Relugolix with oestradiol and norethisterone acetate for treating uterine fibroids [ID3842]

Produced by Aberdeen HTA Group

Authors

Moira Cruickshank[1]
Elisabet Jacobsen[2]
Thenmalar Vadivaloo[1]
Mari Imamura[1]
David Cooper[1]
Paul Manson[1]
Kevin Cooper[3]
Dwayne Boyers[2]
Miriam Brazzelli[1]

1 Health Services Research Unit, University of Aberdeen, UK
2 Health Economics Research Unit, University of Aberdeen, UK
3 NHS Grampian, Aberdeen, UK

Correspondence to Miriam Brazzelli
Health Services Research Unit, University of Aberdeen
3rd Floor, Health Sciences Building, Foresterhill
Aberdeen, AB25 2ZD
m.brazzelli@abdn.ac.uk

Date completed **13 January 2022**
Version **2.0 (post company factual accuracy check)**

Source of funding: This report was commissioned by the NIHR Systematic Reviews Programme as project number 134488.

Declared competing interests of the authors

No competing interests to disclose.

Acknowledgements

Copyright is retained by Gedeon Richter for Figures 1-8, Tables 6-21, 26, 32 and 33 and text referenced on page 45.

Rider on responsibility for report

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

This report should be referenced as follows:
Cruickshank M, Jacobsen E, Vadivaloo T, Imamura M, Cooper D, Manson P, Cooper K, Boyers D, Brazzelli M. Relugolix with oestradiol and norethisterone acetate for treating uterine fibroids [ID3842]. Aberdeen HTA Group, 2021.

Contribution of authors

Mari Imamura and Moira Cruickshank summarised and critiqued the clinical effectiveness evidence; David Cooper and Thenmalar Vadiveloo checked and critiqued the statistical analyses presented in the company submission; Dwayne Boyers and Elisabet Jacobsen reviewed and critiqued the cost-effectiveness evidence; Paul Manson checked and critiqued the company's search strategies; Kevin Cooper provided clinical guidance and comments on the draft report. Miriam Brazzelli oversaw and coordinated all aspects of this appraisal. All authors contributed to the writing of this report and approved its final version.

Table of contents

List of tables

List of figures

List of abbreviations

AE	Adverse event
AH	Alkaline haematin
BMD	Bone mineral density
BMI	Body mass index
CI	Confidence interval
CS	Company submission
CSR	Clinical study report
CT	Combination therapy
DECA scan	Dual-energy X-ray absorptiometry scan
EMA	European Medicines Agency
EQ-5D	European quality of life five dimension
ERG	Evidence review group
GnRH	Gonadotropin-releasing hormone
GnRHa	Gonadotropin-releasing hormone analogue
HMB	Heavy menstrual bleeding
HRQoL	Health-related quality of life
ICER	Incremental cost-effectiveness ratio
KOL	Key opinion leader
LNG-IUS	Levonorgestrel-releasing intrauterine system
LS	Least squares
MBL	Menstrual blood loss
MHRA	Medicines and Healthcare products Regulatory Agency
mITT	Modified intention to treat
NHS	(UK) National Health Service
NICE	National Institute for Health and Care Excellence
NRS	Numerical rating scale
NSAID	Non-steroidal anti-inflammatory drug
OLS	Ordinary least squares
PSSRU	Personal social services research unit
QALY	Quality adjusted life year

QoL	Quality of life
RCT	Randomised controlled trial
SAE	Serious adverse event
SD	Standard deviation
SE	Standard error
SF-36	Short form 36-item survey
SmPC	Summary of product characteristics
TEAE	Treatment emergent adverse event
TTO	Time trade off
UAE	Uterine artery embolisation
UF	Uterine fibroids
UFS-QoL	Uterine fibroid health and symptom-related quality of life
UPA	Ulipristal acetate
UFV	Uterine fibroid volume
UV	Uterine volume
WTP	Willingness to pay

1. Executive Summary

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

Section 1.1 provides an overview of the key issues. Section 1.2 provides an overview of key model outcomes and the modelling assumptions that have the greatest effect on the ICER. Sections 1.3 to 1.6 explain the key issues in more detail. Background information on the condition, technology and evidence and information on non-key issues are in the main ERG report.

All issues identified represent the ERG's view, not the opinion of NICE.

1.1 Overview of the ERG's key issues

The focus of the submission received from Gedeon Richter is relugolix with oestradiol and norethisterone acetate (referred to throughout as relugolix combination therapy or relugolix CT) for treating moderate to severe symptoms associated with uterine fibroids (UF).

The clinical evidence submitted by the company consists of two recent multicentre Phase-3 trials, LIBERTY 1 and LIBERTY 2, and one Phase-3 open-label extension study of LIBERTY 1 and LIBERTY 2, LIBERTY 3. The clinical outcomes used in the economic model are menstrual blood loss (MBL) volume, change in MBL volume, adverse effects, and quality of life. In LIBERTY 1 and LIBERTY 2 bleeding control was achieved by a higher proportion of participants treated with relugolix CT (73.4% in LIBERTY 1 and 71.2% in LIBERTY 2) compared with those treated with placebo (18.9% and 14.7%, respectively; $p<0.0001$ for both comparisons). Similarly, amenorrhea was achieved by 52% and 50% of participants treated with relugolix CT in LIBERTY 1 and LIBERTY 2, respectively, compared with 6% and 3% of those treated with placebo ($p<0.001$ for both comparisons).

Since the absence of head-to-head RCTs comparing relugolix CT with GnRH agonists, the company conducted an ITC. Apart from LIBERTY 1 and LIBERTY 2,

the other trials included in the ITC were two Phase-3 RCTs: PEARL I and PEARL II that assessed women who were waiting for surgery. PEARL I compared UPA versus placebo and PEARL II UPA versus leuprolide acetate. The company present the results of an ITC of relugolix CT versus UPA but not of relugolix CT versus GnRHa. Results were only presented for the mean difference in percentage change from baseline in MBL and hence uncertainty surrounding the treatment effect was not incorporated into the economic model, substantially under estimating uncertainty in the ICER. The ITC results suggest that relugolix CT and UPA are equally effective in reducing MBL volume. The ERG notes, however, that the patient populations in the LIBERTY and PEARL I trials are different in terms of planned surgery. The company did not present any other comparisons for relugolix CT apart from that versus placebo despite several other outcomes being listed in their scope.

The cost-effectiveness evidence presents a Markov state transition model to calculate expected costs and quality-adjusted life-years (QALYs) associated with relugolix CT or GnRH agonists for the medical management of moderate or severe symptomatic fibroids in pre-menopausal women (average age 42). The cohort enters the model in the 'on treatment' state where they receive either relugolix CT or GnRH agonists. Upon treatment discontinuation, informed by the LIBERTY study (relugolix CT) and PEARL II study / clinical expert opinion (GnRHa) the cohort either enter best supportive care (defined as minimal treatment with iron supplements and NSAIDs) or are listed for surgery. The company base case model assumes that patients can only be listed for surgery following treatment discontinuation and must enter a waiting list of 15 months before surgery is delivered. A maximum of two rounds of surgery are modelled. After age 51, the full cohort enters the menopause health state where they are assumed to be cured, incurring general population utility values, and can only exit this state to enter the "death" state based on all-cause mortality rates. The ERG raises several key issues of uncertainty surrounding the company's modelling approach and data inputs (See Section 1.5 and Chapter 4 of the ERG report).

Table 1 presents a summary of the key issues identified by the ERG.

Table 1 Summary of the key issues identified by the ERG

Issue no.	Summary of key issues	Report sections
Issue 1	Differences between the LIBERTY and PEARL trials in terms of the patient population and the use of relugolix CT and GnRH agonists in UK clinical practice.	Section 2.3 & 4.2.3.
Issue 2	Lack of formal comparison between relugolix CT and GnRH agonsits.	Section 3.4, 3.5, 4.2.6 and 5.1
Issue 3	The appropriateness of using "treatment" rather than "health" states in the economic model structure.	Section 4.2.2
Issue 4	The most appropriate assumptions about treatment discontinuation in UK clinical practice for both relugolix CT and GnRH agonists	Section 4.2.6
Issue 5	The appropriateness of a 'waiting time' health state post-treatment discontinuation	Section 4.2.2 and 4.2.6
Issue 6	The role of surgery in the treatment pathway and the lack of data to inform transitions to the surgery health state	Section 4.2.2 and 4.2.6
Issue 7	Uncertainty surrounding the utility function	Section 4.2.7
Issue 8	Monitoring and follow up resource use in UK clinical practice	Section 4.2.8

The key differences between the company's preferred assumptions and evidence, and the ERG's preferred assumptions and evidence are:

- The company have provided a cost-effectiveness case that appears to be primarily for a group of women who are not scheduled to have surgery but to inform the economic model use data from the comparator study (PEARL II) where women are scheduled for surgery at baseline. The ERG would have preferred these two populations to be considered separately in the economic model.

- The company's ITC only considers one outcome (% change in menstrual blood loss from baseline) for the comparison between relugolix CT and GnRH

agonists but fails to provide estimates or measures of uncertainty surrounding the treatment effect. The ERG is of the opinion that more complete ITCs should have been undertaken to assess relugolix CT versus GnRH agonists, including all relevant clinical outcomes and with results accompanied by appropriate confidence intervals.

- The company prefers an economic model structure based on 'treatment' states whereas the ERG prefers an economic model structure based on 'health' states, defined according to symptom control.

- The company prefers to modify treatment discontinuation data from the LIBERTY study, based on the assumptions of clinical expert opinion that discontinuation in the trial over-estimates discontinuation in real-world clinical practice. The ERG prefers the use of relugolix CT treatment discontinuation data sourced directly from the LIBERTY study because it is more consistent with the costs required to deliver the modelled treatment benefit and also ensures consistency with the data collected in the PEARL II study for GnRH agonists.

- The company prefers a modelling assumption where women can only be listed for surgery after treatment discontinuation, when they enter a 'waiting time' state of duration 15 months. The ERG considers it more appropriate to remove the waiting time state because, in clinical practice, most women listed for surgery would continue to receive the primary treatment in preparation for surgery.

- The company has included the key clinical outcome from the ITC (MBL) as a fixed-point estimate in the economic model, but the ERG prefers full incorporation of uncertainty surrounding the treatment effects for relugolix CT vs. GnRH agonists and relugolix CT vs. BSC into the probabilistic analyses.

- The company uses a mapping algorithm to transform disease-specific quality of life (UFS-QoL) to generic EQ-5D and uses a linear (OLS) utility function to model the impact of MBL on mapped EQ-5D values. The ERG would prefer more details in support of the chosen model structure and how it was derived.

Based on the currently available information, the ERG considers data from the repeated measures model provided by the company in response to clarification queries (with reporting error corrected post FAC) to be more appropriate to allow estimation of appropriate standard errors for inclusion in the probabilistic analysis.

- The company assume that all patients (whether on active treatment or BSC) will receive annual examination scans, but only patients on active treatment will receive gynaecologist appointments (6-monthly). The ERG would ideally prefer a model structure that allows follow-up resource use to be linked to the patient's symptom control ('health' states) rather than their 'treatment' received (other than for Dexa- scans). In a 'treatment' state model, the ERG prefers lower resource use: a one-off gynaecologist appointment and scan to make a treatment plan whenever treatment is started or discontinued.

1.2 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 quality of life associated with improved symptom control (MBL) over a longer treatment duration on relugolix CT compared to GnRHa (obtained from a linear additive utility function estimating the effect of changes in MBL on EQ-5D mapped utilities).
- Reducing the proportion and duration of BSC treatment (with lower utility compared to active treatment) for relugolix CT compared to GnRHa.
- Reducing disutilities associated with surgery-related health states, including surgery waiting time, experience of surgery, surgery adverse events, loss of uterus by treating people with active treatment for longer, until menopause. Lower utilities are partially offset by applying general population utilities in the proportion of people assumed to be cured after surgery.

- A negligible impact on QALYs of treatment-related adverse events and a slightly reduced risk of surgical mortality in the relugolix CT arm due to longer treatment duration preventing surgery, but the impact on total QALYs is negligible.

Overall, the technology is modelled to affect costs by:

- Increasing the treatment acquisition costs, due to longer treatment duration with relugolix CT compared to GnRH agonists
- Reducing the costs of BSC and surgery due to longer time on treatment.

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

- Decisions about the role of surgery in the treatment pathway.
- Assumptions regarding treatment discontinuation for both relugolix CT and GnRHa over time.
- The assumption that people can only be listed for surgery after treatment discontinuation and must enter a waiting time state of duration = 15 months prior to surgery where no active treatment is provided.
- The uncertainty surrounding the menstrual blood loss treatment effect for relugolix CT versus GnRH agonists and versus best supportive care.
- Decisions about the most appropriate utility function used to estimate the impact of MBL on mapped utility values.
- Assumptions about the most appropriate follow-up resource use for patient monitoring and what constitutes BSC in UK clinical practice.

1.3 The decision problem: summary of the ERG's key issues

The ERG notes that the patient population in the LIBERTY trials does not match that of the PEARL trials in terms of planned surgery (see Issue 1 below).

