS in subgroup 2. The best fit for the alpelisib + fulvestrant is log-normal, followed by generalised gamma, gamma and log-logistic distributions. These distiributions provide a reasonable visual fit to the KM, and similar PFS projections. The company choose the log-normal for their base case and report scenarios with generalised gamma and gamma distributions.

The KM estimates for elacestrant are more uncertain, due to the small sample and number of observed progression events in this subgroup. The best statistical fit is the generalised gamma, but this has a poor visual fit. The company select the log-normal for their base case,

which has a good statistical and visual fit, and report scenarios with log-logistic and exponential extrapolations.

Table 25 PFS extrapolations: subgroup 2

Distribution	Model fit			Survival estimates (year)				
	AIC	BIC	Rank	1	2	3	5	10
Alpelisib + fulvestrant								
Kaplan-Meier	-	-	-	30.2%	5.0%			-
Exponential	163.80	165.29	7	27.8%	7.6%	2.1%	0.2%	0.0%
Gen. gamma	156.23	160.72	2	21.2%	5.0%	1.9%	0.5%	0.1%
Gompertz	161.48	164.47	6	28.5%	1.4%	0.0%	0.0%	0.0%
Log-logistic	156.73	159.72	4	20.3%	4.6%	1.8%	0.5%	0.1%
Log-normal	154.52	157.51	1	21.0%	3.5%	0.8%	0.1%	0.0%
Weibull	157.98	160.97	5	24.9%	1.5%	0.0%	0.0%	0.0%
Gamma	156.42	159.41	3	22.7%	1.8%	0.1%	0.0%	0.0%
Elacestrant (weighted to everolimus + exemestane)								
Kaplan-Meier	-	-	-	21.1%	-	-	-	-
Exponential	84.72	86.01	4	30.7%	9.2%	2.8%	0.3%	0.0%
Gen. gamma	73.32	77.20	1	21.5%	12.4%	9.0%	6.1%	3.5%
Gompertz	86.66	89.25	7	30.6%	10.5%	4.2%	0.9%	0.1%
Log-logistic	84.16	86.75	3	23.0%	8.5%	4.6%	2.0%	0.7%
Log-normal	82.84	85.43	2	24.3%	8.0%	3.5%	1.0%	0.1%
Weibull	86.46	89.05	6	29.8%	7.2%	1.6%	0.1%	0.0%
Gamma	86.06	88.65	5	28.4%	6.2%	1.3%	0.1%	0.0%

Source: Table collated by the EAG from CS Tables 52, 53, 58 and 59 and the company's model
Company base case distributions in bold

EAG conclusions on PFS extrapolations for subgroup 2

- We consider the company's choice of log normal PFS extrapolations for both arms in subgroup 2 (Figure 7) to be reasonable.

Figure 7 PFS extrapolations for the company's base case: subgroup 2

Source: Produced by the EAG from the company's model
Log-normal extrapolations for elacestrant and alpelisib + fulvestrant

4.2.4.3.3 Time to treatment discontinuation

Table 26 summarises fit statistics and survival estimates for elacestrant TTD for subgroup 2.

As the data are mature, the company use the KM curve directly in the base case analysis.

They also report scenarios with log-normal and log-logistic extrapolations.

Table 26 Elacestrant TTD: subgroup 2

Distribution	Model fit			Survival estimates (year)				
	AIC	BIC	Rank	1	2	3	5	10
Kaplan-Meier	-	-	-	[redacted]	[redacted]	-	-	-
Exponential	121.56	122.78	5	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gen. gamma	108.23	111.89	1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	120.73	123.17	4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	111.60	114.04	2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	113.09	115.53	3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	123.32	125.76	6	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gamma	123.48	125.91	7	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: Table collated by the EAG from CS Tables 54 and 55 and the company's model
Company base case distributions in bold

Due to the lack of data on treatment duration for alpelisib + fulvestrant in this subgroup, the company assume that TTD is equal to PFS (Figure 8).

Figure 8 TTD extrapolations for the company's base case: subgroup 2

Source: Produced by the EAG from the company model
KM and fitted generalised gamma distribution for elacestrant; and base case fitted distribution for PFS (log-normal) assumed for alpelisib + fulvestrant

EAG conclusion for TTD extrapolations in subgroup 2

As in subgroup 1, we agree with the direct use of the mature KM data from EMERALD to model treatment duration for elacestrant. However, the assumption that TTD is equal to PFS for the comparator arm in subgroup 2 results in a longer treatment duration for alpelisib + fulvestrant than for elacestrant, despite elacestrant having a longer projected time to progression. This is counterintuitive and we explore the use of a hazard ratio to reduce TTD relative to PFS in the comparator arm (see section 6.1.1).

4.2.5 Health related quality of life

4.2.5.1 Systematic literature review for utilities

The company conducted a systematic review to identify HRQoL utility data for patients with breast cancer (CS Appendix H). The searches were performed between January 2010 and April 2023, and the inclusion criteria are shown in CS Appendix H Table 27 and CS Appendix H Figure 9 (PRISMA diagram).

Eight studies were identified and summarised in CS Appendix H Table 28. These studies provided the health state utilities and AE disutilities used in the company's scenario analysis. Three studies referred to metastatic breast cancer: Hagiwara et al. 2018[28] conducted in Japan; Mistry et al. 2018[29] conducted in the USA; and Lloyd et al. 2006[30] conducted in the UK. The economic evaluation presented by Zeng et al. 2023 [23] (see section 4.1) used the progression-free state utility from Mistry et al. 2018 (0.837, range 0.753-0.921) and progressed disease state utility from Lloyd et al. 2006 (0.443, range 0.399-0.487), although EMERALD trial results were used to develop their model.

4.2.5.2 Study-based health related quality of life

Patients in the EMERALD trial were asked to complete the EQ-5D-5L questionnaire at study baseline, during treatment cycles, and at post-treatment, end of trial and safety follow-up assessments. EQ-5D-5L data for EMERALD patients with an ESR1 mutation were used to estimate health state utilities in the company's base case analysis (see CS B.3.4.1 and B.3.4.2, and company response to clarification question B4). We note that the utility analysis was not restricted to the company's specific target population for elacestrant (subgroup 1, ESR1-mut with at least 12 months of prior ET + CDK4/6i) and did not differentiate between subgroup 1 and the dual mutation subgroup 2.

In response to clarification question B4, the company provided further information about the data, methods and results of the utility analysis. Data for 187/228 (82%) of patients from the EMERALD trial with an ESR1 mutation were included: 222 were considered in the data preparation stage, 35 of whom were excluded due to missing data (company clarification response Table 6). The company used the NICE recommended Hernández-Alava et al. algorithm to map from EQ-5D-5L data to EQ-5D-3L UK utility values.[25]

The data were analysed using a linear mixed-effects regression to account for repeated observations (the dataset included 886 EQ-5D-5L observations from 187 patients). The utility regression model estimated the relationship between the EQ-5D utility score, progression status, concurrent adverse events, three baseline co-variates (age, utility and

number of prior lines of therapy), and patient ID as the random effect term (company clarication response Table 7). The company note that they also considered including treatment arm, but that this was likely to be correlated with adverse events.

Simple descriptive statistics for utility by health state are reported in CS Table 61. The regression coefficient estimates are reported in Table 8 of the company's clarification response, and residual plots in Figure 1. Predicted health state utilities from the regression model are reported in CS Table 62: progression-free [redacted] (95% CI: [redacted]) and progressed disease [redacted] (95% CI: [redacted]. These values are used in the company's model for the base case analysis.

4.2.5.3 Adverse event disutilities

The company considered adverse events grade 3+ with an incidence of at least 2% for elacestrant or the comparators (CS B.3.4.4). As the utility regression equation included an AE term, the company did not include AE disutilities in their base case, but they did include them in scenario analyses, applied as a one-off QALY decrement by treatment arm. CS Table 63 shows the AE frequencies and Table 64 the disutility values, durations and sources. In response to clarification questions B5 and B6, the company amended CS Table 64 with the following corrections:

- Use the correct Telford et al. 2016 [31] reference (update in the CS document B) (clarification question B5)
- Anaemia disutility reference source from Telford et al. 2019 [31] to Swinburn et al. 2010.[32] Disutility and duration values remained the same (clarification question B6a).
- Disutility value and duration for dyspnoea from Telford et al. 2019 [31] instead of considering an assumption (equal to ATL increase) (duration from 28 to 12.7 days. Disutility remained the same value) (clarification question B6b)
- Hyperglycaemia disutility value instead of hypoglycaemia value from Smith-Palmer et al. 2016 [33] (from -0.122 to -0.081) (clarification question B6c)
- Thrombocytopenia disutility from -0.110 to -0.108 (clarification question B6d)

4.2.5.4 Health state utility values used in the economic model

Health state utility values in the company's base case are taken from the EMERALD trial (CS Table 65): progression-free [redacted] and post-progression [redacted]. The company report results for a scenario using a post-progression utility of 0.601 reported by Lloyd et al. (2006)[30], included as an absolute value in combination with the progression-free utility from the EMERALD trial. The company's model also includes an option to use a relative

decrement for the post-progression state, as well as pre- and post-progression from three previous NICE appraisals TA496[34], NICE TA503[35], and TA563[36].

The utilities in the model are adjusted for general population utility values, which were taken from Ara and Brazier, 2010.[37]

Disutilities were not applied in the company's base case (section 4.2.5.3). The company presented one scenario analysis with AE disutilities. These were estimated by multiplying the disutility by the frequency and duration of the AEs. The total disutility is considered only in the first model cycle.

We summarise the sources for utility parameters in Table 27.

Table 27 Summary of utility parameters used in the economic model

Parameter	Reference	Source	Comments
Health state utility	CS Table 65	EMERALD trial (data on file)	Analysis of prospective EQ-5D data taken from the trial. Lloyd et al. 2006[30] utilities were used in a scenario analysis.
Age and sex-matched general Population Utility	CS B.4.2.7.3	Ara and Brazier 2010	As per the NICE recommendation
AE disutility	CS Table 64	Literature (see CS Table 64)	Used only in scenario analysis, as the AE was considered in the regression analysis.

Source: produced by the EAG from information in the CS
Abbreviations: AE adverse event; PD progressed disease; PF progression free;

EAG conclusion on utilities

The company's approach to estimating utility values is reasonable and consistent with the NICE reference case. We report additional scenario analyses using health state utilities from previous NICE appraisals, see section 6.1.2 below.[34-36]

4.2.6 Resources and costs

4.2.6.1 Drug acquisition

The company presented the drug acquisition costs in CS B. 3.5.1.1. CS Table 66 summarises the unit drug costs.

Elacestrant is administered orally, and patients receive a 345 mg dose daily. Elacestrant is available in packages of 28 tablets (345 mg or 86 mg each tablet) with a proposed list price of [redacted] (345 mg) and [redacted] (86 mg). Elacestrant is available with a patient access scheme (PAS) prices of [redacted] (345 mg) and [redacted] (86 mg).

For each subgroup, we have different comparators:

- **Subgroup 1**: everolimus and exemestane are administered orally, and patients receive a 10 mg tablet of everolimus and 25 mg of exemestane daily. Everolimus is available in packages of 30 tablets (2.5 mg, 5 mg or 10 mg) with the lowest list price (BNF) [38] of £1,020.00 (2.5 mg), £1,912.50 (5 mg) and £2,272.05 (10 mg) per package. Exemestane is available in packages of 30 tablets (25 mg each tablet) with a list price (eMIT 2023) [39] of £4.25. The EAG observed that everolimus has lower prices in eMIT [39]: 30 tablet pack costs £403.03 (2.5 mg tablet), £471.99 (5 mg tablet), and £536.65 (10 mg tablet) than the BNF prices considered by the company.
- **Subgroup 2:** alpelisib is administered orally, and patients receive a 300 mg dose daily. Alpelisib is available in a 56-tablet package (150 mg tablet) and a 28-tablet package (200 mg tablet), both with a list price (BNF) of £4,082.14. Fulvestrant is administered via intramuscular injections of 500 mg. Patients receive the loading doses on days 1, 15 and 29 of the treatment. After that, the maintenance dose is administered monthly. Fulvestrant is available in packages with two vials of 250 mg each and a list price (eMIT 2023) [39] of £80.18 per package.

The company included relative dose intensity (RDI) adjustments for the costs of elacestrant and the comparators, see CS Table 67. The RDI estimate for elacestrant ([redacted]) is from the EMERALD trial results for subgroup 1 and the comparator estimates are from the literature: everolimus 98%, exemestrane, 100% (Jerusalem et al. 2016[40]), alpelisib and fulvestrant 94% (Alaklabi et al. 2022)[41]).

4.2.6.2 Drug administration

Costs by method of administration are shown in Table 28. Oral treatments are assumed to have no administration cost. Intramuscular injections were assumed to take 10 minutes of a primary care nurse's time, with costs from the PSSRU 2022.[42] The cost of Intravenous injections required for subsequent treatments is taken from the NHS Cost Collection 2021/22 (SB12Z: Deliver simple parenteral chemotherapy at first attendance).[43]

Table 28 Drug administration costs per method

Treatments	Method admin.	Admin. cost
Elacestrant, everolimus, exemestane, alpelisib Subsequent treatment: capecitabine	Oral	£0.00
Fulvestrant	Intra muscular	£8.67
Subsequent treatments: docetaxel, paclitaxel	IV infusion	£286.71

4.2.6.3 Health state costs

Health state costs include consultations with health and social service care professionals, hospital resource use, and treatment follow-up. The frequency of resource use was taken from the NICE TA619 (Palbociclib with fulvestrant for treating hormone receptor-positive, HER2-negative, advanced breast cancer)[44] manufacturer's submission, converted to the model cycle length: see CS Table 68.

Clinical advice to the EAG suggested some differences in the frequency of investigations and consultations, including less frequent GP visits and more frequent oncology specialist consultations (four appointments a year instead of two for a progression-free health state and a higher number of visits to the post-progression health state to allow treatment changes). The EAG assessed a scenario with these modifications, see section 6.1.4.

Healthcare unit costs were taken from the PSSRU 2022[42] report and NHS Cost Collection 2021/22[43] data (CS Table 69). In response to clarification question B9, the company updated the unit cost for physiotherapy in CS Table 69 from £45.50 to £48.50. With this correction, the total healthcare cost per cycle is £51.80 for the progression-free health state, and £101.12 for the progressed disease health state.

4.2.6.4 Subsequent treatment

Patients who progress to the progressed disease (PD) health state may commence chemotherapy. The unit costs for the chemotherapies that are included in the company's model (capecitabine, docetaxel, and paclitaxel) are shown in CS Table 73. The EAG notes a minor discrepancy in the list price of paclitaxel 100 mg in CS Table 73. This was corrected in response to clarification question B10 and updated in the economic model (see section 5.3.1).

CS Table 74 shows the proportion of each chemotherapy assumed in the company's base case; and the duration and treatment costs. The EAG notes discrepancies in the subsequent treatment costs in CS Table 74, which the company amended in response to clarification question B11. Table 29 below summarises the corrected subsequent treatment costs.

Table 29 Subsequent treatment costs with EAG corrections

Chemotherapy	Dose per cycle (mg)	Admin. Per cycle	Duration (cycle)	Total Drug cost (£)	Total Admin. Cost (£)	One-off cost (£)
Capecitabine	2036 mg	9.33	13.64	£110.93	£0.00	£110.93
Docetaxel	136 mg	0.33	24.00	£108.87	£2,293.68	£2,402.55
Paclitaxel	471 mg	0.33	20.90	£190.15	£1,997.41	£2,187.57

Source: Based on CS Document B Table 74 and section CS B.3.5.4.2

Based on EMERALD [17] results, the company assumed that only a proportion ([redacted]) of patients would start subsequent treatment after disease progression. The company assumed that all patients starting subsequent treatment would receive capecitabine. Therefore, the one-off cost of subsequent treatment applied on disease progression is [redacted].

Clinical advice to the EAG is that:

- Patients with slow progressing disease are most likely to be candidates for third line treatment.
- The majority of patients who have chemotherapy might receive capecitabine.
- Docetaxel would be used infrequently in this subsequent treatment setting.
- Patients receiving paclitaxel should usually receive weekly treatment with 70 to 80 mg/m^2 for 12 to 18 weeks.
- Eribulin should be considered as an option for chemotherapy.

NICE TA423[45] states that eribulin is only indicated to treat metastatic breast cancer after two or more chemotherapies. The economic model is not set up to consider multiple lines of chemotherapy. Therefore, we did not include eribulin as an additional option in the scenario analysis.

Although the company reported that subsequent treatment distributions were explored in scenario analyses, results for these scenarios were not included in the CS. We explore alternative proportions of subsequent treatments, including the proportion described in Telford et al. 2016 [31] (see section 6.1.1).

4.2.6.5 ESR1-mut testing costs

The company notes that genomic testing for ESR1 mutations is not currently funded in the NHS, but they anticipate that funding would be introduced in a similar way as for PIK3CA

mutation testing after NICE approval of alpelisib (TA816).[8] The cost of ESR1 testing in the company's model is based on the following assumptions:

- £300 per test using digital PCR (CS Table 72). In response to clarification question B8, the company reported that the digital PCR test cost was based on feedback from clinical pathologists.
- 50% of the target population will test positive for an ESR1 mutation, based on results from trials of Imlunestrant (Jhaveri et al. 2023)[46] and palazestrant (Lin et al. 2023)[47]
- 100% of patients are currently tested for the PIK3CA mutation, so no additional cost is included in the model for testing in the dual-mutation subgroup.

The company base case assumes a prevalence-based cost of £600 per person treated in their base case analysis (£300 / 50% = £600), because two people would need to be tested to identify one patient with an ESR1 mutation for whom elacestrant would be suitable.

We note that CS Table 72 also cites a prevalence-based cost of £857.46, but the basis for this estimate is unclear and it is not included in the company's model.

Clinical advice to the EAG regarding ESR1 testing is that:

- ESR1 testing in the EMERALD study was conducted with a Guardant 'liquid biopsy' assay, which uses a blood sample for circulating tumour DNA (ctDNA) analysis. Guardant360 CDx is FDA approved as a companion diagnostic for elacestrant, but an NHS price and pathway for this test is not currently available.
- The ESR1 mutation test would have to be conducted separately from current genetic testing used prior to breast cancer treatment (which identifies whether a PIK3CA mutation is present). As ESR1 is an acquired mutation that can develop after initial treatment, analysis of the primary tumour sample may not be accurate.
- ESR1 mutation testing could be conducted using the same analytical method (digital PCR based on a tissue sample) that is currently used for PIK3CA testing in the NHS, estimated to cost approximately £300.
- However, this approach has disadvantages, including either reliance on a historical tissue sample or a single site repeat biopsy, which may not reflect disease status due to tumour heterogeneity. Repeat sample collection and reporting of the result might delay the start of treatment, as the current reporting time for digital PCR is about a week. Adding ESR1 mutation testing might also burden the testing laboratories further, which could further delay the test results.

- Between 10% to 20% of patients are expected to have the dual mutation (ESR1 and PIK3CA mutations).

In response to a request from NICE, the NHS Genomic Medicine Service (GMS) provided estimates of the possible cost of ctDNA tests for ESR1 mutation in the NHS. They suggested that for the purpose of modelling the impact on the NHS, the cost of providing this testing would be in the region of [redacted]. This assumes that in the future, there would be additional targets to be tested for these patients, and that therefore the testing approach would be to use large next generation sequencing (NGS) panel testing of ctDNA samples. Currently this can be delivered by the North Thames NHS Genomic Laboratory Hub (GLH) using the Marsden360 assay at a cost of [redacted]. We understand that a number of NHS GLHs are currently exploring this or a similar delivery model for ctDNA testing. We report additional EAG scenario analysis using these GMS estimated costs, see section 6.1.5.

4.2.6.6 Adverse event costs

Adverse event costs are calculated by multiplying the total frequency of the adverse events by their unit cost. These costs are applied as a one-off in the first treatment cycle only.

The unit costs of treating each adverse event are taken from the NHS Collection Cost 2021/22[43] and are available in CS Table 70. The adverse event frequency for each treatment arm is shown in CS Table 63. The total adverse event cost for each treatment arm is shown in CS Table 71. The EAG noted some errors in CS Table 63, where adverse events frequencies were misplaced. The company corrected this table in CS document B as requested by the EAG in response to clarification question B7.

4.2.6.7 End-of-life costs

The company's model includes a cost of £8,061 for end-of-life care for deaths related to breast cancer. This estimate was taken from Round et al. 2015 [48] updated to 2021/22 prices using the NHS PSSRU cost inflation index.[42]

The PSSRU Unit Costs for Health and Social Care 2022 manual [42] reports end-of-life health and social care costs based on the Nuffield Trust report by Georghiou et al. (2012) [49], with a cost of £13,113 in the final year of life for cancer patients. The EAG ran a scenario using this source in section 6.3.

Table 30 End of life cost for health and social care

Source	Cost £ per person in the final year of life	
	Original estimate	2021/22 prices
Round et al. 2015	£7,189, 2013/14 prices	£8,061
Georghiou et al. 2012	£10,844, 2010/11 prices	£13,113

EAG comment on resources and costs

- The company's approach to estimating resources and costs in the economic model is consistent with the NICE reference case and previous technology appraisals for metastatic breast cancer.
- The EAG identified some minor errors in resource use costs (physiotherapy), subsequent treatment costs (paclitaxel 100 mg list price, total costs per treatment in CS Table 74), and adverse events (AE frequency in CS Table 63). The company corrected these errors in response to clarification questions B7, B9, B10 and B11.
- We assessed the impact of uncertainty over subsequent treatment costs in two scenarios, varying the proportions to select the most expensive treatment and the proportions in Telford et al. 2016[31]. We also tested scenarios varying the cost of ESR1 testing, healthcare resource use and the cost of end-of-life care. See section 6.1.

5 COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results

CS section 3.9 Table 81 reports the base case results for elacestrant vs everolimus + exemestane (EVE + EXE) for the ESR1-mut + >12 months of prior ET + CDK4/6i population (subgroup 1) and elacestrant vs alpelisib + fulvestrant (ALP+FUL) for ESR1-mut, PIK3CA-mut+>12 months of prior ET + CDK4/6i population (subgroup 2). The company made corrections to their model in response to clarification questions and reported in an updated CS document B.

Revised deterministic base case results are reported in Table 31 below. Note that we report costs and health outcomes, including life years (LYs) and QALYs, discounted at 3.5% per year. Total and incremental QALYs are reported without the severity modifier of 1.2 applied by the company for subgroup 1 (see section 7 for further details). We report ICERs for subgroup 1 both with and without the severity modifier.

- For subgroup 1, the company's base case ICER is £24,893 per QALY gained including the severity modifier; and £29,872 per QALY gained without the severity modifier.
- For subgroup 2, the company's base case result indicates that elacestrant is dominant: with a lower expected cost and higher expected QALYs compared to alpelisib + fulvestrant. The net monetary benefit (NMB) of elacestrant is £17,803 at a cost-effectiveness threshold of £20,000 per QALY gained; and £20,570 at a threshold of £30,000 per QALY gained.

The company's base case results and all other cost-effectiveness results in this report are conducted with a proposed confidential patient access scheme (PAS) price discount for elacestrant. However, they do not include confidential discounts for any other medications. Therefore, the ICERs do not reflect the actual prices that would be paid by the NHS. Results including all available NHS price discounts for comparator and subsequent medications in addition to the proposed PAS discount for elacestrant are presented in a separate confidential addendum to this report.

Table 31 Company's base case results with PAS price for Elacestrant

Technologies	Total costs (£) [a]	Total LYG [a]	Total QALYs [a]	Incremental costs (£) [a]	Incremental LYG [a]	Incremental QALYs [a]	ICER (£/QALY) no severity modifier	ICER (£/QALY) with severity modifier (1.2)
Subgroup 1 - ESR1-mut and ≥12 months of prior ET + CDK4/6i								
Everolimus + exemestane	[redacted]	[redacted]	[redacted]					
Elacestrant	[redacted]	[redacted]	[redacted]	£18,883	0.892	0.632	£29,872	£24,893
Subgroup 2 - ESR1-mut, PIK3CA-mut and ≥12 months of prior ET + CDK4/6i								
Alpelisib + fulvestrant	[redacted]	[redacted]	[redacted]					
Elacestrant	[redacted]	[redacted]	[redacted]	-£12,269	0.394	0.277	Dominant	---

Source: CS Table 81
Abbreviations: CDK4/6i, cyclin-dependent kinase 4/6 inhibitor; ESR1, oestrogen receptor 1 gene; ET, endocrine therapy; ICER, incremental cost-effectiveness ratio; LYG, life years gained; NMB, net monetary benefit; PAS, patient access scheme; PIK3CA, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha; QALYs, quality-adjusted life years.
[a] Discounted at 3.5 % per year, with no severity modifier applied to QALYs

5.2 Company's sensitivity analyses

5.2.1 Deterministic sensitivity analyses

CS section B.3.10.2 reports the deterministic sensitivity analysis (DSA) results for elacestrant vs everolimus + exemestane (subgroup 1) and elacestrant vs alpelisib + fulvestrant (subgroup 2). The parameters varied in the DSA are listed in CS Table 79. The company notes that parametric survival model coefficients were only varied only in the PSA, not in the DSA, because these coefficients are correlated. The EAG considers that this is reasonable for testing the sensitivity of individual parameters.