Issue 1 Differences between the LIBERTY and PEARL trials in terms of the patient population

Report section	Section 2.3 (Table 3) & Section 4.2.3.
Description of issue and why the ERG has identified it as important	The patient population assessed in the LIBERTY trials does not match that assessed in the PEARL trials. In the PEARL trials, all women had surgery planned after 13 weeks while planned surgery was an exclusion criterion for the LIBERTY trials and, therefore, it is unlikely that in the LIBERTY trials women would be receiving surgery and certainly not within 13 weeks. The company submission suggests that the company wish to position relugolix CT as a treatment for women who wish to delay or avoid surgery which is similar to the LIBERTY trials (relugolix CT), but the ERG note that it may also be used in clinical practice as a 'pre surgery' treatment which would be more consistent with the population in the PEARL II study (GnRHa).
What alternative approach has the ERG suggested?	As the trials have been conducted in different patient populations the ERG does not have an alternative approach to suggest. However, as the results of the ITC are used in the economic model there are possible scenarios analyses to consider addressing this concern (see Issue 6 below)
What is the expected effect on the cost-effectiveness estimates?	It is difficult to judge the exact impact on the ICER, but the ERG notes that scenarios that remove "waiting time" and "surgery" states from the economic model (approximates subgroup A) increase the ICER substantially. For subgroup B, short-term treatment for 6 months pre-surgery, the company submission provides no evidence of a difference in clinical effectiveness, so it would be reasonable to consider an analysis assuming equal effectiveness. In this case, the alternative with the lowest treatment acquisition cost is likely to be the optimal treatment strategy.
What additional evidence or analyses might help to resolve this key issue?	There is nothing the company can do to address the differences in the study populations. The ERG has provided several scenarios that may help to approximate the likely impact on the ICER in different subgroups. The ERG accepts that the company wish to seek a recommendation for relugolix CT for women who wish to avoid or delay surgery, but the ERG would welcome further consultation with a range of clinical experts to help determine whether relugolix CT would also be used as a 'pre-surgery' treatment in clinical practice.

1.4 The clinical effectiveness evidence: summary of the ERG's key issues

The CS does not provide a full account of the clinical effectiveness evidence. The company present the results of ITCs of relugolix CT versus UPA and UPA versus leuprolide acetate GnRH agonist but not of relugolix CT versus GnRH agonist. Results were only presented for the mean difference in percentage change from baseline in MBL but not for other relevant clinical outcomes (see sections 3.4 and 3.5 of this report).

Issue 2 Lack of formal comparison between relugolix CT and GnRH agonists

Report section	Section 3.4, 3.5, 4.2.6 and 5.1
Description of issue and why the ERG has identified it as important	Lack of formal comparison between relugolix CT and GnRH agonists. The company present the results of an ITC of relugolix CT versus UPA and UPA versus leuprorelin acetate but not of relugolix CT versus GnRHa. Results were only presented for MBL volume despite several other outcomes were listed in their scope. Furthermore, uncertainty surrounding the treatment effect was no reported or included in the economic model.
What alternative approach has the ERG suggested?	An NMA would have been the most appropriate method for addressing this issue. The ERG has attempted to illustrate the impact of incorporating uncertainty surrounding the treatment effect by re-creating the ITC and approximating standard errors for the comparison of relugolix CT versus BSC for inclusion in the probabilistic analysis of the economic model.
What is the expected effect on the cost-effectiveness estimates?	There is unlikely to be any direct impact on the deterministic ICER as the ERG has been able to back calculate the MBL data used in the model from the ERG's reproduction of the company's ITC for MBL. However, uncertainty surrounding point estimates of MBL treatment effect for relugolix CT vs. GnRH agonists and versus. BSC (from the LIBERTY trials) were not incorporated into the economic model's probabilistic analysis. Therefore, the company's model substantially under-estimates the uncertainty surrounding the company's preferred base case ICER.
What additional evidence or analyses might help to resolve this key issue?	A more complete presentation of the evidence from the company, including an NMA, that estimates a treatment effect and standard error for MBL should be incorporated into the economic model. A pooled estimate of MBL effect for relugolix CT compared to BSC from the LIBERTY study should also be provided and fully incorporated into the model probabilistic analysis. Given that the company have access to the relevant trials data, it would be preferable if they provided a complete set of ITC results (and standard errors) for inclusion within the probabilistic analyses.

1.5 The cost-effectiveness evidence: summary of the ERG's key issues

The ERG raises several issues surrounding the appropriateness of the company's base case model structure (the choice of treatment rather than health states and the appropriateness of assuming listing for surgery can only take place after discontinuation, with a further waiting time of 15 months), assumptions about treatment discontinuation in clinical practice, uncertainty surrounding clinical effectiveness parameters used in the model (i.e., MBL), resource use assumptions for routine follow up, and the utility function used to estimate the impact of MBL on utilities mapped from UFS-QoL to EQ-5D. These issues would benefit from the company providing further data from their studies where possible as well as broader engagement with clinical experts around the use of relugolix CT in UK clinical practice and the associated role of surgery within the treatment pathway.

Issue 3 The appropriateness of using "treatment" rather than "health" states in the economic model structure.

Report section	Section 4.2.2
Description of issue and why the ERG has identified it as important	The model structure is built around 'treatment' states (relugolix CT / GnRH agonist and best supportive care) to reflect the treatment pathway. The ERG would have preferred a model built around "health" states defined according to symptom control because it would a) allow the model clinical effectiveness inputs to more closely reflect the trial data (i.e., avoiding the application of MBL data from the trial's mITT analysis directly to an 'on treatment' cohort) and b) allow routine monitoring to reflect patient health / symptom control rather than treatment received and thus would be more reflective of patient management in UK clinical practice. This is potentially an important driver of the ICER, but further modelling would be required to determine the impact.
What alternative approach has the ERG suggested?	The ERG believes adopting a model structure defined according to 'health' rather than 'treatment' states would generate a more accurate estimate of the ICER and would more appropriately reflect decision-making in UK clinical practice.
What is the expected effect on the cost-effectiveness estimates?	The direction and magnitude of any biases are unclear, but it is likely that MBL data used in the company base case analysis, based on an intention to treat analysis of the LIBERTY trial data, would over-estimate the MBL in an on-treatment cohort. However, the cost savings of avoiding BSC may be overestimated in the company's model. The net impact is unclear, and it could bias in favour or against relugolix CT.
What additional evidence or analyses might help to resolve this key issue?	The ERG would ideally like to see a model structured around 'health' rather than 'treatment' received states but appreciates this would be a significant undertaking. If this is not possible, an alternative, second-best option would be for the company to provide a more accurate estimate of the MBL in an 'on treatment' cohort from both LIBERTY and PEARL II studies to help determine the likely magnitude of any bias.

Issue 4 The most appropriate assumptions about treatment discontinuation in UK clinical practice for both relugolix CT and GnRH agonists

Report section	Section 4.2.6
Description of issue and why the ERG has identified it as important	The company preferred base case analysis has modified relugolix CT treatment discontinuation data from the LIBERTY study (up to 24 months of follow-up data available), based on clinical expert opinion and subjective judgment about whether study withdrawals would have continued treatment in clinical practice. No such adjustments were made to GnRH agonist discontinuation up to 3 months, sourced from the PEARL II study. This is important because it impacts on treatment acquisition costs, the costs of follow-on treatment (BSC / surgery), and the duration with which the cohort receives the benefits of relugolix CT. Therefore, it has an important impact on the ICER.
What alternative approach has the ERG suggested?	The ERG prefers the application of treatment discontinuation rates from the trial to ensure that the costs incurred are consistent with the use of relugolix CT that was required to deliver the modelled treatment benefit.
What is the expected effect on the cost-effectiveness estimates?	The implication of applying unmodified discontinuation rates is to increase the discontinuation rate for relugolix CT relative to the company's base case ICER following clarification, reducing treatment acquisition costs, and increasing the proportion receiving BSC or surgery. The magnitude of the impact on the ICER, therefore, depends on the most appropriate assumptions about other modelling parameters (e.g., resource use incurred in BSC and utilities).
What additional evidence or analyses might help to resolve this key issue?	The ERG is satisfied that the company has provided all the necessary evidence on which to make an informed judgment about the most appropriate treatment discontinuation data to apply in the model.

Issue 5 The appropriateness of a 'waiting time' health state post-treatment discontinuation

Report section	Section 4.2.3 and 4.2.6
Description of issue and why the ERG has identified it as important	Transition to the 'surgery' health states is conditional on A) having discontinued medical treatment prior to being listed for surgery and B) having transitioned through a 15-month waiting time state where no active treatment is provided. The ERG's clinical expert advice is that, in clinical practice, patients remain on their primary treatment whilst waiting for a scheduled surgery to ensure maximum fibroid shrinkage to improve chances of surgical success. Different assumptions about the inclusion/removal of the waiting time state and its duration if included lead to substantial variation in the ICER.
What alternative approach has the ERG suggested?	The ERG prefers the removal of the waiting time state to better reflect the use of treatment in UK clinical practice.
What is the expected effect on the cost-effectiveness estimates?	Removal of the 'waiting time' state, therefore, leads to a substantial increase in the ICER.
What additional evidence or analyses might help to resolve this key issue?	Further clinical expert advice from a range of clinicians experienced in treating fibroids to confirm whether patients would usually remain on treatment up until they receive surgery. Further validation of the assumption that surgery would not be scheduled past the age of 46 would also be useful.

Issue 6 The role of surgery in the treatment pathway and the lack of data to inform transitions to the surgery health state

Report section	Section 4.2.2 and 4.2.6
Description of issue and why the ERG has identified it as important	The ERG considers that some effect on surgery rates may be plausible because of longer treatment duration with relugolix CT compared to GnRHa, but the magnitude of reduction in surgery rates is not evidence-based, and highly uncertain given the data presented. Issues include: • Patient preference plays an important role in the decision to have surgery • Surgery rates were not collected in the LIBERTY studies • Transitions to surgery informed by the PEARL II study where all patients were considered for surgery are unlikely to be generalisable to a cohort of women who are unable or do not wish to have surgery (see Issue 1). The role of surgery in the treatment pathway, and the rates of transition to surgery are important drivers of cost-effectiveness.
What alternative approach has the ERG suggested?	The company has provided sensitivity analyses removing surgery and the ERG conducts further exploratory sub-group analyses in patients: A: who don't wish to or cannot have surgery B: who wish to receive treatment in preparation for surgery
What is the expected effect on the cost-effectiveness estimates?	Group A: removal of surgery states favours GnRHa, increasing the ICER substantially Group B: equalising treatment effectiveness favours relugolix CT
What additional evidence or analyses might help to resolve this key issue?	A comprehensive review of the literature to identify the rates of surgery that might be expected in a patient population similar to that included in the LIBERTY study would help resolve some uncertainty about the likely transitions to surgery following longer-term use of medical treatment. A more complete ITC, particularly around uterine or fibroid volume, would help validate the ERG assumption of equal effectiveness between relugolix CT and GnRH agonists as a treatment in preparation for surgery used for the exploratory subgroup analysis.

Issue 7 Uncertainty surrounding the utility function

Report section	Section 4.2.7
Description of issue and why the ERG has identified it as important	The company have mapped disease-specific quality of life data from the UFS-QoL, collected in the LIBERTY studies, to EQ-5D using an algorithm from a previous assessment. An OLS linear regression model, adjusting for age and MBL, is then used to predict the impact of MBL on mapped EQ-5D utilities to generate time varying utilities while on treatment or BSC. The company have not provided any details about what alternative model specifications were explored, or why the chosen model was used. This is an issue because the ICER is sensitive to changes in the co-efficient on MBL obtained from the utility function.
What alternative approach has the ERG suggested?	The ERG requested and was provided with the results of a repeated measures model at the clarification stage, and a corrected clarification response post FAC, where the co-efficient on MBL was somewhat higher than in the original OLS model. However, the most appropriate specification for the utility function remains unclear. In the absence of a full exploration of the advantages and disadvantages of different approaches, the ERG prefers the repeated measures model because it allows more appropriate exploration of uncertainty and generates utilities closer to general population averages when MBL is low.
What is the expected effect on the cost-effectiveness estimates?	The repeated measures model, with corrected reporting post FAC, generates a slightly higher reduction in utility for every unit increase in MBL compared to the company preferred OLS model. The implication is lower QALYs in both arms of the model, higher incremental QALY gains for relugolix CT and hence a slightly lower ICER compared to the company preferred base case model.
What additional evidence or analyses might help to resolve this key issue?	A complete assessment of the relative advantages and disadvantages of alternative utility functions, including, for example, exploration of squared terms to explore non-linearities in the impact of MBL on utility, discussion of the face validity and model fits of alternative utility functions.