The company presented two tornado diagrams based on the impact on net monetary benefit: see CS Figure 52 (elacestrant vs everolimus + exemestane for subgroup 1) and CS Figure 53 (elacestrant vs alpelisib + fulvestrant for subgroup 2). Parameters relating to the everolimus drug cost, mean age and RDI (elacestrant and everolimus) were the main drivers for the model in subgroup 1, and RDI (alpelisib and elacestrant) was the main driver in subgroup 2.

5.2.2 Scenario analysis

The company coded 59 scenarios to test structural and methodological uncertainties in its economic model (see Appendix 5 for the full list). They reported results for 20 of these scenarios in subgroup 1 (CS Table 85), and for 21 scenarios in subgroup 2 (CS Table 86):

- For subgroup 1, the ICER for elacestrant was less than £30,000 per QALY in all but one scenario: using the gamma distribution for the elacestrant OS (ICER of £43,793).
- For subgroup 2, elacestrant was dominant (positive NMB) in all scenarios.

We discuss additional scenarios of interest in section 6.1.

5.2.3 Probabilistic Sensitivity Analysis

The company's probabilistic sensitivity analysis results were estimated for 5,000 simulations, illustrated in scatterplots (CS Figures 50 and 51) and cost-effectiveness acceptability curves (CEACs, CS Figures 48 and 49).

Mean probabilistic results for the company’s base case are reported in CS Table 82). These results were revised to include corrections after the clarification response (see section 5.3.1). The probabilistic results are stable and consistent with the deterministic results.

The distributions used for the parameters included in the PSA analysis are summarised in CS Table 79:

- **Normal distribution**: patient characteristics (age, proportion female, BSA), drug unit costs (except elacestrant and alpelisib), RDI, administration costs, healthcare resource use costs, healthcare resource use frequency, subsequent treatment costs, subsequent treatment duration, ESR1-mut testing cost, adverse event costs.
- **Beta distribution**: proportion with ESR1-mut, the proportion of PFS events assumed to be in progression, adverse event frequency, and health state utility values.
- **Multinormal distribution**: OS curves (elacestrant and comparators), PFS curves (elacestrant and comparators), and general population utility coefficients (Ara and Brazier equation [37]).
- **Dirichlet**: subsequent treatment distribution

The EAG observed that all cost parameter uncertainties were represented with a normal distribution, instead of gamma or log-normal distributions. We checked the economic model and verified that all cost parameters only allow positive cost values during the PSA iterations. We also note that the subsequent treatment distribution was modelled with a Dirichlet distribution, but this was not active in the PSA.

5.3 Model validation and face validity check

We conducted a range of checks on the company's model using an EAG checklist:

- **Input checks:** comparison of all parameter values in the model against the values stated in the company submission and cited sources.
- **Output checks:** replication of results reported in the CS using the company model. Manually running scenarios and checking model outputs against results reported in the CS for the deterministic sensitivity analyses and scenario analyses.
- **'White box' checks:** checking individual equations within the model.
- **'Black box' checks:** applying a range of extreme value and logic tests to check the plausibility of changes in results when parameters are changed.

The model is generally well-implemented, although we spotted minor discrepancies between the company submission and the initial version of the model, which were corrected in a revised version submitted with the company's clarification response, as described below.

5.3.1 Company's corrections to the company model

In their response to the EAG clarification questions, the company amended some parameters values listed below:

- Mean age at baseline for the ESR1-mut and ≥12 months of prior ET + CDK4/6i (subgroup 1) (CQ B1, see section 4.2.2.1).
- Adverse events disutilities and durations in CS Table 64 (CQ B6, see section 4.2.5.3)
- Adverse event frequency in CS Table 63 (CQ B7, see section 4.2.6.6)
- Resource unit cost for physiotherapy in CS Table 69 (CQ B9, see section 4.2.6.3)
- Unit drug cost for paclitaxel 100 mg in CS Table 73 (CQ B10, see section 4.2.6.4)

The company also corrected two PSA equations (PSA sheet, column AI16:AI5015 and AJ16 to AJ5015) related to the incremental cost and QALYs, where the elacestrant total cost and total QALYs were fixed for the first iteration result values (AI$16 and AJ$16) in all 5,000 iterations. The company provided a revised model considering the clarification response modifications (version 28/05/2024).

The updated results led to a slight increase in the ICER from £24,873 to £24,893 per QALY gained for subgroup 1, including the 1.2 severity modifier. For subgroup 2, elacestrant remained dominant, with a slight increase in the NMB from £20,451 to £20,570.

5.3.2 EAG corrections to the company's model

The EAG identified a minor issue in the scenario results. In the “Scenario analysis” sheet, column BB refers to the incremental QALYs equation. This equation used the severity modifier parameter to calculate the incremental QALYs instead of a fixed value. Therefore, all scenario results change if the severity modifier parameter value is changed. In addition, this makes scenario 5 (severity modifier = 1) in row 23 or scenario 6 (severity modifier = 1.2) in row 24 incorrect, depending on the comparator. However, neither of these scenarios were reported in CS B Tables 85 and 86. We corrected only cells BB23 and BB24. This issue does not affect the base case result.

5.3.3 EAG summary of key issues and additional analyses

We summarise and critique key assumptions in the company's model in Table 32 below.

Table 32 EAG summary and critique of key features of the economic model

Aspect of model	Company assumptions	EAG comment	EAG additional analyses
Decision problem			
Population and subgroups	Target population for elacestrant restricted patients with disease progression after at least 12 months of prior ET + CDK4/6i (subgroup 1). Results also presented for the dual mutated subgroup within the target population (subgroup 2).	Results are based on post-hoc analysis of EMERALD trial data by duration of prior therapy. This improves the estimated cost-effectiveness of elacestrant but increases uncertainty due to the smaller sample sizes.	None
Mean age at baseline	Base case from EMERALD trial. Scenario Flatiron means (CQ response Table 5)	The scenario with Flatiron mean ages is consistent with the use of MAIC-adjusted clinical outcomes in the model	EAG preferred: Flatiron Scenario: EMERALD
Comparators	Everolimus + exemestane for subgroup 1 Alpelisib + fulvestrant for subgroup 2	This is reasonable, although ET with or without chemotherapy, or chemotherapy alone may be used for some patients	None
Clinical effectiveness			
Survival extrapolations	Independent curves fitted to MAIC-weighted EMERALD data and Flatiron KM	Uncertainty due to unanchored MAIC and small sample sizes	Additional scenarios, see Table 33 and Table 34 below
OS distribution	Subgroup 1: log-logistic for elacestrant; gamma for everolimus + exemestane	Subgroup 1 base case predicts long-term OS benefit for elacestrant which is	EAG preferred:

Aspect of model	Company assumptions	EAG comment	EAG additional analyses
	Subgroup 2: Weibull for elacestrant; gamma for alpelisib + fulvestrant	optimistic given current evidence. Agree with base case for subgroup 2	Subgroup 1: gamma both arms Subgroup 2: No change Additional scenarios
PFS distribution	Subgroup 1: log-normal for both arms Subgroup 2: log-normal for both arms	Agree	Additional scenarios
Treatment duration	KM from EMERALD trial for elacestrant Assume TTD = PFS for comparator arms	Agree with use of KM for elacestrant. But potential bias against comparators if some patients discontinue prior to disease progression	Exploratory scenarios with adjustment of comparator TTD relative to the PFS
Health-related quality of life			
Health state utilities	Estimates from the EMERALD trial, mapped from EQ-5D-5L to EQ-5D-3L (CS Table 65)	We agree	Additional scenarios with utilities from previous NICE appraisals [34-36]
Adverse event disutilities	AE disutility and duration presented in CS Table 64. Utility regression includes AE term, so additional AE disutility was not included in the company's base case	We agree	No change
Age-related utility decrement	Adjustment from Ara and Brazier 2010 [37]	We agree	No change

Aspect of model	Company assumptions	EAG comment	EAG additional analyses
Resource use and costs			
Treatment cost	CS B.3.5.1.1 and Table 66. Everolimus and alpelisib were sourced from BNF 2024 and exemestane and fulvestrant from eMIT 2023	The eMIT tool presented a lower acquisition price for everolimus.	EAG preferred: everolimus price from eMIT
Relative dose intensity (RDI)	CS B.3.5.1.1 and Table 67. Parameters were collected from the EMERALD trial for elacestrant and from the literature for the comparators.	We agree	No change
Administration cost	CS B.3.5.1.2 and Table	We agree	No change
Resource use and costs	Based on NICE TA619[44] and presented in CS Table 68.	We agree	Additional scenario based on clinical advice regarding resource use frequency (see Table in section 0).
Subsequent treatments	The proportions of patients receiving chemotherapies were based on assumptions.	Uncertainty over % use of each chemotherapy for progressed disease health state.	Additional scenarios for distribution of subsequent treatments (see Table 29 in section 6.1.1).
ESR1 mutation testing	Cost based on digital PCR testing (~£300). Prevalence-based cost £600 per person treated, assuming 50% of tested have ESR1-	There is uncertainty over the cost of introducing ESR1 testing in the NHS. Potential service implications due to the	Exploratory scenarios varying for ESR1 test cost and

Aspect of model	Company assumptions	EAG comment	EAG additional analyses
	mut (£300 / 50%). Potential for introduction of liquid-based biopsy as companion diagnostic.	delay of results for digital PCR, and pressure on genomic testing facilities	number needed to test to find one positive (see section 6.1).
Adverse event	Costs in CS Table 70 based on NHS Cost Collection 2021/22[43]. AE frequency is in CS Table 63 with estimates from the literature.	We agree	No change
End-of-life	Based on estimates from Round et al. 2015[48]	We agree	Additional scenario with Georghiou et al. 2012[49] cost (see Table 30 in section 4.2.6.7)

6 EAG'S ADDITIONAL ANALYSES

6.1 Exploratory and sensitivity analyses undertaken by the EAG

Based on the EAG critique of the company's model assumptions (Table 32), we performed a range of additional scenario analyses, which are summarised in the following subsections.

6.1.1 Exploratory scenarios: survival curves (OS, PFS and TTD)

See section 4.2.4 above for EAG discussion and conclusions on the selection of survival curves for OS, PFS and TTD. We summarise the company's base case and scenarios and EAG additional scenarios for subgroup 1 and 2 in Table 33 and Table 34 respectively.

Table 33 Survival analysis – scenario analysis (subgroup 1)

	Elacestrant	**Everolimus + exemestane**
OS		
Company base case	Log-logistic	Gamma
Company scenarios	Gamma, Log-normal	Weibull, exponential
EAG scenarios	Weibull	+ MAIC HR
	Gamma	+ MAIC HR
	Generalised gamma	+ MAIC HR
	Log-logistic	+ MAIC HR
PFS		
Company base case	Log-normal	Log-normal
Company scenarios	Log-logistic	log-logistic, gamma
EAG scenarios	Exponential	
		Weibull
TTD		
Company base case	KM curve	Assumed equal to PFS
Company scenarios	log-normal, log-logistic	
EAG scenarios	Generalised gamma	
		HR of 0.8 for TTD versus PFS

Table 34 Survival analysis – scenario analysis (subgroup 2)

	Elacestrant	**Alpelisib + fulvestrant**
OS		
Company base case	Weibull	Gamma
Company scenarios	Gamma, log-normal	Weibull, log-normal

	Elacestrant	Alpelisib + fulvestrant
EAG scenarios	Gompertz	+ MAIC HR
	Weibull	+ MAIC HR
	Gamma	+ MAIC HR
PFS		
Company base case	Log-normal	Log-normal
Company scenarios	Log-logistic, exponential	Generalised gamma, gamma
EAG scenarios		Weibull
TTD		
Company base case	KM curve	Equal to PFS
Company scenarios	Log-normal, log-logistic	
EAG scenarios		HR of 0.5 for TTD versus PFS

6.1.1.1 EAG survival scenario results for subgroup 1

The EAG exploratory scenarios for survival curves in subgroup 1 had the following results (company base case: ICER £24,893 per QALY, with the 1.2 QALY weight).