Issue 8 Monitoring and follow up resource use in UK clinical practice

Report section	Section 4.2.8
Description of issue and why the ERG has identified it as important	The ERG considers the company's base case routine monitoring and resource use to be an over-estimate of UK clinical practice. In addition to dexa-scans to monitor BMD, the company assumes all patients would have annual scans (ultrasound [100%], MRI [25%], hysteroscopy [25%]) whether on or off treatment. A six-monthly gynaecologist consultation was assumed for those on treatment, but not for those on BSC. This is important because the frequency of scanning and consultations leads to important changes to the ICER, particularly when these differ between the on and off treatment cohorts.
What alternative approach has the ERG suggested?	The ERG's clinical expert considers the company's use of dexa-scans to be appropriate, but the remaining examinations and consultations to be an over-estimate. The ERG considers a one-off consultation with a gynaecologist and a scan to assess progress and make a long-term treatment plan would be more appropriate and would be applied whether on treatment or after discontinuation (i.e., upon entry to the BSC state).
What is the expected effect on the cost-effectiveness estimates?	The ERG preferred resource use reduces total costs in both arms of the model, and also reduces the incremental costs associated with relugolix CT, by removing the additional six-monthly gynaecologist consultation compared to BSC. The impact is a reduction in incremental costs and a reduction in the ICER compared to the company's preferred base case.
What additional evidence or analyses might help to resolve this key issue?	The ERG would consider it more appropriate to link resource usage to symptom control rather than on/off treatment and believe this could be incorporated into a model defined by 'health' states (See Issue 3). The ERG is of the opinion that further engagement with a wide range of clinical experts would help to better understand the heterogeneity in how frequently patients have contact with hospital services in UK clinical practice.

1.6 Summary of ERG's preferred assumptions and resulting ICER

Table 2 below outlines the ERG's preferred modelling assumptions. The table demonstrates the impact of changing each assumption from the company's base case individually. There are several uncertainties that the ERG has not been able to resolve at this stage and the ERG's preferred ICER may therefore change following technical engagement if further evidence is provided by the company. The ERG notes that there are many uncertainties surrounding modelling assumptions, and limited data to inform the model. Several assumptions are associated with advantages and disadvantages. Whilst the ERG provides some suggested alternative assumptions, it may be more appropriate to consider a plausible range of ICERs that more appropriately reflect the uncertainty in the underlying assumptions. The magnitude of uncertainty is more appropriately captured using the ERG's revised probabilistic analyses.

Given that the ERG agrees with the company's assumption that all GnRH agonists have equal effectiveness, the cheapest GnRHa (goserelin monthly) dominates all other GnRH agonists at current list prices. For simplicity of reporting, ICERs are only reported for relugolix CT versus goserelin monthly.

Table 2 Summary of ERG's preferred assumptions and ICER (relugolix CT vs. Goserelin monthly)

Scenario	Incremental cost (£)	Incremental QALYs	ICER (£ / QALY)
Company's base case, submitted following clarification	2112	0.364	5,796
+ Application of unmodified treatment discontinuation rates from the LIBERTY study (Issue 4)	444	0.103	4,311
+ Removal of waiting time state for surgery (Issue 5)	407	0.046	8,784
+ Utilities sourced from a repeated measures model (Issue 7)	407	0.07	5,846
+ Female specific UK general population utility norms	407	0.069	5,866
+ Resource use adapted to reflect UK clinical practice (Issue 8)	194	0.069	2,795
ERG's suggested base case deterministic	194	0.069	2,795
ERG's suggested base case probabilistic (including Issue 2)	197	0.069	2,833

Further details of the ERG's additional exploratory and sensitivity analyses, including a full set of updates to the probabilistic analyses can be found in Chapter 6.

2 INTRODUCTION AND BACKGROUND

2.1 Introduction

The relevant health condition for the submission received from Gedeon Richter Limited is moderate to severe symptoms associated with uterine fibroids (UF). The company's description of this health condition in terms of prevalence and complications appears generally accurate and in line with the decision problem. However, the company's focus on heavy menstrual bleeding as "moderate to severe symptoms" is questioned by the ERG's clinical expert, who is of the opinion that pressure symptoms are relevant in this context and should have been specified in the company's inclusion criteria. The relevant intervention for this submission is relugolix in combination with oestradiol and norethisterone (Ryeqo®, Gedeon Richter Limited).

2.2 Background

The company submission (CS) describes uterine fibroids (UF) as benign tumours that develop in or around the uterus. The majority of UF (correctly known as leiomyomas or myomas)[1] are asymptomatic but, for those people who do experience symptoms, treatment can be necessary. There are three distinct classes of symptoms: prolonged or heavy menstrual bleeding, pelvic pressure and pain, and reproductive dysfunction. Bleeding symptoms can be related to the location of the UF, with submucosal the most likely cause. Pelvic pressure is due to increase in the size of the uterus.[2] Other symptoms experienced by some people include abdominal pain, frequent need to urinate, constipation and pain or discomfort during sex.[3] Although the aetiology of UF is not currently known, their development has been linked to oestrogen;[3-5] accordingly, UF tend to develop in people aged between 16 and 50 years, when oestrogen levels are high and shrink after the menopause, when oestrogen levels drop.[3, 5]

Risk factors for UF include race (in particular, black women are disproportionately affected, with UF being three times more common in black women than white women, and more severe symptoms in black women), age, obesity (which increase the risk of UF due to the metabolic function of adipose tissues), having never been pregnant (with each subsequent child possibly lowering the risk further in multiparous women), hypertension, and vitamin D deficiency and diet.[5-7]

Uterine fibroids may be discovered during routine gynaecological examinations, otherwise, diagnosis is usually by tests such as ultrasound scan, hysteroscopy, or laparoscopy.[8]

Uterine fibroids are the most common neoplasms in women worldwide[6] but their actual incidence is difficult to estimate because they are often asymptomatic.[4, 9] Hospital Episode Statistics for the year 2020-21 in England report a total of 15,646 finished consultant episodes for leiomyoma of the uterus (codes D25.0: Submucous leiomyoma of uterus, D25.1: Intramural leiomyoma of uterus, D25.2: Subserosal leiomyoma of uterus, D25.9: Leiomyoma of uterus, unspecified).[10]

The CS cites the NICE pathway for managing heavy menstrual bleeding as the most relevant clinical pathway (presented in Document B, Figure 3 of the CS and reproduced as Figure 1 below).[11]

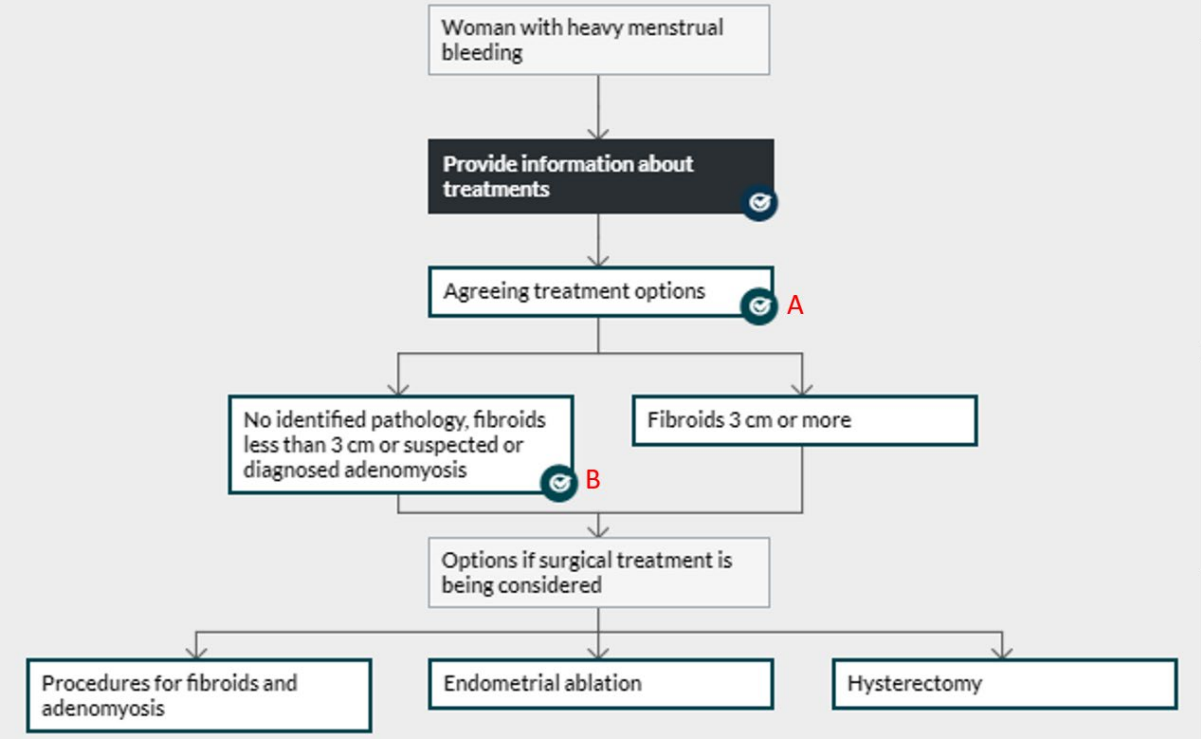


Figure 1 NICE pathway for managing heavy menstrual bleeding [reproduced from Figure 3, Document B of the CS]

The NICE pathway makes the following recommendations:

No identified pathology, fibroids <3cm in diameter which are not causing any distortion of the uterine cavity or suspected or diagnosed adenomyosis.

The NICE pathway recommends levonorgestrel-releasing intrauterine system (LNG-IUS) as the first treatment for heavy menstrual bleeding (HMB) in these women.
For women who decline LNG-IUS, or for whom it is not suitable, pharmacological treatments should be considered:

- Non-hormonal:
    - Tranexamic acid
    - Non-steroidal anti-inflammatory drugs (NSAID)
- Hormonal:
    - Combined hormonal contraception
    - Cyclical oral progestogens.

If treatment is unsuccessful, pharmacological treatment is declined or symptoms are severe; referral to specialist care should be considered:

- Investigations to diagnose the cause of HMB, if needed, taking account of any investigations already undergone and
- Alternative treatment choices, including:
    - Pharmacological options not already tried
    - Surgical options:
        - Second generation endometrial ablation
        - Hysterectomy.

For women with submucosal fibroids, hysteroscopic removal should be considered.

Fibroids 3cm or more in diameter

Taking into account the size, location, and number of fibroids, and severity of symptoms, the following treatments should be considered:

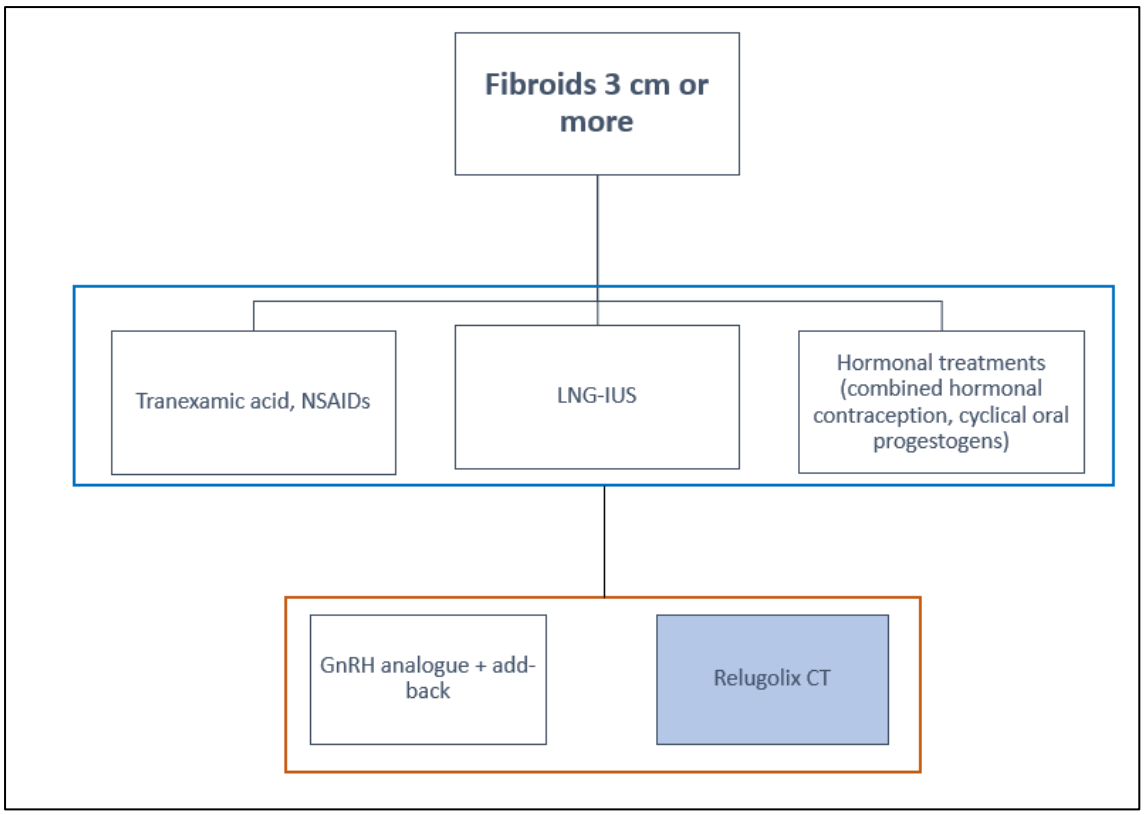
- Non-hormonal (tranexamic acid and/or NSAIDs) should be offered whilst investigations and definitive treatment are being organised; use of these treatments should be continued for as long as they are found to be beneficial
- Hormonal treatment (LNG-IUS, combined hormonal contraception, cyclical oral progestogens, ulipristal acetate [UPA])
- Uterine artery embolisation (UAE)

- Surgical: myomectomy or hysterectomy
- UPA should only be considered for the intermittent treatment of moderate to severe symptoms of UF in premenopausal women if surgery and UAE are not suitable, declined, or have failed surgery or UAE
- Second-generation endometrial ablation should be considered for those who meet the criteria
- Pre-treatment with gonadotropin-releasing hormone (GnRH) analogues before hysterectomy and myomectomy should be considered if UF are causing an enlarged or distorted uterus.

The CS states that there is a current unmet need for pharmacological treatments for moderate to severe UF due to a lack of satisfactory medical treatments. The CS further states that there is no other treatment currently available that meets the unmet need and with an indication that is not time restricted in premenopausal women with moderate to severe UF. The ERG's clinical expert agrees with the company's position.

The CS provides a description of the relevant intervention for this appraisal, relugolix CT (relugolix in combination with oestradiol and norethisterone acetate) in Document B, Table 2 of the CS. Then company describes relugolix as a non-peptide GnRH antagonist that binds to, and inhibits, GnRH receptors in the anterior pituitary gland. Such inhibition results in a dose-dependent decrease in the release of luteinizing hormone and follicle stimulating hormone. By reducing their circulating concentrations, follicular growth and development are prevented and ovulation and development of the corpus luteum are prevented, resulting in reduction of oestrogen production and progesterone, respectively. Relugolix CT was granted marketing authorisation from the EMA on 16th July 2021 and from the MHRA on 9th August 2021.

The proposed place of relugolix CT in the treatment pathway is presented in Document A, Figure 1 of the CS, and is reproduced below as Figure 2. The ERG agrees that the company's proposed pathway is representative of current clinical practice and the anticipated positioning of relugolix CT is within its licensed indication.



NSAIDs; non-steroidal anti-inflammatory drugs, LNG-IUS; levonorgestrel-releasing intrauterine system, GnRH; gonadotrophin-releasing hormone

Figure 2 Company's proposed treatment pathway and positioning of relugolix CT for treating uterine fibroids [reproduced from Figure 1, Document A of the CS]

2.3 Critique of company's definition of the decision problem

A summary of the company's decision problem in relation to the NICE final scope is presented in Table 3 below. A critique of adherence of the company's economic modelling to the NICE reference case is presented in Chapter 4.

Table 3 Summary of the company's decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
Population	People with moderate to severe symptoms associated with uterine fibroid(s) (UF)	Same as scope	N/A	The ERG agrees that the population included in the LIBERTY trials is appropriate for this appraisal. The ERG notes, however, that the patient population of the LIBERTY trials does not match that of the PEARL trials, which were used for the ITC with UPA. In the PEARL trials, all women had surgery planned after 13 weeks while planned surgery was an exclusion criterion for the LIBERTY trials. The ERG believes that it may be relevant to consider the clinical and cost-effectiveness of relugolix CT within two different settings to reflect the differing treatment goals: a) in women who wish to improve symptoms but do not intend to undergo surgery, and b) in women who have already been listed for surgery (see Section 4.2.2 for a critique of the populations used in the economic modelling). The ERG notes that symptoms associated with UF include both menstrual- and pressure-related. At

				clarification, the company reiterated that HMB is the most common symptom of UF and most people with "moderate symptoms" will have heavy bleeding. The company further stated that HMB is one of the only symptoms which can be assessed in an objective and quantifiable way and is the most accurate indicator of severity, with other symptoms being supplementary and supportive. Nonetheless, the ERG's clinical expert is of the opinion that some participants with moderate to severe symptoms, in particular, pressure symptoms, may have been excluded from this population.
Intervention	Relugolix with oestradiol and norethisterone acetate (also known as norethindrone acetate), alone, or as an add on to non-hormonal pharmacological treatments *[Please note that relugolix in combination with oestradiol and norethisterone acetate is referred to as 'relugolix CT' throughout this submission; 'CT' is the abbreviation for 'combination therapy']*	Same as scope	N/A	The ERG questioned the fixed 1 mg dosage of oestrogen in the relugolix CT as titrating the dose of oestrogen to gain vasomotor symptom relief for individual patients is current clinical practice. At clarification, the company explained that the dosages of relugolix 40 mg, oestradiol 1mg and norethisterone 0.5 mg were selected to achieve a balance of reproductive hormones to treat the UF symptoms whilst maintaining bone health, minimising vasomotor symptoms and protecting the endometrium from the

				effects of unopposed oestrogen. The company further stated that the combined doses of 40 mg relugolix and 1mg oestradiol achieves systemic oestradiol concentrations of 10 to <60 pg/ml which was sufficient to prevent hypoestrogenic symptoms and maintain bone health in most people. In addition, combining 1mg oestradiol and 0.5 mg norethisterone ensures oestradiol levels are within the pre-follicular phase level of 20 to 50 pg/mL, providing control of UF symptoms whilst minimising side effects. The ERG's clinical expert is of the opinion that the 1 mg dose of oestradiol in the company's combined therapy would be effective in addressing the osteoporosis side effects. However, using a static dose control, vasomotor symptoms is not considered reasonable as people metabolise at different rates and, in current clinical practice, the oestrogen dose is varied to the level required to control symptoms. The dose of 1mg of oestrogen is that which protects against BMD loss and will help control vasomotor symptoms in some, but not all, users.

Comparator(s)	Hormonal treatments, including: • levonorgestrel-releasing intrauterine system (LNG-IUS; off-label for some LNG-IUS) • combined hormonal contraception (off-label for some combined hormonal contraceptives) • cyclical oral progestogens gonadotrophin-releasing hormone analogues (off-label for some gonadotrophin-releasing hormone analogues)	The submission will focus on gonadotrophin-releasing hormone (GnRH) agonists as the relevant comparator for relugolix CT.	N/A	The ERG's clinical expert notes that GnRH antagonists, as opposed to GnRH agonists, would be relevant in this context. At clarification, the company stated that GnRH agonists were the most relevant comparators for relugolix CT, and that these are the existing treatment options that are expected to be displaced by relugolix CT in the NICE pathway for managing HMB. The company further reported that four GnRH antagonists were identified in its systematic literature review (relugolix, elagolix, linzagolix and cetrorelix) and provided justification for the latter three antagonists as not being relevant comparators for this appraisal. The ERG agrees that it is justifiable to exclude these treatments. The ERG's clinical expert also questions the omission of Esmya as a comparator, given that it is an oral preparation that targets symptoms and causes fibroid shrinkage. At clarification, the company stated that Esmya's indication has become limited due to safety concerns about liver injuries and is currently only indicated for intermittent treatment in

				this population when UF embolisation and/or surgery are not suitable or have failed. The company further stated that use of Esmya is currently low, a fact which supports GnRH as the most relevant comparators in this appraisal. The ERG agrees with the company's position in that it is unlikely that many people with UF requiring treatment would agree to randomisation to Esmya, given the level of monitoring required and potential risks of liver damage.
Outcomes	The outcome measures to be considered include: • change in menstrual blood loss (MBL) volume • time to MBL response • pain • uterine fibroid volume (UFV) / uterine volume (UV) • haemoglobin levels • change in bone mineral density (BMD) • rates and route of surgery • impact on fertility and pregnancy and teratogenic effects • mortality • adverse effects of treatment, including but not limited to	The outcome measures in the clinical effectiveness section include: • change in MBL volume • time to MBL response • pain • UFV/UV • haemoglobin levels • adverse effects of treatment, including but not limited to vasomotor symptoms, incontinence and pelvic organ prolapse • health-related quality of life. The outcome measures in the cost-effectiveness model include:	The following measures are not included in the clinical effectiveness section as they were not collected in the relugolix CT clinical trials: • rates and route of surgery • impact on fertility and pregnancy and teratogenic effects Rates and route of surgery are, however, included in the economic model. Mortality is not included as no deaths were reported during the relugolix CT clinical trials.	The ERG's clinical expert considers the outcomes reported in the CS to be appropriate for addressing the topic of this appraisal. However, for the ITCs the company provides only results for the mean difference in percentage from baseline in MBL and not results for other relevant outcomes. Moreover, they failed to consider the uncertainty of clinical effectiveness results within their economic model, appropriately. Despite the company's assertion that mortality is not included, the CS does indeed report that there were no deaths during the relevant trials.

	vasomotor symptoms, incontinence and pelvic organ prolapse • health-related quality of life	• MBL volume and change in MBL volume (used to derive utility) • Adverse effects • Quality of life	Whilst 'change in BMD' was explored in the relugolix CT clinical trials, it is not a relevant outcome in the economic model. In this submission, BMD is not an outcome in the economic model as it is assumed that BMD may resolve once treatment with GnRH agonist therapy (the comparator for relugolix CT) ceases and thus there may be no additional benefit to favour relugolix CT on this outcome. Despite this assumption, and as stated in section B.2.13, there is evidence to suggest that BMD may not be fully recoverable from GnRH agonist use which may underestimate the potential benefit that relugolix CT would provide to women with UF.	
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating	Same as scope	N/A	A critique of the company's economic analyses against the NICE reference case is provided in Section 4.2.1.

	clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective. The availability of any commercial arrangements for the intervention, comparator and subsequent treatment technologies will be taken into account.			
Subgroups	Not specified	Not specified	Not specified	At clarification, the company stated that subgroup analyses of the primary efficacy endpoint were conducted for LIBERTY 1, LIBERTY 2 and LIBERTY 3. The company noted that no subgroups were used in the economic analyses and provided the results of all the analyses at clarification. The ERG is satisfied with the company's response regarding the LIBERTY studies.
Special considerations including issues related to equity or equality		Black African and African-Caribbean origin, who are 2-3 times more likely to develop UF than white women, may be more opposed to surgery due to cultural and religious beliefs.		The ERG's clinical expert is in agreement with the company's position

		Additionally, some women will choose to decline surgery in order to avoid impacting their personal circumstances with respect to work and family commitments such as childcare, etc.		

3 CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

Full details of the methods used to identify and select the clinical evidence relevant to this appraisal are reported in Appendix D of the CS. The ERG'S appraisal of the company's systematic review methods is summarised in Table 4.

Table 4 ERG's appraisal of the systematic review methods presented in the CS

Review process ERG	ERG response	Comments
Were appropriate searches (e.g., search terms, search dates) performed to identify all relevant clinical and safety studies?	Yes	The CS provides full details of the searches used to identify the studies for the clinical effectiveness review. The search strategies include relevant controlled vocabulary and text terms with appropriate use of Boolean operators and are fully reproducible. Details provided in Appendix D of the CS.
Were appropriate bibliographic databases/sources searched?	Yes	Sources included Embase, Medline, and CENTRAL for primary research, DARE and CDSR for evidence syntheses. Relevant conference proceedings were also searched. Full details are provided in Appendix D of the CS.
Were eligibility criteria consistent with the decision problem outlined in the NICE final scope?	Yes	Although the submission focused on GnRH agonists as the comparator, the searches for clinical evidence included all therapeutic options so all relevant results will have been found.
Was study selection conducted by two or more reviewers independently?	Yes	Appendix D1.1, page 234: "*Abstracts and titles were reviewed by two independent reviewers in a double-blind process against the inclusion to identify potentially relevant studies*"
Was data extraction conducted by two or more reviewers independently?	Yes	At clarification: "*Two reviewers were involved in data extraction for both the initial and update SLRs and worked independently*"
Were appropriate criteria used to assess the risk of bias of identified studies?	Yes (for the RCTs)	Document B, page 63: "*LIBERTY 1, LIBERTY 2 and LIBERTY 3 were assessed for quality using*

		the York Centre for Reviews and Dissemination [CRD] guidance for undertaking reviews in healthcare". The ERG considers these criteria to be appropriate. LIBERTY 3 is not an RCT so the CRD criteria are mainly not applicable. The CRD criteria were also used for the assessment of PEARL I and PEARL II.
Was the risk of bias assessment conducted by two or more reviewers independently?	Yes	At clarification: "*Two reviewers conducted the risk of bias assessment. The reviewers worked independently then came together to discuss and agree the assessment findings"*
Was identified evidence synthesised using appropriate methods?	No	The ERG believes a network meta-analysis should have been used for the primary efficacy outcome and that a comparison of relugolix CT versus GnRHa presented for the secondary outcomes. Full details of the ITC for MBL should have been provided and the associated uncertainty incorporated into the economic model.

The ERG conducted a quality assessment of the methods used by the company for the systematic review of clinical evidence using the Centre for Review and Dissemination (CRD) criteria. The results are presented in Table 5.