For the OS curves, we tested the following distributions for elacestrant with the MAIC hazard ratio used to estimate curves for the comparator EVE + EXE:

- Weibull distribution: ICER £44,266 per QALY
- Gamma distribution: ICER £36,925 per QALY
- Generalised Gamma: ICER £51,802 per QALY
- Log-logistic distribution: ICER £27,070 per QALY

For the PFS curves (independently fitted curves as in the company's base case):

- Exponential for elacestrant PFS: ICER £25,174 per QALY
- Weibull for the EVE + EXE PFS: ICER £25,627 per QALY

For the TTD curves:

- Generalised gamma for elacestrant TTD: ICER £30,457 per QALY
- HR of 0.8 for EVE + EXE TTD vs PFS: ICER £27,782 per QALY

6.1.1.2 EAG survival scenario results for subgroup 2

The EAG exploratory scenario for survival curves in subgroup 2 had the following results (company base case: elacestrant dominant, NMB £20,570 at £30,000 per QALY threshold).

For the OS curves, we tested the following distributions for elacestrant with the MAIC hazard ratio used to estimate curves for the ALP + FUL comparator:

- Gompertz: elacestrant dominant, NMB £16,697 at £30,000 per QALY threshold
- Weibull: elacestrant dominant, NMB £19,341 at £30,000 per QALY threshold
- Gamma: elacestrant dominant, NMB £21,438 at £30,000 per QALY threshold

With the EAG scenario using an independent Weibull distribution for the ALP+FUL PFS curve, elacestrant remained dominant, with an NMB £20,737 at the £30,000 per QALY threshold.

The EAG scenario with ALP+FUL TTD estimated assuming a 0.5 hazard ratio relative to the ALP+FUL PFS curve resulted in an ICER of £4,362 per QALY (elacestrant not dominant).

6.1.2 Exploratory scenarios: utilities

The company reported one scenario for health state utilities (pre- and post-progression) using values from Lloyd et al. (2006).[30] We considered additional scenarios that were included in the model but not reported in the CS, with health state utilities taken from previous NICE appraisals of untreated advanced HR+ breast cancer:

- NICE TA496 (ribociclib)[34], based on MONALEESA-2 trial data;
- NICE TA503 (fulvestrant)[35] based on FALCON trial data; and
- NICE TA563 (abemaciclib)[36], based on MONARCH 3 trial data.

Table 35 Utility values – scenario analysis

Health state	EMERALD	NICE TA496	NICE TA503 [a]	NICE TA563	Lloyd et al. 2006
PFS on treatment	[redacted]	0.774	0.751	0.690	0.715
PFS off treatment	[redacted]	0.774	0.751	0.690	0.715
Post-progression	[redacted]	0.505	0.691	0.505	0.600
Progression decrement	[redacted]	0.269	0.060	0.185	0.115

Source: Partly reproduced from CS Table 65 and economic model
[a] Telford et al. 2019[31] also based their utilities on the FALCON trial, so their utilities are equal to NICE TA503

For subgroup 1, the ICER varied between £24,968 (NICE TA503) to £28,958 (NICE TA563) including the QALY weight of 1.2. The scenarios with NICE TA496 and Lloyd et al. 2006 health state utilities had an ICER of £26,547 and £26,937, respectively.

For subgroup 2, elacestrant remained dominant for all utility scenarios. The NMB varied between £18,497 (NICE TA563) and £20,425 (NICE TA503) for a WTP of over £30,000 (base case NMB £20,570). The scenarios with Lloyd et al. 2006 and NICE TA496 health state utilities had an NMB of £19,603 and £18,658 for a WTP of over £30,000, respectively.

6.1.3 Exploratory scenarios: subsequent treatment distribution

To address the observations in section 4.2.6.4 about the distribution of subsequent treatments, we explored two scenarios, including the distribution in Telford et al. 2019[31] for second-line treatment and a scenario with a more expensive treatment (see Table 36 below). Although the subsequent treatment costs increased in these scenarios, the difference between arms was very small (see Table 37).

Table 36 Subsequent treatment distribution

Chemotherapy	Company submission	EAG scenario 1 - Telford et al. 2019 (2nd line treatment)[a]	EAG scenario 2
Capecitabine	100%	48%	0%
Docetaxel	0%	28%	0%
Paclitaxel	0%	24%	100%

Table 37 EAG scenarios: Subsequent treatment costs variation

Scenario	Subsequent treatment costs		Difference between arms
	Elacestrant	Comparator	
Subgroup 1 – elacestrant vs everolimus + exemestane			
Company base case	[redacted]	[redacted]	-£2
EAG scenario 1	[redacted]	[redacted]	-£21
EAG scenario 2	[redacted]	[redacted]	-37
Subgroup 2 – elacestrant vs alpelisib + fulvestrant			
Company base case	[redacted]	[redacted]	-£1
EAG scenario 1	[redacted]	[redacted]	-£3
EAG scenario 2	[redacted]	[redacted]	-£4

Source: produced by the EAG from the company's model

6.1.4 Exploratory scenario: resource use

As per clinical advice (section 4.2.6.3), the EAG explored a scenario adjusting the number of visits to the GP and specialist oncologist, as shown in Table 38 below. For subgroup 1, this scenario increased the ICER by £231. For subgroup 2, the resource use cost increment was £191, and elacestrant remained dominant.

Table 38 Healthcare resource use (frequency per month) – scenario analysis

Resource	Company base case		EAG scenario	
	Progression free	Post progression	Progression free	Post progression
GP visit	1.00	1.50	0.25	0.38
Oncology specialist	0.17	0.50	0.33	1.00

Source: Partly reproduced from CS Document B Table 68

6.1.5 Exploratory scenario: ESR1 mutation test

The EAG conducted four exploratory scenarios to assess the impact of uncertainty over the cost of ESR1 mutation testing. The scenarios are summarised in Table 39 below.

Table 39 ESR1 mutation test costs – scenario analysis

Scenario	Source	Non prevalence based cost	Prevalence based cost [a]
Company base case	Digital PCR	£300	£300/0.5 = £600
Company scenario	Exclude ESR1 mutation testing cost		
EAG scenarios	Estimated NHS GMS	[redacted]	[redacted] /0.5 = [redacted]
	Marsden360 assay	[redacted]	[redacted] /0.5 = [redacted]

Source: Produced by the EAG using information from the CS and GMS estimates
[a] Assuming 50% prevalence of ESR1 mutation at the point of testing

For subgroup 1 (QALY weight of 1.2 applied):

- Non-prevalence based
  - NHS GMS estimate: ICER £25,223 per QALY.
  - Marsden360 assay: ICER £26,343 per QALY
- Prevalence-based
  - NHS GMS estimate: ICER £26,343 per QALY.
  - Marsden360 assay: ICER £28,585 per QALY

For subgroup 2, elacestrant is dominant for all scenarios and:

- Non-prevalence based
  - NHS GMS estimate: NMB £20,320 at £30,000 per QALY threshold
  - Marsden360 assay: NMB £19,470 at £30,000 per QALY threshold
- Prevalence-based
  - NHS GMS estimate: NMB £19,470 at £30,000 per QALY threshold
  - Marsden360 assay: NMB £17,770 at £30,000 per QALY threshold

6.2 EAG's preferred assumptions

Based on the EAG critique of the company's model discussed in Table 32, we have identified four key aspects of the company base case with which we disagree. Our preferred model assumptions are the following:

- Mean age from the Flatiron database (see section 4.2.2.1)
- Everolimus prices from eMIT 2023 instead of the BNF (see section 4.2.6.1). This only affects subgroup 1.
- The proportion of positive ESR1 tests for subgroup 2 (dual mutated) based on clinical advice estimate of 20% (see section 4.2.6.5). The proportion of positive cases for subgroup 1 remains at 50%.
- OS extrapolations: subgroup 1 independent gamma for both arms
  (no change to company base case for subgroup 2)

Table 40 shows the cumulative cost-effectiveness results for subgroup 1 of adding the EAG's preferred model assumptions one at a time to the corrected company's base case. Including all of the EAG's preferred assumptions increases the ICER from £24,893 to £73,224 per QALY (including the QALY weight of 1.2).

Table 40 EAG's preferred assumptions: cumulative change to ICER for subgroup 1

Preferred assumption	Treatment	Total costs	Total QALYs	ICER £/QALY No QALY weight	ICER £/QALY With QALY weight
Company's revised base case	EVE + EXE	[redacted]	[redacted]	£29,872	£24,893
	Elacestrant	[redacted]	[redacted]		
+ Mean age from Flatiron ([redacted] yrs)	EVE + EXE	[redacted]	[redacted]	£29,942	£24,952
	Elacestrant	[redacted]	[redacted]		
	EVE + EXE	[redacted]	[redacted]	£47,723	£39,769

Preferred assumption	Treatment	Total costs	Total QALYs	ICER £/QALY No QALY weight	ICER £/QALY With QALY weight
+ Everolimus price from eMIT 2023	Elacestrant	[redacted]	[redacted]		
+ OS Independent gamma both arms	EVE + EXE	[redacted]	[redacted]	£87,869	£73,224
	Elacestrant	[redacted]	[redacted]		
EAG base case	EVE + EXE	[redacted]	[redacted]	£87,869	£73,224
	Elacestrant	[redacted]	[redacted]		

Source: Produced by the EAG from the company's model

Table 41 shows the cumulative cost-effectiveness results of adding the EAG's preferred model assumptions for subgroup 2. Elacestrant remains dominant, with a small reduction in the NMB from £20,570 to £19,670 at the £30,000 per QALY threshold.

Table 41 EAG's preferred assumptions: cumulative change to ICER for subgroup 2

Preferred assumption	Treatment	Total Costs	Total QALYs	ICER £/QALY	NMB (£) at WTP £30,000
Company's revised base case	ALP + FUL	[redacted]	[redacted]	Dominant	£20,570
	Elacestrant	[redacted]	[redacted]		
+ Mean age from Flatiron ([redacted] yrs)	ALP + FUL	[redacted]	[redacted]	Dominant	£20,570
	Elacestrant	[redacted]	[redacted]		
+ Proportion of positive ESR1-mut tests (20%)	ALP + FUL	[redacted]	[redacted]	Dominant	£19,670
	Elacestrant	[redacted]	[redacted]		
EAG base case	ALP + FUL	[redacted]	[redacted]	Dominant	£19,670
	Elacestrant	[redacted]	[redacted]		

Source: Produced by the EAG from the company's model

We confirmed that the severity modifier QALY weight is unchanged (1.2 for subgroup 1 and no weight for subgroup 2), see section 7.2.

We reran the probabilistic sensitivity analysis (PSA) with the EAG base case model. The cost-effectiveness scatterplot is shown in Figure 9 (subgroup 1) and Figure 10 (subgroup 2). The probabilistic results are aligned with the deterministic results (see Table 42), with a 3% difference in the ICER for subgroup 1 and a 0.7% difference in NMB (£19,808 for a WTP of £30,000) for subgroup 2.