Table 5 Quality assessment of the company's systematic review of clinical effectiveness evidence

CRD quality item	Yes/No/Unclear
1. Are any inclusion/exclusion criteria reported relating to the primary studies, which address the review question?	Yes
2. Is there evidence of a substantial effort to search for all of the relevant research?	Yes
3. Is the validity of included studies adequately assessed?	Yes
4. Are sufficient details of the individual studies presented?	Yes
5. Are the primary studies summarised appropriately?	Yes

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

3.2.1 Included studies

Details of the key clinical effectiveness evidence are provided in Document B, Section B.2 of the CS. The company presents clinical effectiveness evidence from two phase-3, multicentre, international, double-blind RCTs, LIBERTY 1, LIBERTY 2, conducted between March 2017 and July 2019 and one phase 3 open-label extension study of the LIBERTY 1 and LIBERTY 2 trials, LIBERTY 3. The methods of the three trials are summarised in Document B, Table 6 of the CS and reproduced in Table 6 below. Details of LIBERTY 1, LIBERTY 2 and LIBERTY 3 are reported in sections B.2.2 and B.2.3 of the CS and the participant flow of the studies are presented in Appendix D.1.2. LIBERTY 1 was conducted at 80 sites (USA, Brazil, Italy, Poland, South Africa and the UK) and LIBERTY 2 at 99 sites (USA, Belgium, Brazil, Chile, Czech Republic, Hungary, Poland and South Africa). All three LIBERTY studies were funded by Myovant Sciences. The objective of LIBERTY 1 and LIBERTY 2 was to assess the effectiveness of relugolix combination therapy (CT) compared with placebo for 24 weeks and the methods used in the two trials were identical. Participants were randomly assigned in a 1:1:1 ratio to receive either:

- relugolix CT for 24 weeks: 40 mg relugolix in combination with 1 mg oestradiol and 0.5 mg norethisterone or
- delayed relugolix CT: 40 mg relugolix monotherapy for 12 weeks followed by relugolix CT (as above) for 12 weeks or
- placebo for 24 weeks.

The purpose of the relugolix delayed arm was to allow for the comparison of BMD and vasomotor symptoms in the combination and monotherapy arms during the first 12 weeks of the trial. This arm is not further considered in the clinical effectiveness evidence for this appraisal, which will focus on the relugolix CT versus placebo comparison, as per the licensed indication for relugolix CT. The key eligibility criteria for LIBERTY 1 and LIBERTY 2 are reported in Document B, Table 7 and the eligibility criteria in full are presented in Appendix M1.1, Table 116 of the CS. The study population in LIBERTY 1 and LIBERTY 2 was premenopausal women aged 18

to 50 years with HMB associated with UF (≥80mL per cycle for two cycles or ≥160mL for one cycle as measured by the alkaline haematin [AH] method during the screening period). People who were expected to undergo gynaecological surgery or ablation procedures for UF within 6 months of enrolment into the study were excluded.

LIBERTY 3 is a 28-week open-label extension to LIBERTY 1 and LIBERTY 2. Eligible participants were those who completed LIBERTY 1 or LIBERTY 2 and all received open-label relugolix CT.

Table 6 Comparative summary of the methodology of the relugolix CT studies [reproduced from Table 6, Document B of the CS]

Trial number (acronym)	MVT-601-3001 (LIBERTY 1)	MVT-601-3002 (LIBERTY 2)	MVT-601-3003 (LIBERTY 3)
Location	80 centres globally, including centres in the USA, Brazil, Italy, Poland, South Africa and the UK. Approximately 25% of patients were enrolled at sites outside of North America.	99 centres globally, including centres in the USA, Belgium, Brazil, Chile, Czech Republic, Hungary, Poland and South Africa. Approximately 25% of patients were enrolled at sites outside of North America.	149 centres in the USA, Belgium, Brazil, Chile, Czech Republic, Hungary, Italy, Poland and South Africa.
Trial design	Phase-3 randomised, double-blind, placebo-controlled trial		Phase-3, open-label, single-arm, long-term efficacy and safety extension study
Eligibility criteria for participants	Premenopausal women 18 to 50 years of age with regularly occurring menstrual periods of <14 days' duration with cycle of 21 to 38 days; who had a diagnosis of fibroids as confirmed on ultrasonography and who had HMB, as assessed by the AH method, were eligible		Women who completed 24 weeks of study drug treatment and study participation in either LIBERTY 1 or LIBERTY 2. They were not expected to undergo gynaecological surgery or ablation procedures for UF within the study period, including during the Safety Follow-up period. Negative urine pregnancy test at Week 24/Baseline visit.
Trial drugs	Participants were randomly assigned, in a 1:1:1 ratio, by means of an interactive website to receive blinded placebo for 24 weeks, relugolix CT (40 mg relugolix in combination with 1 mg oestradiol and 0.5 mg norethisterone acetate) for 24 weeks, or relugolix-delayed CT (relugolix monotherapy followed by relugolix CT, each for 12 weeks).* • LIBERTY 1: 388 randomised: relugolix CT (128), placebo (128), relugolix-delayed CT (132) • LIBERTY 2: 382 randomised: relugolix CT (126), placebo (129), relugolix-delayed CT (127)		477 women enrolled to receive open-label relugolix CT (40 mg relugolix in combination with 1 mg oestradiol and 0.5 mg norethisterone acetate) for 28 weeks. This comprised >75% of patients who completed one of the parent studies (LIBERTY 1 or LIBERTY 2).

	Trial visits occurred at baseline and every 4 weeks for 24 weeks.		
Primary outcomes	The proportion of women 'responding' in the relugolix CT versus the placebo group where a 'responder' was classified as a woman who achieved an MBL volume of < 80 mL and at least a 50% reduction from baseline MBL volume over the last 35 days of treatment, as measured by the AH method.		The proportion of women who achieved or maintained an MBL volume of < 80 mL and at least a 50% reduction from parent study baseline MBL volume to the last 35 days of treatment, as measured by the AH method
Other outcomes used in the economic model/specified in the scope	Outcomes in the model: • MBL volume and change in MBL volume (used to derive utility) • Adverse events • Quality of life Other outcomes in the scope: • Achievement of amenorrhoea • Uterine volume • Uterine fibroids volume • Pain (associated with uterine fibroids) • Change in haemoglobin		Outcomes in the model: • MBL volume and change in MBL volume (used to derive utility) • Quality of life Other outcomes in the scope: • Adverse events • Achievement of amenorrhoea • Uterine volume • Uterine fibroids volume • Pain (associated with uterine fibroids)
Pre-planned subgroups	N/A	N/A	N/A

* The relugolix-delayed CT group was included to allow for the comparison of BMD and vasomotor symptoms in the combination and monotherapy groups during the first 12 weeks of the trial. This arm does not relate to the licenced indication for relugolix CT.

The company assessed the risk of bias of LIBERTY 1, 2 and 3 using an adapted version of the Centre for Reviews and Dissemination checklist for RCTs and concluded that the LIBERTY 1 and LIBERTY 2 RCTs and the LIBERTY 3 open-label extension study were of good quality.[12] In general, the ERG agrees with the findings of the company's assessment.

Details of the baseline characteristics of the modified ITT (mITT) populations of LIBERTY 1 and LIBERTY 2 (i.e. randomised participants who received any amount of the study drug; efficacy analyses were performed by treatment group as randomised) and the safety population of LIBERTY 3 (i.e. participants who received any amount of open-label study drug; safety data were analysed by parent study treatment group by actual treatment received) are presented in Document B Tables 9, 10, 11 and 14 of the CS and summarised in Table 7 below.

Table 7 Baseline characteristics of participants in the modified ITT populations of LIBERTY 1 and LIBERTY 2 and the safety population of LIBERTY 3 [adapted from Tables 9, 10, 11, and 14, Document B of the CS, and the LIBERTY 3 CSR]

	LIBERTY 1[a]			LIBERTY 2[a]			LIBERTY 3[b]		
	Relugolix CT (n=128)	**Relugolix-delayed CT (n=132)**	**Placebo (n=127)**	**Relugolix CT (n=125)**	**Relugolix-delayed CT (n=127)**	**Placebo (n=129)**	**Relugolix CT (n=163)**	**Relugolix-delayed CT (n=149)**	**Placebo (n=164)**
Age, years, mean (SD)	42.5 (5.0)	41.3 (5.4)	42.2 (5.7)	42.4 (5.4)	42.1 (5.3)	41.8 (5.3)	42.6 (5.1)	42.1 (5.6)	41.9 (5.4)
Race, n (%)									
White	64 (50.0)	53 (40.2)	56 (44.1)	58 (46.4)	50 (39.4)	49 (38.0)	85 (52.1)	51 (34.2)	71 (43.3)
Black or African American	59 (46.1)	67 (50.8)	65 (51.2)	62 (49.6)	66 (52.0)	74 (57.4)	69 (42.3)	81 (54.4)	88 (53.7)
Other	5 (3.9)	12 (9.1)	6 (4.7)	2 (1.6)	8 (6.3)	5 (3.9)	6 (3.7)	15 (10.1)	4 (2.4)
Not reported	0 (0)	0 (0)	(0)	3 (2.4)	3 (2.4)	1 (<1%)	3 (1.8)	2 (1.3)	1 (<1%)
BMI, kg/m^2, mean (SD)	31.4 (7.6)	31.4 (7.3)	32.3 (7.5)	31.0 (6.6)	30.8 (5.7)	32.1 (7.6)	31.4 (7.0)	31.0 (6.4)	32.6 (7.5)
MBL volume, mL, mean (SD)	239.4 (180.3)	228.9 (159.6)	218.8 (125.0)	246.7 (186.0)	227.4 (134.4)	211.8 (129.9)	248.7 (197.0)	238.8 (155.3)	216.0 (123.8)
Haemoglobin, g/dL, mean (SD)	11.2 (1.6)	11.1 (1.7)	11.4 (1.4)	11.3 (1.5)	11.1 (1.6)	11.1 (1.6)	11.4 (1.5)	11.0 (1.6)	11.2 (1.5)
Index UF volume, cm^3, mean (SD)	71.9 (128.1)	93.8 (143.8)	71.8 (124.0)	73.7 (126.7)	78.9 (157.5)	74.1 (123.0)	80.0 (145.1)	91.5 (137.8)	74.2 (128.1)

	LIBERTY 1[a]			LIBERTY 2[a]			LIBERTY 3[b]		
	Relugolix CT (n=128)	**Relugolix-delayed CT (n=132)**	**Placebo (n=127)**	**Relugolix CT (n=125)**	**Relugolix-delayed CT (n=127)**	**Placebo (n=129)**	**Relugolix CT (n=163)**	**Relugolix-delayed CT (n=149)**	**Placebo (n=164)**
Uterine volume, cm^3, mean (SD)	379.1 (316.8)	469.9 (427.9)	397.8 (324.9)	387.7 (344.0)	402.7 (371.1)	407.9 (402.0)	386.7 (320.5)	442.4 (370.9)	401.5 (351.5)
Surgery for UF									
Yes	20 (15.6)	15 (11.4)	13 (10.2)	11 (8.8)	15 (11.8)	11 (8.5)	21 (12.9)	14 (9.4)	17 (10.4)
No	108 (84.4)	117 (88.6)	114 (89.8)	114 (91.2)	112 (88.2)	118 (91.5)	142 (87.1)	135 (90.6)	147 (89.6)
UAE							NR	NR	NR
Yes	2 (1.6)	2 (1.5)	1 (0.8)	3 (2.4)	0 (0)	0 (0)			
No	126 (98.4)	130 (98.5)	126 (99.2)	122 (97.6)	127 (100)	129 (100)			
UFS-QoL (BPD subscale), mean (SD)	66.8 (22.1)	68.5 (22.9)	71.4 (21.3)	70.7 (20.8)	72.0 (22.9)	70.0 (20.3)	67.2 (21.0)	72.7 (19.0)	72.6 (19.7)

Note. [a]mITT population, [b]Safety population
CT: combined therapy, SD: standard deviation, BMI: body mass index, MBL: menstrual blood loss, UF: uterine fibroids, UAE: uterine artery embolisation

In general, baseline characteristics were balanced within and across LIBERTY 1 and LIBERTY 2. Mean age was 42 years in LIBERTY 1 and 42.1 in LIBERTY 2. The majority of participants were Black or African American in both trials. Mean BMI of all randomised groups was ≥30, indicating that participants were generally in the obese range. The ERG's clinical expert is of the opinion that this is not representative of women seen in clinical practice and that women of healthy weight are equally likely to have uterine fibroids (UF). Adipose tissue produces oestrogen and obese women have a greater proportion of adipose tissue than women of healthy weight, but UF treatments targeting oestrogen production tackle only the oestrogen produced by the ovaries and not that produced by adipose tissue. Therefore, in the population of LIBERTY 1 and LIBERTY 2, the effects of relugolix CT may have been attenuated due to the mean BMI of participants being in the obese range.

The CS states that the disease-specific characteristics of participants in LIBERTY 1 and LIBERTY 2 are consistent with the population relevant to this appraisal – in particular, mean menstrual blood loss (MBL) at baseline ranged from 211.8mL (LIBERTY 2, placebo arm) to 246.7mL (LIBERTY 2, relugolix CT arm). Overall, the ERG's clinical expert is satisfied that the disease-specific baseline characteristics of the participants in LIBERTY 1 and LIBERTY 2 are representative of women with UF seen in clinical practice in the UK.

3.2.2 Primary and secondary efficacy endpoints

The outcome measures listed in the NICE final scope for this appraisal were: change in MBL volume; time to MBL response; pain; uterine fibroid volume (UFV) / uterine volume (UV); haemoglobin levels; rates and route of surgery; impact on fertility and pregnancy and teratogenic effects; change in bone mineral density (BMD); mortality; adverse effects of treatment, including but not limited to vasomotor symptoms, incontinence, and pelvic organ prolapse; and health-related quality of life. Rates and route of surgery and the impact on fertility and pregnancy and teratogenic effects were not measured in the CS. Results for the primary and secondary endpoints assessed in the CS are presented below.