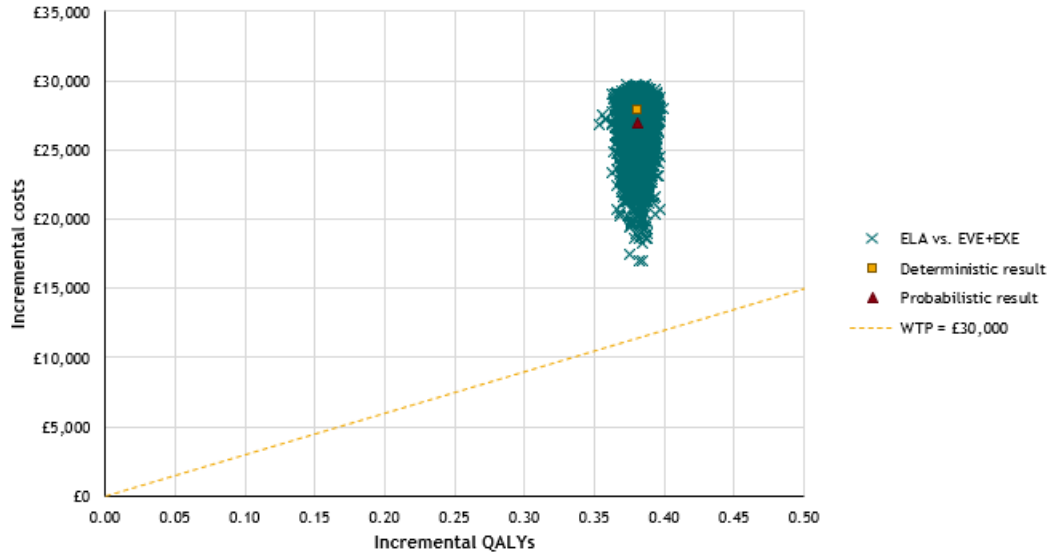

Figure 9 PSA scatterplot graph for subgroup 1 using the EAG preferred assumptions

PSA: probabilistic sensitivity analysis, QALY Quality-adjusted life year, WTP: willingness to pay, ELA: elacestrant, EVE + EXE: everolimus with exemestane

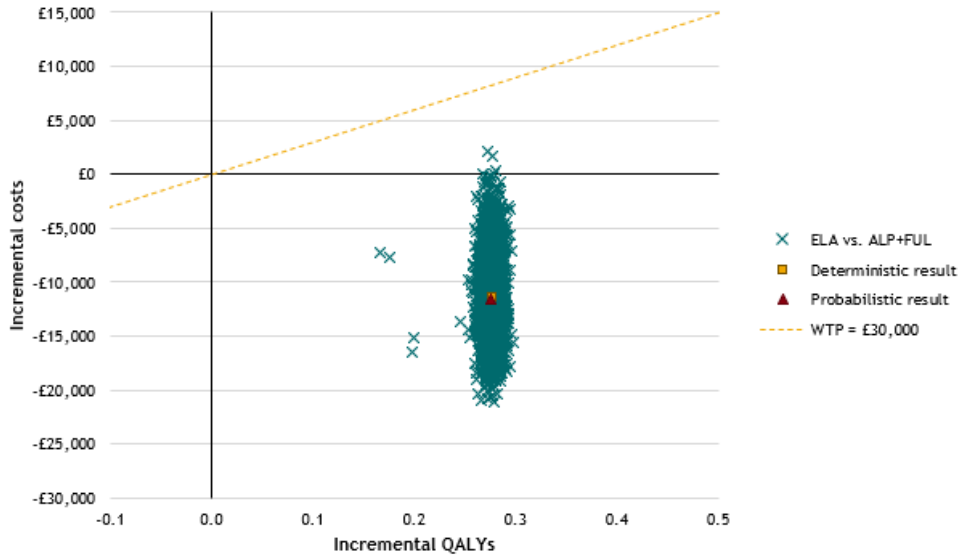

Figure 10 PSA scatterplot graph for subgroup 2 using the EAG preferred assumptions

PSA: probabilistic sensitivity analysis, QALY Quality-adjusted life year, WTP: willingness to pay, ELA: elacestrant, ALP + FUL: alpelisib with fulvestrant

Table 42 Probabilistic sensitivity analysis results – EAG base case

Technologies	Total Costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY) no severity modifier	ICER (£/QALY) with severity modifier 1.2
Subgroup 1 - ESR1-mut and ≥12 months of prior ET + CDK4/6i								
EVE + EXE	[redacted]	[redacted]	[redacted]					
Elacestrant	[redacted]	[redacted]	[redacted]	£26,953	0.422	0.317	£84,914	£70,762
Subgroup 2 - ESR1-mut, PIK3CA-mut and ≥12 months of prior ET + CDK4/6i								
ALP + FUL	[redacted]	[redacted]	[redacted]				Dominant	Not applicable
Elacestrant	[redacted]	[redacted]	[redacted]	-£11,522	0.393	0.276		

Source: Produced by the EAG from the company's economic model

6.3 Scenario analyses conducted with the EAG's preferred assumptions

We performed a range of scenario analyses with the EAG base case to analyse the impact of changing some of the model assumptions. The scenarios in Table 43 and Table 44 are divided into four groups:

- Company base case assumptions that were modified in the EAG preferred analysis (section 6.1)
- Selection of relevant company scenarios described in section 5.2.2
- Selection of relevant additional company scenarios described in Appendix 9.5
- Selection of relevant EAG exploratory scenarios described in section 6.1

6.3.1 Subgroup 1

Table 43 below summarises the results of the scenarios on the EAG base case for subgroup 1. The ICER varied from £35,240 (elacestrant OS – log-normal) to £262,288 (elacestrant OS - Gompertz), assuming a 1.2 QALY weight.

The scenarios that have the most significant effect on the cost-effectiveness are:

- Changes to the elacestrant OS distribution. All five scenarios varied the ICER by more than 45%:
  - The log-normal, exponential and log-logistic distributions decreased the ICER to £35,240, £35,966, and £39,769, respectively.
  - The Weibull and Gompertz distributions increased the ICER to £107,211 and £262,288, respectively.
- Taking the everolimus price from the BNF 2024, instead of eMIT 2023, reduced the ICER by £29,416 (40% decrease).
- Using MAIC hazard ratios, instead of independent parametric survival extrapolations, decreased the ICER by £9,641
- Assuming extrapolation curves for the elacestrant TTD (instead of the KM curve) decreased the ICER by £4,537 using the log-normal distribution and by £5,929 using the log-logistic distribution.
- Varying the ESR1 mutation test cost, with or without adjustment for prevalence, varied the ICER from £73,880 (< 1% increase) to £80,573 (10% increase).

Table 43 EAG scenario analyses for subgroup 1

EAG base case	Scenario	Treatment	Total cost (£)	Total QALYs	ICER (£/QALY) without the severity modifier	ICER (£/QALY) with the 1.2 severity modifier
EAG base case		EVE + EXE	[redacted]	[redacted]	£87,869	£73,224
		Elacestrant	[redacted]	[redacted]		
Company base case assumptions						
Mean age from Flatiron	Mean age from EMERALD	EVE + EXE	[redacted]	[redacted]	£87,838	£73,198
		Elacestrant	[redacted]	[redacted]		
Everolimus price from eMIT 2023	Everolimus price from BNF 2024	EVE + EXE	[redacted]	[redacted]	£52,570	£43,808
		Elacestrant	[redacted]	[redacted]		
Independent gamma for OS curve - both arms	Elacestrant OS – log-logistic	EVE + EXE	[redacted]	[redacted]	£47,723	£39,769
		Elacestrant	[redacted]	[redacted]		
Selected scenarios presented in the submission						
MAIC approach – independent PSM extrapolation	HR	EVE + EXE	[redacted]	[redacted]	£76,300	£63,583
		Elacestrant	[redacted]	[redacted]		
Elacestrant OS – gamma distribution	Log-normal	EVE + EXE	[redacted]	[redacted]	£42,288	£35,240
		Elacestrant	[redacted]	[redacted]		
EVE + EXE OS – Gamma distribution	Weibull	EVE + EXE	[redacted]	[redacted]	£89,199	£74,332
		Elacestrant	[redacted]	[redacted]		

EAG base case	Scenario	Treatment	Total cost (£)	Total QALYs	ICER (£/QALY) without the severity modifier	ICER (£/QALY) with the 1.2 severity modifier
	Exponential	EVE + EXE	[redacted]	[redacted]	£99,295	£82,746
		Elacestrant	[redacted]	[redacted]		
Elacestrant PFS – log-normal distribution	Log-logistic	EVE + EXE	[redacted]	[redacted]	£87,845	£73,204
		Elacestrant	[redacted]	[redacted]		
EVE + EXE PFS – log-normal distribution	Log-logistic	EVE + EXE	[redacted]	[redacted]	£87,838	£73,198
		Elacestrant	[redacted]	[redacted]		
	Gamma	EVE + EXE	[redacted]	[redacted]	£87,935	£73,279
		Elacestrant	[redacted]	[redacted]		
Elacestrant TTD – KM curve	Log-normal	EVE + EXE	[redacted]	[redacted]	£82,424	£68,687
		Elacestrant	[redacted]	[redacted]		
	Log-logistic	EVE + EXE	[redacted]	[redacted]	£80,754	£67,295
		Elacestrant	[redacted]	[redacted]		
Progressed utility source – EMERALD EQ-5D analysis ([redacted])	Lloyd et al. (2006), absolute approach (0.601)	EVE + EXE	[redacted]	[redacted]	£84,919	£70,766
		Elacestrant	[redacted]	[redacted]		
Company's additional scenario analysis presented in the economic model						
Elacestrant OS: Gamma	Weibull	EVE + EXE	[redacted]	[redacted]	£128,654	£107,211
		Elacestrant	[redacted]	[redacted]		

EAG base case	Scenario	Treatment	Total cost (£)	Total QALYs	ICER (£/QALY) without the severity modifier	ICER (£/QALY) with the 1.2 severity modifier
	Gompertz	EVE + EXE	[redacted]	[redacted]	£314,746	£262,288
		Elacestrant	[redacted]	[redacted]		
	Exponential	EVE + EXE	[redacted]	[redacted]	£43,159	£35,966
		Elacestrant	[redacted]	[redacted]		
EVE+EXE PFS: log-normal	Weibull	EVE + EXE	[redacted]	[redacted]	£87,931	£73,276
		Elacestrant	[redacted]	[redacted]		
Elacestrant TTD: KM curve	generalised gamma	EVE + EXE	[redacted]	[redacted]	£94,411	£78,676
		Elacestrant	[redacted]	[redacted]		
EAG exploratory scenarios						
EVE + EXE TTD: equal to PFS	HR for TTD vs. PFS = 0.8	EVE + EXE	[redacted]	[redacted]	£89,509	£74,591
		Elacestrant	[redacted]	[redacted]		
Subsequent treatment cost: 100% capecitabine	Scenario 1 (Telford et. al. 2019) (section 6.1.3)	EVE + EXE	[redacted]	[redacted]	£87,815	£73,179
		Elacestrant	[redacted]	[redacted]		
ESR1-mut testing cost: £300, prevalence based (50%) = £600	NHS GMS, prevalence-based: [redacted] /0.5= [redacted]	EVE + EXE	[redacted]	[redacted]	£91,333	£76,111
		Elacestrant	[redacted]	[redacted]		
	NHS GMS, non-prevalence base: [redacted]	EVE + EXE	[redacted]	[redacted]	£88,656	£73,880
		Elacestrant	[redacted]	[redacted]		

EAG base case	Scenario	Treatment	Total cost (£)	Total QALYs	ICER (£/QALY) without the severity modifier	ICER (£/QALY) with the 1.2 severity modifier
	Marsden360 assay cost, prevalence-based: [redacted] /0.5 = [redacted]	EVE + EXE	[redacted]	[redacted]	£96,688	£80,573
		Elacestrant	[redacted]	[redacted]		
	Marsden 360 assay, non-prevalence base: [redacted]	EVE + EXE	[redacted]	[redacted]	£91,333	£76,111
		Elacestrant	[redacted]	[redacted]		
ESR1-mut testing – proportion of positive tests (50%)	25%	EVE + EXE	[redacted]	[redacted]	£89,759	£74,799
		Elacestrant	[redacted]	[redacted]		
End of life cost: Round et al. 2015	Georghiou et al. 2012	EVE + EXE	[redacted]	[redacted]	£87,638	£73,031
		Elacestrant	[redacted]	[redacted]		
Utilities (PF and PD) from EMERALD EQ-5D analysis	Utilities From Lloyd et al. 2006 (PF and PFD)	EVE + EXE	[redacted]	[redacted]	£89,547	£74,622
		Elacestrant	[redacted]	[redacted]		

Source: Produced by the EAG from the company's model

6.3.2 Subgroup 2

Table 44 below summarises the results of the scenarios on the EAG base case for subgroup 2. The NMB varied from £9 (ALP+FUL TTD: HR for TTD vs. PFS = 0.2775) to £37,469 (elacestrant OS – log-normal). Elacestrant remained dominant in all scenarios, except when we assumed a hazard ratio for ALP+FUL TTD vs. ALP+FUL PFS of less than 0.6. Fourteen of 24 scenarios varied the ICER by more than 5%.