Primary endpoint: LIBERTY 1 and LIBERTY 2

The primary endpoint of LIBERTY 1 and LIBERTY 2 was achieving a response, defined as both a volume of MBL of less than 80 ml and a reduction of at least 50% from the baseline volume of MBL, as measured by the alkaline haematin (AH) method, over the last 35 days of the treatment period. In LIBERTY 1, the primary efficacy endpoint was achieved at a higher frequency in the relugolix CT group (94 participants, 73.4%) compared with the placebo group (24 patients, 18.9%), and the difference between the groups was statistically significant (54.5%, 95% CI 44.3% to 64.78%, $p < 0.0001$). Similarly, in LIBERTY 2, a greater proportion of participants in the relugolix CT group achieved the primary endpoint (89 participants, 71.2%) compared with the placebo group (19 participants, 14.7%) with a statistically significant difference between the groups (56.47%, 95% CI 46.45% to 66.49%, $p < 0.0001$).

Secondary endpoint: LIBERTY 1 and LIBERTY 2

The secondary efficacy endpoints reported in the CS are the following:

- MBL volume: Figure 8, Section B.2.6 of the CS, shows that the least-squares (LS) mean percent reduction from baseline to Week 24 in MBL volume was greater in the relugolix CT group than that in the placebo group in both LIBERTY 1 (-84.3% versus -23.2%) and LIBERTY 2 (-84.3% versus -15.1%) and the difference between the groups was statistically significant ($p < 0.001$ for both comparisons). Figure 9, Section B.2.6 of the CS, shows that a significant reduction in MBL volume occurred by Week 4, the first post-baseline assessment, and was sustained through Week 24.
- Time to MBL response (MBL volume < 80 mL and ≥ 50% reduction from baseline): The CSRs report that, based upon a Kaplan-Meier analysis, the median time to achieve a first response (the primary endpoint) in the relugolix CT group was 8.3 weeks in LIBERTY 1 and 8.4 weeks in LIBERTY 2, compared with 25.1 weeks and 27.1 weeks, respectively, in the placebo groups (nominal $p < 0.0001$ for both comparisons) (Figure 8, Section 5.1.2.1.1, page 102 of the LIBERTY 1 CSR; Figure 9, Section 5.1.2.1.1, page 106 of the

LIBERTY 2 CSR).[13, 14] This should be interpreted with caution due to the small number of participants in the analysis.

- Amenorrhoea: 67 (52.3%) and 63 (50.4%) of women who received relugolix CT in LIBERTY 1 and LIBERTY 2, respectively, achieved amenorrhea over the last 35 days of treatment compared with 7 (5.5%) and 4 (3.1%) women who received placebo ($p<0.001$ for both comparisons). Additionally, a greater proportion of participants in the relugolix CT group compared with the placebo group in both trials achieved sustained amenorrhoea, defined as the maintenance of amenorrhea at every subsequent visit after the initial achievement of amenorrhoea, at Weeks 8, 12, 16, 20, and 24 (nominal $p < 0.0001$).
- Bleeding and pelvic discomfort (BPD) was defined as the LS mean change from baseline to Week 24 as measured by the uterine fibroid health and symptom-related quality of life (UFS-QoL) BPD scale score (score range 0-100 with higher score value indicating greater distress). In LIBERTY 1, the UFS-QoL BPD score decreased (improved) by -45.0 points in the relugolix CT group, which was greater than the change observed in the placebo group (-16.1 points) ($p < 0.0001$). Similar results were reported for LIBERTY 2, with a UFS-QoL BPD score reduction of -51.7 points in the relugolix CT group compared with a reduction of -18.3 points in the placebo group ($p < 0.0001$).
- Pain associated with uterine fibroids: Pain associated with uterine fibroids was assessed in the subset of pain evaluable participants who had moderate-to-severe pain at baseline (maximum numerical rating scale [NRS] score ≥ 4). Approximately 50% of the participants were considered evaluable for pain (for relugolix CT and placebo, n = 58 and 69, respectively, in LIBERTY 1; and n = 68 and 82, respectively, in LIBERTY 2). In both trials, the proportions of evaluable participants who had achieved reductions to minimal or no pain (maximum NRS ≤1) were higher in the relugolix CT group than in the placebo group (LIBERTY 1: 43.1% versus 10.1%, $p < 0.0001$; LIBERTY 2: 47.1% versus 17.1%, $p < 0.0001$).
- Uterine volume (UV) / Primary uterine fibroid volume (UFV): based on the LS mean percent change from baseline to Week 24, the overall UV reduction in both the LIBERTY 1 and LIBERTY 2 trials was greater for relugolix CT

than placebo (LIBERTY 1: -12.9% versus 2.2%; p<0.001; LIBERTY 2: -13.8% versus -1.5%; p = 0.008). The reduction for primary UFV was numerically favourable for relugolix CT compared with placebo, although the difference between groups did not reach statistical significance (LIBERTY 1: -12.4% versus -0.3%; p = 0.09; LIBERTY 2: -17.4% versus -7.4.0%; p = 0.22)

- Change in haemoglobin levels: Defined as the proportion of women with anaemia (haemoglobin ≤10.5 g/dL) at baseline who achieve an increase of >2 g/dL from baseline to Week 24. Among the participants who had baseline anaemia (30 and 23 women in the relugolix CT group and placebo group, respectively, in LIBERTY 1 and 31 and 37 in LIBERTY 2), the outcome was significantly better with relugolix CT than with placebo (LIBERTY 1: 50.0% vs. 21.7%, p = 0.0377; LIBERTY 2: 61.3% vs 5.4%, p < 0.0001).

A summary of key outcomes in LIBERTY1 and LIBERTY2 is presented in Table 8.

Table 8 Summary of LIBERTY 1 and LIBERTY 2 outcomes [adapted from Table 18, Document B of the CS]

Endpoint	LIBERTY 1		LIBERTY 2	
	Placebo (n=127)	Relugolix CT (n=128)	Placebo (n=129)	Relugolix CT (n=125)
Primary Efficacy Endpoint				
Proportion of women with MBL volume < 80 mL & ≥ 50% reduction*				
n (%) Difference 95% CI (unadjusted) p-value	24 (19%)	94 (73%) 55% (44%, 65%) < 0.001	19 (14.73%)	89 (71%) 56% (46%, 66%) < 0.001
Secondary Efficacy Endpoint				
Proportion of women who achieved amenorrhoea over the last 35 days of treatment				
n (%) Difference (95% CI) p-value	7 (6%)	67 (52%) 47% (37%, 56%) < 0.001	4 (3%)	63 (50%) 47% (38%, 57%) < 0.001
% change in MBL volume (baseline to Week 24)				
LS mean (SD) Difference (95% CI) p-value	-23.2 (±4.6)	-84.3 (±4.7) -61.1 (-73.5, -48.6) < 0.001	-15.1 (±5.5)	-84.3 (±5.5) -69.2 (-84.1, -54.3) < 0.001
Change in UFS-QoL BPD score (baseline to Week 24º)				
LS mean (SD) Difference (95% CI) p-value	-16.1 (±2.8)	-45.0 (±2.9) -28.9 (-36.3, -21.5) < 0.001	-18.3 (±2.9)	-51.7 (±2.9) -33.4 (-41.2, -25.5) < 0.001
Proportion of women with anaemia (i.e. ≤10.5 g/dL) at baseline who achieved a Hb increase of > 2 g/dL (baseline to Week 24)				
n/N (%) Difference (95% CI) p-value	5/23 (22%)	15/30 (50%) 28% (4%, 53%) 0.04	2/37 (5%)	19/31 (61%) 56% (37%, 75%) < 0.001
Proportion of women who achieved a maximum NRS score ≤ 1 for UF-associated pain over the last 35 days of treatment in the subset of women with a maximum pain score ≥ 4 during the 35 days prior to randomisation				
n/N (%) Difference (95% CI) p-value	7/69 (10%)	25/58 (43%) 33% (18%, 48%) < 0.001	14/82 (17%)	32/68 (47%) 30% (16%, 44%) < 0.001
% change in primary UFV (baseline to Week 24)				

LS mean (SD) Difference (95% CI) p-value	-0.3 (±5.40)	-12.4 (±5.62) -12.1 (-26.3, 2.0) 0.09	-7.4 (±5.9)	-17.4 (±5.9) -10.0 (-25.8, 5.8) 0.2153
% change in UV (baseline to Week 24)				
LS mean (SD) Difference (95% CI) p-value	2.2 (±3.01)	-12.9 (±3.1) -15.1 (-23.0, -7.3) <0.001	-1.5 (±3.4)	-13.8 (±3.4) -12.2 (-21.3, -3.2) 0.008

Note: * from baseline MBL volume. º score as measured by the UFS-QoL (Q1, Q2, Q5).
CI: Confidence Interval; CT: Combination Therapy; Hb: Haemoglobin; LS: least-squares; MBL: menstrual blood loss; NRS: numerical rating scale; UFS-QoL BPD: uterine fibroid health and symptom-related quality of life bleeding and pelvic discomfort; UFV: uterine fibroid volume; UV: uterine volume

Primary and secondary endpoints in LIBERTY 3

The company also presents the long-term results of the 28-week LIBERTY 3 extension study, in which women who completed one of 24-week parent studies LIBERTY 1 or LIBERTY 2 entered the open-label phase.

The primary endpoint was the proportion of responders, defined as women who achieved or maintained an MBL volume of < 80 mL and at least a 50% reduction from parent study baseline MBL volume to the last 35 days of treatment, as measured by the AH method. Key secondary endpoints in LIBERTY 3 included achievement of amenorrhoea, improvement of anaemia assessed by changes in haemoglobin concentrations, UFS-QoL score, uterine volume and uterine fibroid volume. Primary and secondary endpoints are summarised in Table 9 below.

Table 9 Summary of outcomes assessed in the LIBERTY 3 extension study

	Randomisation in parent trial	
	Placebo (N = 164)	**Relugolix CT (N = 163)**
Proportion of responders[a] at Week 52, n (%)	[redacted]	[redacted]
Proportion of patients who achieved amenorrhoea at Week 52, n (%)	[redacted]	[redacted]
Proportion of women with anaemia (i.e. ≤10.5 g/dL) at parent study baseline who achieved a Hb increase of > 2 g/dL at Week 52, n/N (%)	[redacted]	[redacted]
Change in UFS-QoL BPD scale score from parent study baseline to Week 52	[redacted]	-51.3 points
Proportion of responders[b] on the UFS-QoL BPD scale at Week 52	[redacted]	[redacted]
Percent change in uterine volume from parent study baseline to Week 52	[redacted]	[redacted]
Percent change in uterine fibroid volume from parent study baseline to Week 52	[redacted]	[redacted]

[a] MBL volume < 80 mL and ≥ 50% reduction from baseline over the last 35 days of treatment
[b] At least a 20-point reduction
CT: Combination Therapy; Hb: Haemoglobin; UFS-QoL BPD: uterine fibroid health and symptom-related quality of life bleeding and pelvic discomfort;

3.2.3 Subgroup analyses related to the primary endpoint

Subgroup analyses were not specified in the NICE final scope. Section B.2.4 of the CS stated that subgroup analyses of the primary efficacy endpoint were conducted in LIBERTY 1, LIBERTY 2, and LIBERTY 3 trials. Details of these subgroup analyses were provided by the company at the clarification stage.

In LIBERTY 1 and LIBERTY 2, subgroup analyses were conducted for the following groups: geographical region, age, baseline MBL volume, race, body mass index (BMI), uterine volume at baseline, maximum NRS score at baseline and history of prior pregnancy (Figure 2 and Figure 3 in the clarification response). LIBERTY 3 included the following four additional subgroups: MBL volume at parent study baseline, uterine fibroid volume, and alcohol use and smoking status (see Figure 4 in the clarification response).

In all three studies, the direction of effect across all subgroups appears generally consistent with that observed in the overall study population. However, across all studies, the size of the effect was smaller for Black/African American women relative to White women, and for women with larger uterine volumes (≥ 300 cm^3) relative to those with smaller uterine volume (<300 cm^3). Smaller effect size was also observed for women with greater MBL volume at baseline (≥ 225 mL) in LIBERTY 1 and 3.

3.2.4 Adverse events

LIBERTY 1 and LIBERTY 2

The safety population of LIBERTY 1 and LIBERTY 2 included all participants who received any amount of study drug (LIBERTY 1: relugolix CT, n = 128; placebo, n = 127; LIBERTY 2: relugolix CT, n = 126; placebo, n = 129). The methods used to assess safety are reported in Sections B.2.4 and B.2.10, Document B of the company submission, and are considered appropriate by the ERG. Tables 28 and 29 in Document B of the CS show adverse events for LIBERTY 1 and LIBERTY 2 and are reproduced as Table 10 below. The ERG's clinical expert considers the overall incidence and the types of adverse events for relugolix CT akin to those expected in clinical practice in this clinical population.