The scenarios that have the biggest effects on the cost-effectiveness are:

- The results for subgroup 2 are sensitive to assumptions regarding the comparator TTD. We examined this by varying the ALP+FUL TTD relative to the ALP+FUL PFS using an assumed hazard ratio (HR). Elacestrant remains dominant with an assumed HR between 0.6 and 1. Elacestrant is not dominant but has an ICER below £30,000 per QALY with an HR is between 0.2775 and 0.5785. And elacestrant has an ICER above £30,000 per QALY threshold with an HR of less than 0.2775. We estimate the HR at which the mean TTD for elacestrant and ALP+FUL are similar at approximately 0.46, which yields an ICER of £11,519 per QALY.
- Assuming Gamma or log-normal distributions for elacestrant OS increases the NMB by £4,952 and £17,799, respectively.
- Assuming a Weibull distribution for the ALP + FUL OS increases the NMB by £1,372. Whereas a log-normal ALP + FUL OS decreases the NMB by £5,132.
- Assuming a log-normal distribution for elacestrant TTD instead of the KM curve increases the NMB by £2,730, and the log-logistic distribution increases the NMB by £5,455.
- Elacestrant remained dominant for all ESR1 mutation test scenarios. Varying the ESR1-mut testing cost inversely affects the NMB. Increasing the total cost by £7,000 (Marsden assay cost, prevalence-based) decreases the NMB by £7,000. The NMB varied from £12,670 to £20,320.

Table 44 EAG scenario analyses for subgroup 2

EAG base case	Scenario	Treatment	Total cost (£)	Total QALYs	ICER (£/QALY) without the severity modifier	Net monetary benefit (£) at £30,000 per QALY gained
EAG base case		ALP + FUL	[redacted]	[redacted]	Dominant	£19,670
		Elacestrant	[redacted]	[redacted]		
Company base case assumptions						
Mean age from Flatiron	Mean age from EMERALD	ALP + FUL	[redacted]	[redacted]	Dominant	£19,670
		Elacestrant	[redacted]	[redacted]		
Proportion of positive cases after ESR1-mut testing (20%)	Proportion of positive cases after ESR1-mut testing (50%)	ALP + FUL	[redacted]	[redacted]	Dominant	£20,570
		Elacestrant	[redacted]	[redacted]		
Company scenarios presented in the submission						
MAIC approach – independent PSM extrapolation	HR	ALP + FUL	[redacted]	[redacted]	Dominant	£18,441
		Elacestrant	[redacted]	[redacted]		
Elacestrant OS - Weibull	Gamma	ALP + FUL	[redacted]	[redacted]	Dominant	£24,622
		Elacestrant	[redacted]	[redacted]		
	Log-normal	ALP + FUL	[redacted]	[redacted]	Dominant	£37,469
		Elacestrant	[redacted]	[redacted]		

EAG base case	Scenario	Treatment	Total cost (£)	Total QALYs	ICER (£/QALY) without the severity modifier	Net monetary benefit (£) at £30,000 per QALY gained
ALP+FUL OS - Gamma	Weibull	ALP + FUL	[redacted]	[redacted]	Dominant	£21,042
		Elacestrant	[redacted]	[redacted]		
	Log-normal	ALP + FUL	[redacted]	[redacted]	Dominant	£14,538
		Elacestrant	[redacted]	[redacted]		
Elacestrant PFS – log-normal	Log-logistic	ALP + FUL	[redacted]	[redacted]	Dominant	£19,634
		Elacestrant	[redacted]	[redacted]		
	Exponential	ALP + FUL	[redacted]	[redacted]	Dominant	£20,070
		Elacestrant	[redacted]	[redacted]		
ALP+FUL PFS – log-normal	Generalised gamma	ALP + FUL	[redacted]	[redacted]	Dominant	£21,266
		Elacestrant	[redacted]	[redacted]		
	Gamma	ALP + FUL	[redacted]	[redacted]	Dominant	£19,390
		Elacestrant	[redacted]	[redacted]		
Elacestrant TTD – KM curve	Log-normal	ALP + FUL	[redacted]	[redacted]	Dominant	£22,400
		Elacestrant	[redacted]	[redacted]		
	Log-logistic	ALP + FUL	[redacted]	[redacted]	Dominant	£25,125
		Elacestrant	[redacted]	[redacted]		
Progressed utility source – EMERALD EQ-5D ([redacted])	Lloyd et al. 2006 absolute approach	ALP + FUL	[redacted]	[redacted]	Dominant	£18,767
		Elacestrant	[redacted]	[redacted]		

EAG base case	Scenario	Treatment	Total cost (£)	Total QALYs	ICER (£/QALY) without the severity modifier	Net monetary benefit (£) at £30,000 per QALY gained
Company additional scenarios presented in the economic model						
ALP+FUL PFS – log-normal	Weibull	ALP + FUL	[redacted]	[redacted]	Dominant	£19,837
		Elacestrant	[redacted]	[redacted]		
EAG additional scenarios						
ALP + FUL TTD: equal to PFS	HR for TTD vs. PFS = 0.2775	ALP + FUL	[redacted]	[redacted]	£29,969	£9
		Elacestrant	[redacted]	[redacted]		
	HR for TTD vs. PFS = 0.46	ALP + FUL	[redacted]	[redacted]	£11,519	£5,114
		Elacestrant	[redacted]	[redacted]		
	HR for TTD vs. PFS = 0.5785	ALP + FUL	[redacted]	[redacted]	£9	£8,299
		Elacestrant	[redacted]	[redacted]		
	HR for TTD vs. PFS = 0.6	ALP + FUL	[redacted]	[redacted]	Dominant	£8,873
		Elacestrant	[redacted]	[redacted]		
Subsequent treatment cost: 100% capecitabine	Telford et al. 2019 (section 6.1.3)	ALP + FUL	[redacted]	[redacted]	Dominant	£19,672
		Elacestrant	[redacted]	[redacted]		
ESR1-mut testing cost: £300, prevalence-based (20%)=£1,500	NHS GMS, prevalence-based: [redacted] /0.2= [redacted]	ALP + FUL	[redacted]	[redacted]	Dominant	£16,920
		Elacestrant	[redacted]	[redacted]		
	NHS GMS, non-prevalence base: [redacted]	ALP + FUL	[redacted]	[redacted]	Dominant	£20,320
		Elacestrant	[redacted]	[redacted]		

EAG base case	Scenario	Treatment	Total cost (£)	Total QALYs	ICER (£/QALY) without the severity modifier	Net monetary benefit (£) at £30,000 per QALY gained
	Marsden360 assay cost, prevalence-based: [redacted] /0.2 = [redacted]	ALP + FUL	[redacted]	[redacted]	Dominant	£12,670
		Elacestrant	[redacted]	[redacted]		
	Marsden360 assay, non-prevalence base: [redacted]	ALP + FUL	[redacted]	[redacted]	Dominant	£19,470
		Elacestrant	[redacted]	[redacted]		
ESR1-mut testing – 20% of positive tests	10%	ALP + FUL	[redacted]	[redacted]	Dominant	£18,170
		Elacestrant	[redacted]	[redacted]		
End of life cost: Round et al. 2015	Georghiou et al. 2012	ALP + FUL	[redacted]	[redacted]	Dominant	£19,738
		Elacestrant	[redacted]	[redacted]		
Utilities (PF and PD) from EMERALD EQ-5D analysis	Utilities from Lloyd et al. 2006	ALP + FUL	[redacted]	[redacted]	Dominant	£18,706
		Elacestrant	[redacted]	[redacted]		

Source: Produced by the EAG from the company's model

6.4 Conclusions on the cost-effectiveness evidence

The EAG identified a set of assumptions and input parameter values that we prefer to those used in the company's base case analysis. See Table 32 for description and justification for these assumptions.

For subgroup 1, the EAG's preferred assumptions increased the ICER for elacestrant versus everolimus with exemestane from £24,893 to £73,224 per QALY (including a severity modifier of 1.2). The results are most sensitive to changes in the overall survival curve for elacestrant, the everolimus price, and using the MAIC hazard ratio approach, instead of independent parametric distributions for the survival curves.

For subgroup 2, elacestrant remained dominant with the EAG's preferred assumptions, with an NMB of £19,670 at the £30,000 per QALY threshold (no severity modifier is applicable for subgroup 2). The results are most sensitive to changes in the ALP+FUL TTD assumption (assumed equal to the PFS curve in the base case and varied relative to the PFS in EAG scenario analysis), the elacestrant OS and ALP+FUL OS distributions, as well as the ESR1-mut testing cost and proportion of positive ESR1 mutation cases after testing.

The main uncertainties regarding the cost-effectiveness of elacestrant are the following:

- Structural uncertainty relating to the use of a post-hoc subgroup analysis to define the target population and outcomes on the basis of duration of prior treatment (progression after at least 12 months of ET+CDK4/6i).
- The lack of comparative data for elacestrant versus the most relevant current treatment options; and reliance on treatment effects from an unanchored MAIC, with small sample sizes and limited availability of prognostic data.
- Selection of overall survival extrapolations for the company's target population (subgroup 1) and the assumed persistence of the relative treatment benefit.
- Assumptions regarding the duration of treatment for comparators, particularly for patients with a dual ESR1 and PIK3CA mutation (subgroup 2).
- The source used for the price of everolimus (BNF versus eMIT). The cost-effectiveness results in this report are based on a confidential discounted price proposed for elacestrant, but only publicly available prices for other drugs. We present results using all drug price discounts available in the NHS in a confidential addendum to this report.

Finally, we note that there is uncertainty over the cost and practical implications for the NHS of introducing a test for ESR1 mutation when treatment with elacestrant is being considered; using either digital PCR methods that would require a repeat tissue biopsy, or with a ctDNA blood test. ctDNA testing is currently available from the North Thames NHS GLH using the Marsden360 assay. We understand that other NHS GLHs are exploring this or a similar approach., and that the cost could fall if NGS panel testing were to be introduced for ESR1 and additional treatment targets as they become available.

7 SEVERITY MODIFIER

7.1 Severity modifier for the company's base case

The company presented their rationale for applying a severity modifier for QALYs in CS section B.3.6. This was calculated using the QALY shortfall calculator estimator (Schneider et al., 2021).[50] This calculator follows NICE recommended methods in the NICE Health Technology Evaluations manual, section 6.2.[24] The following information is required:

- **Mean age of the patient population:** the calculator only accepts integer numbers for age. Therefore, the company considered [redacted] years old for subgroup 1 and [redacted] for subgroup 2 (see CS Table 39).
- **Discount rate:** 3.5% (cost and QALYs) (see CS B.3.2.2.3)
- **The proportion of females in the patient population:** [redacted] (CS Table 39)
- **Remaining QALYs with the disease (discounted):** the company considered the total discounted QALYs from the comparators' results of [redacted] for subgroup 1 and [redacted] for subgroup 2 (see CS Table 81).
- **Scenario:** "Reference case - MVH value set + HSE 2014 ALDVMM model (Hernandez Alava et al.)"

The EAG verified the severity modifier results reported for the company's base case (CS Table 78). Subgroup 1 met the criteria for a QALY severity weight of 1.2 on the basis of proportional shortfall (85% to 95%), but subgroup 2 did not (see Table 45). Neither subgroup met the requirement for a QALY weight based on absolute QALY shortfall (≥12).

Table 45 Severity modifier estimates for the company's base case

	Subgroup 1	Subgroup 2
Mean age of the patient population	[redacted]	[redacted]
Remaining QALYs without the disease	[redacted]	[redacted]
Remaining QALYs with the disease	[redacted]	[redacted]
Absolute shortfall	[redacted]	[redacted]
Proportional shortfall	[redacted]	[redacted]
QALY weight	**1.2**	**1.0**

Source: Produced by the EAG using the Schneider QALY shortfall calculator and information in the CS and model

We assessed the sensitivity of the severity modifier to the baseline age of the modelled population. Varying the mean age did not affect the severity modifier estimate.

7.2 Severity modifier for the EAG's preferred analysis

Using the EAG's preferred analysis, assumptions (see section 6.2) considered the mean age of the population from the Flatiron instead of the EMERALD estimate. Table 46 below shows that the QALYs' weight remained the same for both subgroups.

Table 46 Severity modifier estimates for the EAG's assumptions

	Subgroup 1	Subgroup 2
Mean age of the patient population	[redacted]	[redacted]
Remaining QALYs without the disease	[redacted]	[redacted]
Remaining QALYs with the disease	[redacted]	[redacted]
Absolute shortfall	[redacted]	[redacted]
Proportional shortfall	[redacted]	[redacted]
QALY weight	**1.2**	**1.0**

Source: Produced by the EAG using the Schneider QALY shortfall calculator and information in the CS and model

8 REFERENCES

1. Chandarlapaty S, Chen D, He W, et al. Prevalence of ESR1 Mutations in Cell-Free DNA and Outcomes in Metastatic Breast Cancer: A Secondary Analysis of the BOLERO-2 Clinical Trial. JAMA Oncol 2016;2(10):1310-15. doi: 10.1001/jamaoncol.2016.1279

2. Turner N, Rugo HS, Ciruelos EM, et al. Abstract PD15-01: Impact of ESR1 mutations on endocrine therapy (ET) plus alpelisib benefit in patients with hormone receptor-positive (HR+), human epidermal growth factor receptor 2-negative (HER2-), PIK3CA-mutated, advanced breast cancer (ABC) who progressed on or after prior cyclin-dependent kinase inhibitor (CDK4/6i) therapy in the BYLieve trial. Cancer Research 2022;82(4_Supplement):PD15-01. doi: 10.1158/1538-7445.SABCS21-PD15-01

3. Clatot F, Perdrix A, Augusto L, et al. Kinetics, prognostic and predictive values of ESR1 circulating mutations in metastatic breast cancer patients progressing on aromatase inhibitor. Oncotarget 2016;7(46):74448-59. doi: 10.18632/oncotarget.12950

4. Kalinsky K, Accordino MK, Chiuzan C, et al. Randomized Phase II Trial of Endocrine Therapy With or Without Ribociclib After Progression on Cyclin-Dependent Kinase 4/6 Inhibition in Hormone Receptor-Positive, Human Epidermal Growth Factor Receptor 2-Negative Metastatic Breast Cancer: MAINTAIN Trial. J Clin Oncol 2023;41(24):4004-13. doi: 10.1200/JCO.22.02392

5. Turner NC, Swift C, Kilburn L, et al. ESR1 Mutations and Overall Survival on Fulvestrant versus Exemestane in Advanced Hormone Receptor-Positive Breast Cancer: A Combined Analysis of the Phase III SoFEA and EFECT Trials. Clin Cancer Res 2020;26(19):5172-77. doi: 10.1158/1078-0432.CCR-20-0224

6. Cardoso F, Costa A, Norton L, et al. ESO-ESMO 2nd international consensus guidelines for advanced breast cancer (ABC2). Breast 2014;23(5):489-502. doi: 10.1016/j.breast.2014.08.009

7. NICE. Everolimus with exemestane for treating advanced breast cancer after endocrine therapy: TA421: National Institute for Health and Care Excellence, 2016.

8. NICE. Alpelisib with fulvestrant for treating hormone receptor-positive, HER2-negative, PIK3CA-mutated advanced breast cancer: TA816: National Institute for Health and Care Excellence, 2022.

9. Bardia A. PS17-02 Elacestrant vs standard-of-care in ER+/HER2- advanced or metastatic breast cancer (mBC) with ESR1 mutation: key biomarkers and clinical subgroup analyses from the phase 3 EMERALD trial. SABCS 2023

10. Bardia A, Bidard F-C, Neven P, et al. Abstract GS3-01: GS3-01 EMERALD phase 3 trial of elacestrant versus standard of care endocrine therapy in patients with ER+/HER2- metastatic breast cancer: Updated results by duration of prior CDK4/6i in metastatic setting. Cancer Research 2023;83(5_Supplement):GS3-01. doi: 10.1158/1538-7445.SABCS22-GS3-01

11. Bidard F-C, Kaklamani VG, Neven P, et al. Elacestrant (oral selective estrogen receptor degrader) Versus Standard Endocrine Therapy for Estrogen Receptor-Positive, Human Epidermal Growth Factor Receptor 2-Negative Advanced Breast Cancer: Results From the Randomized Phase III EMERALD Trial. J Clin Oncol 2022;40(28):3246-56. doi: 10.1200/JCO.22.00338

12. Rugo HS, Neven P, Saffie I, et al. Abstract PD13-05: Alpelisib + fulvestrant in patients with PIK3CA-mutated, HR+, HER2— advanced breast cancer (ABC) who received chemotherapy or endocrine therapy (ET) as immediate prior treatment: BYLieve Cohort C primary results and exploratory biomarker analyses. Cancer Research 2022;82(4_Supplement):PD13-05. doi: 10.1158/1538-7445.SABCS21-PD13-05

13. Raphael A, Salmon-Divon M, Epstein J, et al. Alpelisib Efficacy in Hormone Receptor-Positive HER2-Negative PIK3CA-Mutant Advanced Breast Cancer Post-Everolimus Treatment. Genes (Basel) 2022;13(10) doi: 10.3390/genes13101763

14. Stemline Menarini. UK Clinical Expert Opinion, Data on File, 2024.

15. NICE. Fulvestrant for the treatment of locally advanced or metastatic breast cancer TA239: National Institute for Health and Care Excellence, 2011.

16. Centre for Reviews and Dissemination. Systematic reviews: CRD's guidance for undertaking reviews in health care: York Publishing Services Ltd.: CRD 2009.

17. Stemline Menarini. Clinical Study Report v.2: Elacestrant monotherapy vs. standard of care for the treatment of patients with ER+/HER2- advanced breast cancer following CDK4/6 inhibitor therapy: A phase 3 randomized, open-label, active-controlled, multicenter trial (EMERALD), 2023:4424.

18. NICE. Final scope: Elacestrant for treating oestrogen receptor-positive, HER2-negative advanced breast cancer with an ESR1 mutation after at least 1 endocrine treatment [ID6225]: National Institute for Health and Care Excellence, 2024.

19. Stemline Menarini. Clinical Study Protocol (RAD1901-308) | Elacestrant monotherapy vs. standard of care for the treatment of patients with ER+/HER2- advanced breast cancer following CDK4/6 inhibitor therapy: a Phase 3 randomized, open-label, active-controlled, multicentre trial (EMERALD), 2020:118.

20. Stemline Menarini. Overall Survival Addendum to Clinical Study Report v.2 RAD1901-308 | Elacestrant monotherapy vs. standard of care for the treatment of patients with ER+/HER2- advanced breast cancer following CDK4/6 inhibitor therapy: a Phase 3 randomized, open-label, active-controlled, multicenter trial (EMERALD) | overall survival addendum as of 2nd September 2022, 2023:14.

21. Cuyún Carter G, Mohanty M, Stenger K, et al. Prognostic Factors in Hormone Receptor-Positive/Human Epidermal Growth Factor Receptor 2-Negative (HR+/HER2-) Advanced Breast Cancer: A Systematic Literature Review. Cancer Manag Res 2021;13:6537-66. doi: 10.2147/cmar.S300869 [published Online First: 20210820]

22. Vidal G, Cinfio F, Badaracco J, et al. CO134 Estimating the Direct and Indirect Resource Burden of Treatment Management with Current Standard of Care or Elacestrant for ER+, HER2-, ESR1-Mutated Advanced or Metastatic Breast Cancer Patients: A Population-Level Provider Model. Value in Health 2023;26(12):S38.

23. Zeng W, Cao X, Lin J, et al. Cost-effectiveness analysis of elacestrant versus standard endocrine therapy for second-/ third-line treatment of patients with HR+/HER2- advanced or

metastatic breast cancer: a US payer perspective. Front Oncol 2023:11. doi: 10.3389/fonc.2023.1272586

24. NICE. Introduction to health technology evaluation | NICE health technology evaluations: the manual | Guidance | NICE. 2022

25. Hernandez M, Pudney S. eq5dmap: A command for mapping between EQ-5D-3L and EQ-5D-5L. The Stata journal 2017;18 doi: 10.1177/1536867X1801800207

26. Woods B, Sideris E, Palmer S, et al. NICE DSU technical support document 19: partitioned survival analysis for decision modelling in health care: a critical review: University of Sheffield, 2017.

27. Guyot P, Ades A, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC medical research methodology 2012;12:1-13.

28. Hagiwara Y, Shiroiwa T, Shimozuma K, et al. Impact of Adverse Events on Health Utility and Health-Related Quality of Life in Patients Receiving First-Line Chemotherapy for Metastatic Breast Cancer: Results from the SELECT BC Study. Pharmacoeconomics 2018;36(2):215-23. doi: 10.1007/s40273-017-0580-7

29. Mistry R, May JR, Suri G, et al. Cost-Effectiveness of Ribociclib plus Letrozole Versus Palbociclib plus Letrozole and Letrozole Monotherapy in the First-Line Treatment of Postmenopausal Women with HR+/HER2- Advanced or Metastatic Breast Cancer: A U.S. Payer Perspective. J Manag Care Spec Pharm 2018;24(6):514-23. doi: 10.18553/jmcp.2018.24.6.514

30. Lloyd A, Nafees B, Narewska J, et al. Health state utilities for metastatic breast cancer. Br J Cancer 2006;95(6):683-90. doi: 10.1038/sj.bjc.6603326

31. Telford C, Bertranou E, Large S, et al. Cost-Effectiveness Analysis of Fulvestrant 500 mg in Endocrine Therapy-Naïve Postmenopausal Women with Hormone Receptor-Positive Advanced Breast Cancer in the UK. Pharmacoecon Open 2019;3(4):559-70. doi: 10.1007/s41669-019-0134-3

32. Swinburn P, Lloyd A, Nathan P, et al. Elicitation of health state utilities in metastatic renal cell carcinoma. Current Medical Research and Opinion 2010;26(5):1091–96. doi: https://doi.org/10.1185/03007991003712258

33. Smith-Palmer J, Bae JP, Boye KS, et al. Evaluating health-related quality of life in type 1 diabetes: a systematic literature review of utilities for adults with type 1 diabetes. Clinicoecon Outcomes Res 2016;8:559-71. doi: 10.2147/CEOR.S114699

34. NICE. Ribociclib with an aromatase inhibitor for previously untreated, hormone receptor-positive, HER2-negative, locally advanced or metastatic breast cancer: TA496: National Institute for Health and Care Excellence, 2017.

35. NICE. Fulvestrant for untreated locally advanced or metastatic oestrogen-receptor positive breast cancer: TA503: National Institute for Health and Care Excellence, 2018.

36. NICE. Abemaciclib with an aromatase inhibitor for previously untreated, hormone receptor-positive, HER2-negative, locally advanced or metastatic breast cancer: TA563: National Institute for Health and Care Excellence, 2019.

37. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. Value Health 2010;13(5):509-18. doi: 10.1111/j.1524-4733.2010.00700.x

38. NICE. BNF: British National Formulary - NICE. 2021

39. eMIT. Drugs and pharmaceutical electronic market information tool (eMIT). GOVUK 2023

40. Jerusalem G, Mariani G, Ciruelos EM, et al. Safety of everolimus plus exemestane in patients with hormone-receptor-positive, HER2-negative locally advanced or metastatic breast cancer progressing on prior non-steroidal aromatase inhibitors: primary results of a phase IIIb, open-label, single-arm, expanded-access multicenter trial (BALLET). Ann Oncol 2016;27(9):1719-25. doi: 10.1093/annonc/mdw249

41. Alaklabi S, Roy AM, Attwood K, et al. Real world outcomes with alpelisib in metastatic hormone receptor-positive breast cancer patients: A single institution experience. Front Oncol 2022;12:1012391. doi: 10.3389/fonc.2022.1012391

42. PSSRU. Unit Costs of Health and Social Care. 2022

43. NHS. National Cost Collection for the NHS. 2022

44. NICE. Palbociclib with fulvestrant for treating hormone receptor-positive, HER2-negative, advanced breast cancer: TA619, 2020.

45. NICE. Eribulin for treating locally advanced or metastatic breast cancer after 2 or more chemotherapy regimens: TA423: National Institute for Health and Care Excellence, 2016.

46. Jhaveri KL, Jeselsohn R, Ma CX, et al. Imlunestrant, with or without everolimus or alpelisib, in ER+, HER2- advanced breast cancer (aBC): Results from the phase 1a/b EMBER study.

47. Lin NU, Borges VF, Patel MR, et al. Results from the phase 1/2 study of OP-1250, an oral complete estrogen receptor (ER) antagonist (CERAN) and selective ER degrader (SERD) in patients (pts) with advanced or metastatic ER-positive, HER2-negative breast cancer. Annals of Oncology;34(S2):S338. doi: https://doi.org/10.1016/j.annonc.2023.09.559

48. Round J, Jones L, Morris S. Estimating the cost of caring for people with cancer at the end of life: A modelling study. Palliat Med 2015;29(10):899-907. doi: 10.1177/0269216315595203

49. Georghiou T, Davies S, Davies A, Bardsley M. Understanding patterns of health and social care at the end of life: Nuffield Trust, 2012.