Table 10 Summary of adverse events in the LIBERTY 1 and LIBERTY 2 safety population [reproduced from Tables 28 and 29, Document B of the CS]

Characteristics N (%)	LIBERTY 1		LIBERTY 2	
	Placebo (N=127)	Relugolix CT (N=128)	Placebo (N=129)	Relugolix CT (N=126)
Any	84 (66%)	79 (62%)	76 (59%)	76 (60%)
Leading to discontinuation	5 (4%)	7 (5%)	6 (5%)	3 (2%)
Serious	2 (2%)	7 (5%)	4 (3%)	1(1%)
Fatal outcome	0	0	0	0
Adverse event reported in >5% of participants in any group				
Hot flush	10 (8%)	14 (11%)	5 (4%)	7 (6%)
Headache	19 (15%)	14 (11%)	15 (12%)	11 (9%)
Hypertension	0	7 (5%)	4 (3%)	5 (4%)
Arthralgia	4 (3%)	4 (3%)	4 (3%)	1 (1%)
Cough	7 (6%)	1 (1%)	4 (3%)	0
Nausea	6 (5%)	4 (3%)	10 (8%)	6 (5%)
URTI	3 (2%)	1 (1%)	7 (5%)	6 (5%)
Anaemia	6 (5%)	4 (3%)	8 (6%)	2 (2%)
Fatigue	5 (4%)	4 (3%)	2 (2%)	1 (1%)

CT: combination therapy; URTI: upper respiratory tract infection

In LIBERTY 1, during the 24-week study period, the proportion of women treated with relugolix CT who experienced 'any' adverse events was 62% compared with 66% of those treated with placebo. In LIBERTY 2 the incidence of adverse events was 60% and 59%, respectively. The most frequently reported adverse events in any treatment group included headache and hot flush.

The most frequently reported vasomotor symptom through week 24, by preferred term, was hot flush, which was reported more frequently in the relugolix CT group than in the placebo group in both trials (14 [11%] versus 10 [8%] in LIBERTY 1; 7 [6%] versus 5 [4%] in LIBERTY 2). The hot flush events were reported mostly to be Grade 1 or Grade 2 in severity.[13, 14]

No deaths were reported across both trials.

Least-squares mean percent changes from baseline in BMD at the lumbar spine (L1 - L4) in the relugolix CT group compared with placebo at week 24 were -0.356% versus 0.052% for LIBERTY 1 and -0.126% versus 0.315% for LIBERTY 2, with no

significant difference observed between the groups. Similarly, the percent change to week 24 in BMD at the total hip was similar in the relugolix CT and placebo groups in both trials (LIBERTY 1: 0.023% versus 0.549%; LIBERTY 2: -0.0173% versus 0.044%) (CSR, Table 32, page 145 for LIBERTY 1; Table 29, page 139 for LIBERTY 2).[13, 14] BMD was measured by means of a dual-energy x-ray absorptiometry (DEXA).

Serious adverse events (SAE) in LIBERTY 1 were reported for 7 participants (5.5%) in the relugolix CT group and 2 participants (1.6%) in the placebo group. In the relugolix CT group two SAEs were related to expulsion/prolapse of uterine fibroid, and one of these events was assessed as related to study drugs. In LIBERTY 2, SAEs were reported for 1 woman (0.8%) in the relugolix CT group and 4 women (3.1%) in the placebo group, none of them were considered to be related to the study drug.

LIBERTY 3

Cumulatively over the 52-week treatment period encompassing the parent (24 weeks) and open-label extension (28 weeks) studies, [redacted] of participants in the relugolix CT group reported at least one treatment-emergent adverse event (TEAE). [redacted] of the participants in this group experienced one TEAE during the open-label extension study. Grade 3 or higher events were reported for [redacted] in the relugolix CT group, with the event first occurring in the open-label extension study for [redacted]. Among those in the placebo group in the parent study, at least one TEAE was reported for [redacted] cumulatively and [redacted] during the extension. [redacted] were reported during the study. [redacted] The ERG agrees with the company's conclusions.

A summary of serious adverse events reported during LIBERTY 3 are provided in Table 32 of the CS and reproduced as Table 11 below.

Table 11 Summary of serious adverse events by System Organ Class and Preferred Term from the extension safety population of LIBERTY 3 [reproduced from Table 32, Document B of the CS]

Characteristics	LIBERTY 3 Randomisation in parent trial			
	Placebo (N=164)		Relugolix CT (N=163)	
	Cumulative	Extension	Cumulative	Extension
No. of patients with at least one serious AE n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Blood and lymphatic disorders	[redacted]	[redacted]	[redacted]	[redacted]
Anaemia	[redacted]	[redacted]	[redacted]	[redacted]
Cardiac disorders	[redacted]	[redacted]	[redacted]	[redacted]
Atrial fibrillation	[redacted]	[redacted]	[redacted]	[redacted]
Eye disorders	[redacted]	[redacted]	[redacted]	[redacted]
Vitreous detachment	[redacted]	[redacted]	[redacted]	[redacted]
Hepatobiliary disorders	[redacted]	[redacted]	[redacted]	[redacted]
Cholecystitis	[redacted]	[redacted]	[redacted]	[redacted]
Cholecystitis acute	[redacted]	[redacted]	[redacted]	[redacted]
Cholelithiasis	[redacted]	[redacted]	[redacted]	[redacted]
Infections and infestations	[redacted]	[redacted]	[redacted]	[redacted]
Appendicitis	[redacted]	[redacted]	[redacted]	[redacted]
Pneumonia	[redacted]	[redacted]	[redacted]	[redacted]
Injury, poisoning and procedural complications	[redacted]	[redacted]	[redacted]	[redacted]
Ankle fracture	[redacted]	[redacted]	[redacted]	[redacted]
Avulsion fracture	[redacted]	[redacted]	[redacted]	[redacted]
Forearm fracture	[redacted]	[redacted]	[redacted]	[redacted]
Radius fracture	[redacted]	[redacted]	[redacted]	[redacted]
Road traffic accident	[redacted]	[redacted]	[redacted]	[redacted]
Wrist fracture	[redacted]	[redacted]	[redacted]	[redacted]
Investigations	[redacted]	[redacted]	[redacted]	[redacted]
Blood pressure increased	[redacted]	[redacted]	[redacted]	[redacted]
Musculoskeletal and connective tissue disorders	[redacted]	[redacted]	[redacted]	[redacted]
Intervertebral disc protrusion	[redacted]	[redacted]	[redacted]	[redacted]
Neoplasms benign, malignant and unspecified (including cysts & polyps)	[redacted]	[redacted]	[redacted]	[redacted]
Uterine leiomyoma	[redacted]	[redacted]	[redacted]	[redacted]
Uterine myoma expulsion	[redacted]	[redacted]	[redacted]	[redacted]
Nervous system disorders	[redacted]	[redacted]	[redacted]	[redacted]
Syncope	[redacted]	[redacted]	[redacted]	[redacted]
Psychiatric disorders	[redacted]	[redacted]	[redacted]	[redacted]
Panic attack	[redacted]	[redacted]	[redacted]	[redacted]

Renal and urinary disorders	[redacted]										
Nephrolithiasis	[redacted]										
Reproductive system and breast disorders	[redacted]										
Menorrhagia	[redacted]										
Metrorrhagia	[redacted]										
Ovarian cyst ruptured	[redacted]										
Uterine haemorrhage	[redacted]										

Abbreviations: AE = adverse event; n = number of patients in subset; N = number of patients.
Note: Percentages are based on the total number of patients in each treatment group.
Note: Patients with multiple events for a given preferred term or system organ class were counted only once for each preferred term and system organ class.
Note: Cumulative represents the entire treatment period since randomisation in study LIBERTY 1 or LIBERTY 2. Data in the Extension columns relate to the treatment period since enrolment into LIBERTY 3 only.

3.2.5 Meta-analyses

The company did not perform a meta-analysis.

3.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

For the ITC, the company presents evidence from LIBERTY 1, LIBERTY 2 and two further Phase-3 double-blind RCTs (PEARL I, PEARL II). PEARL I assessed the efficacy and safety of ulipristal (UPA) versus placebo for the pre-operative treatment of symptomatic UF and PEARL II assessed the efficacy and safety of UPA versus the GnRH agonist leuprolide acetate in the pre-operative treatment of symptomatic UF.

The baseline demographic characteristics of participants in PEARL I and PEARL II are presented in Table 121 and Table 122, Document B of the CS and reproduced as Tables 12 and 13 below.

Table 12 Summary of the baseline demographic characteristics of PEARL I [reproduced from Table 121, Appendix M, Document B of the CS]

Characteristic		Treatment Group			Total (N=241)
		Placebo (N=48)	UPA 5 mg (N=95)	UPA 10 mg (N=98)	
Age	**N**	48	95	98	241
	Mean	41.6	41.2	42.0	41.6
	SD	5.6	5.9	5.5	5.7
	Median	42.5	42.0	43.0	43.0
	Min, Max	26, 50	24, 50	23, 50	23, 50
Ethnic Origin	**White**	41 (85.4%)	84 (88.4%)	87 (88.8%)	212 (88.0%)
	Black	0	0	0	0
	Asian	7 (14.6%)	11 (11.6%)	11 (11.2%)	29 (12.0%)
	Hispanic	0	0	0	0
	Other	0	0	0	0
Fertility Status	**Not of Childbearing Potential**	5 (10.4%)	8 (8.4%)	6 (6.1%)	19 (7.9%)
	Of Childbearing Potential	43 (89.6%)	87 (91.6%)	92 (93.9%)	222 (92.1%)
	Missing	0	0	0	0
Weight (kg)	**N**	48	95	98	241
	Mean	64.70	70.05	67.12	67.79
	SD	12.47	13.60	10.25	12.22
	Median	60.40	68.00	66.00	66.00
	Min, Max	45.0, 106.5	42.0, 120.0	48.9, 95.0	42.0, 120.0
Height (cm)	**N**	48	95	98	241
	Mean	162.3	164.3	163.9	163.7
	SD	6.6	6.5	6.1	6.4
	Median	163.5	164.0	164.0	164.0
	Min, Max	143, 176	150, 178	145, 178	143, 178
Body Mass Index (kg/m^2)	**N**	48	95	98	241
	Mean	24.55	25.93	25.03	25.29
	SD	4.37	4.63	3.92	4.32
	Median	24.49	25.39	24.87	24.96
	Min, Max	18.0, 40.1	18.1, 39.2	18.1, 37.6	18.0, 40.1

Table 13 Summary of baseline demographic characteristics of PEARL II [reproduced from Table 122, Appendix M, Document B of the CS]

Characteristic		Treatment Group			Total (N=301)
		UPA 5 mg (N=97)	UPA 10 mg (N=103)	GnRH-agonist (N=101)	
Age	N	97	103	101	301
	Mean	40.1	40.7	40.3	40.4
	SD	6.2	6.3	6.2	6.2
	Min, Max	25, 50	20, 50	24, 51	20, 51
Ethnic Origin	White	83 (85.6%)	88 (85.4%)	85 (84.2%)	256 (85.0%)
	Black	9 (9.3%)	11 (10.7%)	9 (8.9%)	29 (9.6%)
	Asian	1 (1.0%)	1 (1.0%)	0	2 (0.7%)
	Hispanic	3 (3.1%)	2 (1.9%)	5 (5.0%)	10 (3.3%)
	Other	1 (1.0%)	1 (1.0%)	2 (2.0%)	4 (1.3%)
Fertility Status	Not of Childbearing Potential	4 (4.1%)	4 (3.9%)	3 (3.0%)	11 (3.7%)
	Of Childbearing Potential	93 (95.9%)	99 (96.1%)	98 (97.0%)	290 (96.3%)
Weight (kg)	N	97	103	100	300
	Mean	68.26	68.84	67.92	68.35
	SD	12.28	12.72	12.16	12.36
	Min, Max	48.5, 108.0	46.0, 111.0	48.0, 119.0	46.0, 119.0
Height (cm)	N	97	103	100	300
	Mean	163.7	162.3	165.2	163.7
	SD	6.4	6.7	5.9	6.4
	Min, Max	146, 180	146, 180	147, 178	146, 180
Body Mass Index(kg/m^2)	N	97	103	100	300
	Mean	25.44	26.15	24.86	25.49
	SD	4.08	4.74	4.06	4.33
	Min, Max	19.4, 37.8	18.1, 39.8	18.4, 39.3	18.1, 39.8

Disease-specific baseline characteristics of participants in PEARL I and PEARL II are presented in Table 14. The CS presents a comparison of the patient characteristics of the LIBERTY and PEARL studies in section M1.6 of the Appendices. The demographic characteristics are balanced within the PEARL I and PEARL II studies and appear similar between the studies. The percentage of White ethnic origin was much higher in PEARL studies compared to LIBERTY studies. The BMI of the participants in LIBERTY studies were higher compared to PEARL studies which will have a negative effect on the relative effect of relugolix CT. The ERG is concerned though participants in the PEARL studies were expected to receive surgery after 13

weeks while those in the LIBERTY studies appear unlikely to be receiving surgery. While this is not necessarily shown in the baseline characteristics it does suggest two different populations in the respective studies.