50. Schneider P, McNamara S, Love-Koh J, et al. QALY Shortfall Calculator.

9 APPENDICES

Appendix 1 EAG assessment of company's clinical effectiveness systematic literature review methods

Table 47 EAG appraisal of systematic review methods

Systematic review components and processes	EAG response	EAG comments
Was the review question clearly defined using the PICOD framework or an alternative?	Yes	CS Appendix D Table 4 provides details of the eligibility criteria for the initial clinical SLR (referred to in the submission as "the global clinical SLR"). Criteria were appropriate but broader for interventions and comparators than that of the NICE final scope. CS Appendix D Table 5 provides details of narrower eligibility criteria that aligned with the NICE final scope. These eligibility criteria were appropriate in terms of the appraisal and were used to rescreen included studies identified from the initial SLR
Were appropriate sources of literature searched?	Yes	Searches covered sufficient databases (MEDLINE (Ovid), Embase (Ovid), Cochrane (CENTRAL and CDSR; Ovid)) Relevant grey literature was also searched (conference proceedings from global, US, European and Australasian breast cancer meetings; Government/international bodies; reference lists of included studies)
What time period did the searches span and was this appropriate?	Yes	Database searches were carried out from inception to August 2023. Searches of conference proceedings were limited to meetings held in 2020 to 2023 inclusive. The searches were

Systematic review components and processes	EAG response	EAG comments
		approximately 8 months old when the CS was received by the EAG. The EAG therefore reran the searches with a date limit for the past 8 months.
Were appropriate search terms used and combined correctly?	Yes	Search strategies for MEDLINE, Embase and Cochrane are reported in CS Appendix D.1.1. The searches used an appropriate set of terms to specify the type of breast cancer relevant to the appraisal combined with a broad range of interventions/ comparators including, but not limited to, those for the appraisal. The RCT filter used in the company searches however excludes conference abstracts. The EAG therefore reran the Embase search for the past three years using terms that would include conference abstracts.
Were inclusion and exclusion criteria specified? If so, were these criteria appropriate and relevant to the decision problem?	Yes	CS Appendix D Table 4 provides details of the initial SLR eligibility criteria, which were appropriate but broader for interventions and comparators than that of the NICE final scope. Appendix D Table 5 provides details of the narrower eligibility criteria, which aligned with the NICE final scope. These eligibility criteria were applied to the included studies identified from the broader SLR and were appropriate.
Were study selection criteria applied by two or more reviewers independently?	Yes	For the initial broader SLR, titles and abstracts and full papers were screened by two independent reviewers. Discrepancies between the reviewers

Systematic review components and processes	EAG response	EAG comments
		was reconciled through consensus (titles and abstracts) or a third independent reviewer (titles and abstracts, full papers) The included publications from the initial SLR were rescreened by two independent reviewers using the narrower eligibility criteria aligned with the NICE scope. Any discrepancies were resolved by a third independent reviewer.
Was data extraction performed by two or more reviewers independently?	Yes	Data were extracted by one reviewer and checked by a second reviewer. The EAG considers this acceptable.
Was a risk of bias assessment or a quality assessment of the included studies undertaken? If so, which tool was used?	Yes	The company used the seven-criteria checklist recommended by NICE, based on guidance provided by CRD (CS Appendix D.2.4).
Was risk of bias assessment (or other study quality assessment) conducted by two or more reviewers independently?	Unclear	The CS does not state how the risk of bias assessments were conducted.
Is sufficient detail on the individual studies presented?	Yes	CS section B.2.1 to 2.7, CS Appendix D.2.3 and D.2.4, and CS Appendix E provide methodological details and results from the single relevant trial (EMERALD) identified for this appraisal. The trial CSR was also provided.
If statistical evidence synthesis (e.g. pairwise meta-analysis, ITC, NMA)	Yes	Due to the absence of individual patient-level data for the comparators and a lack of common comparator, an unanchored

Systematic review components and processes	EAG response	EAG comments
was undertaken, were appropriate methods used?		MAIC was implemented to facilitate an ITC for two outcomes (OS and PFS). Our critique of the MAIC is provided in section 3.4 of this report

Source: Table created by the EAG
CDSR, Cochrane Database of Systematic Reviews; CENTRAL, Cochrane Central Register of Controlled Trials; CRD, Centre for Reviews and Dissemination; CS, company submission; CSR, clinical study report; EAG, External Assessment Group; ITC, indirect treatment comparison; MAIC, matching-adjusted indirect comparison; NMA, network meta-analysis; OS, overall survival; PFS, Progression-free survival; PICOD, population, intervention, comparator, outcome, design; RCT, randomised controlled trial; SLR, systematic literature review

Appendix 2 Prognostic factors included in the company's MAIC

Table 48 below compares prognostic factors identified by Cuyún Carter et al (2021) with those proposed by key opinion leaders consulted by the company. For further discussion please see section 3.4.2 of this EAG report.

Table 48 Comparison of prognostic factors identified by a systematic review by Cuyún Carter et al (2021) with factors proposed by key opinion leaders, and their inclusion status in the MAIC

Prognostic factors with strongest evidence of association with[a]:		Prognostic factors/effect modifiers identified by key opinion leaders (KOLs)[b]	Included in MAIC?
worse OS	worse PFS		
Negative progesterone receptor status		ER expression	*Partial* - Included implicitly through population restriction (focus on ESR1-mut)
Higher tumour grade		Not identified by KOLs	No - Not identified by KOLs
Higher circulating tumour cell (CTC) count and higher Ki67 level	Higher circulating tumour cell (CTC) count,	Not identified by KOLs	No - Not identified by KOLs
Number of metastatic sites (e.g. multiple vs single)	Number and sites of metastases	Number of metastatic sites	No - excluded due to lack of data
Sites of metastases (e.g. presence of liver metastases vs absence),		Bone metastases / bone metastases only; Visceral metastases	No - excluded due to lack of data

Prognostic factors with strongest evidence of association with[a]:		Prognostic factors/effect modifiers identified by key opinion leaders (KOLs)[b]	Included in MAIC?
worse OS	**worse PFS**		
Shorter time to recurrence or progression to advanced breast cancer		Time since original diagnosis	No - discrepancy in data available (only time since stage III diagnosis in Flatiron study)
Poor performance status		ECOG performance status	***Partial*** – approx. 25% of patients had missing performance status.
Prior therapy attributes in the early or metastatic setting (type of therapy, treatment line, response of prior therapy)	Absence of prior therapy or higher lines of therapy in the early or metastatic setting	Length of time on prior CDK4/6i;	***Partial*** - Included implicitly through population restriction (prior CDK4/6i ≥12 months)
		Number of treatment lines in metastatic setting;	**Yes** – for ET lines. Number of prior ET included as only number of prior lines of ET available
		Prior chemotherapy	Yes
Race (black vs white).		Not identified by KOLs	No - Not identified by KOLs
		Histology (ductal vs. lobular)	No - Excluded due to lack of data
		De novo vs. recurrent disease (i.e. diagnosed in adjuvant setting)	No - Excluded due to lack of data
		De novo vs. progressed disease	No - Excluded due to lack of data
		Age	**Yes** - Flatiron patients restricted to 50 years or older

Prognostic factors with strongest evidence of association with[a]:		Prognostic factors/effect modifiers identified by key opinion leaders (KOLs)[b]	Included in MAIC?
worse OS	**worse PFS**		
		Menopausal status	***Partial***. Included implicitly through a focus on postmenopausal women in EMERALD and older women in Flatiron.

Source: reproduced, in part, from CS Table 25
[a] as identified by a systematic review of prognostic factors by Cuyún Carter et al (2021).
[b] As identified through consultation by the company with key opinion leaders (see CS Section B.2.9.1 and CS Table 25)
Dark shaded cells indicate that the prognostic factor was not included in the sub-set of factors judged by Cuyún Carter et al (2021) as having the strongest evidence of association with health outcomes.

Appendix 3 Survival extrapolations: Target population (subgroup 1)

Figure 11 Everolimus + exemestane OS for subgroup 1

Source: Produced from the company's model by the EAG

Figure 12 Elacestrant OS for subgroup 1

Source: Produced from the company's model by the EAG

Figure 13 Everolimus + exemestane PFS for subgroup 1

Source: Produced from the company's model by the EAG

Figure 14 Elacestrant PFS for subgroup 1

Source: Produced from the company's model by the EAG

Figure 15 Elacestrant TTD for subgroup 1

Source: Produced from the company's model by the EAG

Figure 16 Everolimus + exemestant outcomes, subgroup 1 (company base case)

Source: Produced by the EAG from the company model

Figure 17 Elacestrant outcomes, subgroup 1 (company base case)

Source: Produced by the EAG from the company model

Appendix 4 Survival extrapolations: Dual mutated (subgroup 2)

Figure 18 Alpelisib + fulvestrant OS for subgroup 2

Source: Produced from the company's model by the EAG

Figure 19 Elacestrant OS for subgroup 2

Source: Produced from the company's model by the EAG

Figure 20 Alpelisib + fulvestrant PFS for subgroup 2

Source: Produced from the company's model by the EAG

Figure 21 Elacestrant PFS for subgroup 2

Source: Produced from the company's model by the EAG

Figure 22 Elacestrant TTD for subgroup 2

Source: Produced from the company's model by the EAG

Figure 23 Alpelisib + fulvestrant outcomes, subgroup 2 (company base case)

Source: Produced by the EAG from the company model

Figure 24 Elacestrant outcomes, subgroup 2 (company base case)

Source: Produced by the EAG from the company model

Appendix 5 Additional company's scenario analysis

The company's economic model has a scenario module with additional scenarios described below:

Severity modifier: do not consider a severity modifier for subgroup 1 (base case 1.2)

- Elacestrant OS:
  - Subgroup 1: additional scenarios with the exponential (worst BIC fit), generalised Gamma, Gompertz and Weibull (best BIC fit) distributions
  - Subgroup 2: additional scenarios with the exponential (worst statistical fit), generalised Gamma, Gompertz (best statistical fit) and log-logistic distributions
- Comparator OS:
  - Subgroup 1: additional scenarios with the generalised gamma (worst statistical fit), Gompertz, Log-logistic and Log-normal distributions
  - Subgroup 2: additional scenarios with the generalised Gamma, Gompertz, Log-logistic and exponential (worst statistical fit) distributions
- Elacestrant PFS:
  - Subgroup 1: additional scenarios with the exponential, generalised gamma (best statistical fit), Gompertz, Weibull (worst statistical fit), and Gamma distributions
  - Subgroup 2: additional scenarios with the generalised gamma (best statistical fit), Gompertz (worst statistical fit), Weibull, and gamma distributions
- Comparator PFS:
  - Subgroup 1: additional scenarios with the exponential (worst statistical fit), generalised Gamma, Gompertz, and Weibull distributions
  - Subgroup 2: additional scenarios with the exponential (worst statistical fit), log-logistic, Gompertz, and Weibull distributions
- **Elacestrant TTD:** additional scenarios with the exponential, generalised gamma (best statistical fit), Gompertz, Weibull and Gamma (worst statistical fit) distributions
- **ESR1-mut testing cost:** consider the user-defined cost (base case: digital PCR cost)
- **ESR1-mut testing cost approach:** consider non-prevalence-based (base case: prevalence-based)
- **PF health state utility source:** use PF utilities from previous assessments as TA563, TA496, TA503 (base case: EMERALD)
- **PD health state utility source:** use PD utilities from previous assessments as TA563, TA496, TA503 (base case: EMERALD)
- **Health state utility source:** consider a user-defined utility (base case: EMERALD)

- **Capecitabine dose:** consider the minimum (1,000 mg/m^2) and maximum doses (1,250 mg/m^2) (base case: average dose, 1,125 mg/m^2)

For subgroup 1, the ICER varied from £22,804 (elacestrant OS extrapolation using exponential) to £151,291 (elacestrant OS extrapolation using Gompertz distribution). The non-cost-effective scenarios are related to the OS extrapolations for elacestrant and the comparator everolimus + exemestane. Two scenarios are not cost-effective, and the relative QALYs shortfall indicated that the severity modifier 1.2 did not apply to them: the log-logistic distribution was the second-best fit, and the log-normal distribution was the second-worst fit to the OS extrapolation for the comparator (everolimus + exemestane).

For subgroup 2, elacestrant is dominant for all additional scenarios. One scenario could not be performed owing to a lack of model convergence to fit the generalised gamma as an OS extrapolation for elacestrant. Three scenarios modified the total discounted QALYs to a value where the severity modifier 1.2 could be applied: progression disease health state utilities from TA563 and TA496 and generalised gamma as OS extrapolation for alpelisib + fulvestrant.