Table 14 Baseline disease-specific characteristics of participants in PEARL I and PEARL II [adapted from Table 1 of Donnez 2012a and Table 1 of Donnez 2012b][15, 16]

	PEARL I			PEARL II		
	Placebo (n=48)	**UPA 5mg (n=95)**	**UPA 10mg (n=98)**	**UPA 5mg (n=97)**	**UPA 10mg (n=103)**	**Leuprolide acetate (n=101)**
PBAC score, median (IQR)	376 (241-608)	386 (235-627)	330 (235-537)	286 (190-457)	271 (183-392)	297 (189-443)
Haemoglobin, g/dL, mean (SD)	9.6 (1.2)	9.3 (1.5)	9.5 (1.6)	12.4 (1.6)	12.4 (1.6)	12.1 (1.8)
Total UF volume, cm^3, median (IQR)	61.9 (24.8-158.9)	100.7 (40.0-205.3)	96.7 (31.7-181.3)	79.6 (30.3-151.0)	47.6 (24.1-110.6)	59.2 (27.8-156.3)
Uterine volume, cm^3, median (IQR)	318.8 (216.0-496.3)	337.6 (236.1-502.8)	325.6 (212.6-453.3)	199.4 (149.6-315.0)	197.8 (120.9-297.7)	199.9 (138.2-271.9)
UFS-QoL (symptom severity subscale), mean (SD)	NR	NR	NR	54.0 (20.0)	48.9 (22.1)	52.5 (21.7)

Note. UPA: [a]Total volume of 3 largest myomas, cm^3; UPA: ulipristal acetate; PBAC: pictorial blood loss assessment chart; SD: standard deviation; IQR: interquartile range; NR: not reported

Comparison of the disease specific characteristics suggest that the participants in PEARL I are in poorer health than those in PEARL II. The PEARL II disease specific characteristics are also similar to the participants in the LIBERTY studies, although the uterine volumes were higher in LIBERTY studies compared to PEARL II.

Tabulated results for the efficacy endpoints for PEARL I and PEARL II were provided by the company at clarification and are reproduced in Tables 15 and 16 below.

Table 15 PEARL I efficacy results for UPA 5mg and placebo groups [reproduced from Table 4 of the company's clarification response]

Endpoint	Placebo (N = 48)	UPA 5 mg (N = 95)	Difference, 5 mg UPA − Placebo (95% CI)†	P Value
Primary endpoints at week 13				
PBAC <75 — no./total no. (%)	9/48 (19)	86/94 (91)	73 (55 to 83)	<0.001
% Change from screening in total fibroid volume‡				0.002
Median	3.0	−21.2	−22.6 (−36.1 to −8.2)	
Interquartile range	−19.7 to 23.0	−41.2 to −1.1		
Secondary endpoints at week 13				
Baseline PBAC				
Median	376	386		
Interquartile range	241 to 608	235 to 627		
Wk 9-12 PBAC				
Median	336	0		
Interquartile range	115 to 543	0 to 5		
Change from baseline to wk 9-12 in PBAC				
Median	−59	−329	−291 (−399 to −194)	<0.001
Interquartile range	−216 to 58	−571 to −205		
Amenorrhea, PBAC ≤2, at wk 9–12 — no./total no. (%)	3/48 (6)	69/94 (73)	67 (50 to 77)	<0.001
Total reduction ≥25% in fibroid volume at wk 13 — no./ total no. (%)	8/45 (18)	35/85 (41)	23 (4 to 39)	0.01
% Change from screening in uterine volume at wk 13				0.001§
Median	5.9	−12.1		
Interquartile range	−3.8 to 18.4	−28.3 to 2.9		
Reduction in uterine volume ≥25% at wk 13 — no./ total no. (%)	3/47 (6)	30/88 (34)	28 (11 to 40)	<0.001
Haemoglobin – g/dl				
Baseline	9.55±1.18	9.32±1.50		
Wk 13	12.61±1.30	13.50±1.32		
Change from baseline to wk 13	3.10±1.68	4.25±1.90	0.92 (0.39 to 1.44)	<0.001
Pain assessment with Short-Form McGill Pain Questionnaire				
Baseline				
Median	8.5	6.5		
Interquartile range	3.0 to 18.0	3.0 to 15.0		
Wk 13				
Median	4.2	1.0		

Interquartile range	1.0 to 10.0	0.0 to 4.0		
Change from baseline to wk 13				
Median	−2.5	−5.0	−2.0 (−4.0 to 0.0)	0.10
Interquartile range	−6.3 to 1.0	−8.0 to −2.0		
Measurement of discomfort questionnaire				
Baseline				
Median	16.0	14.0		
Interquartile range	13.5 to 18.0	10.0 to 19.0		
Wk 13				
Median	11.0	3.0		
Interquartile range	4.0 to 15.0	1.0 to 7.0		
Change from baseline to wk 13				
Median	−6.0	−9.0	−4.0 (−6.0 to −1.0)	0.001
Interquartile range	−9.0 to −2.0	−13.0 to −6.0		

* All confidence intervals and P values have been adjusted for multiplicity (Bonferroni correction) because two doses of ulipristal acetate were compared with placebo (i.e., P values were multiplied by 2). PBAC denotes pictorial blood-loss assessment chart.

† The differences in categories with numbers and percents are percentage-point differences. The differences in categories with medians and interquartile ranges are differences in medians, as calculated with the use of the Hodges–Lehmann estimator.

‡ The percent change from screening in total fibroid volume was assessed in 45 patients in the placebo group, 85 patients in the 5-mg ulipristal acetate group, and 80 patients in the 10-mg ulipristal acetate group.

Table 16 PEARL II efficacy results for UPA 5mg and leuprolide acetate groups (per protocol population) [reproduced from Table 5 of the company's clarification response]

	UPA 5mg (N = 93)	Leuprolide acetate (N = 93)	Difference, 5 mg UPA vs. Leuprolide acetate (95% CI)
Primary efficacy endpoints at week 13			
PBAC <75 — no./total no. (%)	84/93 (90)	82/92 (89)†	1.2 (−9.3 to 11.8)‡
Secondary efficacy endpoints			
Median (IQR)	0 (0 to 2)	0 (0 to 1)	
Change from baseline — median (IQR)	−268 (−412 to −172)	−274 (−430 to −161)	6 (−54 to 63)
≤2, indicating amenorrhea — no./total no. (%)	70/93 (75)	74/92 (80)	−5.2 (−18.7 to 8.6)
Total volume of three largest myomas			
Percent change from baseline — median (IQR)	−36 (−58 to −11)	−53 (−69 to −36)	
Ratio to screening volume — geometric mean	0.66	0.54	1.23 (0.99 to 1.52)
Uterine volume			
Percent change from baseline — median (IQR)	−20 (−40 to −3)	−47 (−57 to −35)	
Ratio to screening volume — geometric mean	0.84	0.57	1.48 (1.25 to 1.74)
Short-Form McGill Pain Questionnaire Score			
Median (IQR)	2.0 (0.0 to 4.0)	0.0 (0.0 to 4.0)	
Change from baseline — median (IQR)	−5.0 (−11.0 to −2.0)	−5.5 (−14.5 to −2.0)	0.2 (−2.0 to 3.0)
Uterine Fibroid Symptom and Quality of Life questionnaire			
Health-related quality of life score	76.4±23.2	73.2±23.0	
Change from baseline	23.7±26.9	23.2±28.2	2.5 (−7.3 to 12.3)
Haemoglobin — g/dl	12.8±1.4	12.7±1.6	−0.02 (−0.3 to 0.3)

† One patient had a missing score on the pictorial blood-loss assessment chart.
‡ A lower limit of the confidence interval of more than −20% (the prespecified noninferiority margin) indicates noninferiority. A lower limit of the confidence interval of more than zero indicates superiority.

Tables 15 and 16 show effect sizes favouring UPA 5mg in comparison with placebo and similar benefits from UPA 5mg and leuprolide acetate. In both PEARL I and PEARL II, UPA 5mg can be seen to reduce MBL and uterine volume and increase the haemoglobin level. PEARL II shows GnRHa reduces MBL and uterine volume.

Tabulated safety results for PEARL I and PEARL II were also provided by the company at clarification and are reproduced in Tables 17 and 18 below.

Table 17 PEARL I summary of adverse events in the UPA 5mg and placebo groups (safety population) [reproduced from Table 6 of the company's clarification response]

Event *	**Placebo (N = 48)** number (%)	**UPA 5 mg (N = 95)** number (%)
At least one serious adverse event	3 (6)	2 (2)
Serious adverse event during treatment period	1 (2)	0
Uterine haemorrhage	0	0
Fibroid protruding through cervix	1 (2)	0
Serious adverse event within 4 wk after treatment period	1 (2)	2 (2)
Uterine haemorrhage	0	1 (1)
Breast cancer	1 (2)	0
Ovarian haemorrhage	0	1 (1)
Serious adverse event from wk 17 to wk 38	1 (2)	0
Menometrorrhagia	1 (2)	0
Uterine haemorrhage	0	0
Adverse event leading to discontinuation of study drug†	1 (2)	1 (1)
At least one adverse event‡	22 (46)	47 (49)
Headache	2 (4)	4 (4)
Breast pain, tenderness, or discomfort	0	2 (2)
Abdominal pain	2 (4)	2 (2)
Pyrexia	2 (4)	3 (3)
Hypercholesterolemia	1 (2)	3 (3)
Hypothyroidism	0	2 (2)
Constipation	1 (2)	4 (4)
Hypertriglyceridemia	1 (2)	3 (3)
Influenza	1 (2)	1 (1)
Dizziness	0	1 (1)
Nasopharyngitis	0	3 (3)
Dysmenorrhoea	2 (4)	0

* All serious adverse events and adverse events occurring in at least 3% of the patients in any group are included. Patients could have more than one adverse event of the same type. There were no significant differences between either ulipristal acetate group and the placebo group for any adverse event, with two-sided P values calculated with the use of Fisher's exact test and no adjustment for multiplicity.

† The adverse events leading to discontinuation of the study drug were breast cancer (one patient in the placebo group), endometrial changes (one patient in the 5-mg ulipristal acetate group, with the event initially reported by the local laboratory as hyperplasia but later diagnosed as benign endometrium by three pathologists who were unaware of the study-group assignments).

‡ Adverse events with onset at or after the first dose of study drug and on or before the last assessment date of week 17 (4 weeks after the end of the treatment period) are included.

Table 18 PEARL II summary of adverse events in the UPA 5mg and leuprolide acetate groups (safety population) [reproduced from Table 7 of the company's clarification response]

Event *	**UPA 5mg (N = 97)** number (%)	**Leuprolide acetate (N = 101)** number (%)
At least one event	8 (8)	6 (6)
Any event during treatment	2 (2)	2 (2)
Headache	1 (1)	0
Fibroid protruding through cervix	0	0
Lung infection	0	1 (1)
Thyroid cancer	1 (1)	0
Uterine haemorrhage	0	1 (1)
Within 4 wk after treatment†	3 (3)	2 (2)
From wk 17 to 38‡	3 (3)	2 (2)
Adverse events		
Leading to study-drug discontinuation	1 (1)	6 (6)
At least one event¶	75 (77)	85 (84)
Hot flash	25 (26)	66 (65)
Headache	25 (26)	29 (29)
Procedural pain	9 (9)	9 (9)
Abdominal pain	6 (6)	14 (14)
Nausea	6 (6)	6 (6)
Fatigue	4 (4)	3 (3)
Anaemia	5 (5)	5 (5)
Nasopharyngitis	6 (6)	2 (2)
Acne	0	5 (5)
Breast pain or tenderness	5 (5)	2 (2)
Influenza	2 (2)	5 (5)
Insomnia	2 (2)	5 (5)
Pharyngitis	5 (5)	2 (2)

* Listed are all serious adverse events and adverse events that occurred in at least 5% of patients in each study group, including events that were considered to be unrelated to the study drug. There were no significant between-group differences for any adverse event except hot flashes (P<0.001 for both doses of ulipristal acetate vs. leuprolide acetate). No adjustment for multiplicity was performed.
† These serious adverse events were operative complications in two patients and sarcoma in one patient (retrospectively diagnosed after further review after premature discontinuation of the study drug) in the group receiving 5 mg of ulipristal acetate; endometrial polyp, haemangioma, and operative complications and lymphocytic choriomeningitis in one patient each in the group receiving leuprolide acetate.
‡ These serious adverse events were spontaneous abortion, surgery for suspected ovarian tumour but intraoperative diagnosis corrected to new uterine myoma, and vaginal haemorrhage in one patient each receiving 5 mg of ulipristal acetate; and uterine haemorrhage in two patients receiving leuprolide acetate.

The ERG does not have any concern over the adverse event rates in PEARL I. However, results from PEARL II suggest lower rate of headaches, hot flushes and abdominal pain in relugolix CT.

3.4 Critique of the indirect comparison and/or multiple treatment comparison

Based on the data from LIBERTY 1, LIBERTY 2, PEARL I and PEARL II trials, Tables 19 and 20 below show the mean difference in percentage change from baseline in MBL from the ITC for relugolix CT versus UPA and leuprorelin acetate (GnRHa) versus UPA. The ITC results indicate that at week 4 and week 12 relugolix CT had a larger mean percentage decrease in MBL compared with UPA. At 8 weeks UPA showed a larger decrease compared with relugolix CT. In both tables, the confidence intervals are very wide indicating uncertainty around the point estimates.

Table 19 ITC results: relugolix CT versus UPA

	Mean difference %-CFB Week 4	**Mean difference %-CFB Week 8**	**Mean difference %-CFB Week 12**	**Mean difference %-CFB Week 24-no hysterectomy* (UPA patients not on treatment)**	**Mean difference %-CFB Week 24-no surgery** (UPA patients not on treatment)**
Relugolix CT vs. UPA (95% CI)	-19.43% (-55.32%, 16.46%)	+4.53% (-22.62%, 31.69%)	-10.73% (-39.41%, 17.94%)	-77.63% (-119.79%, -35.46%)	-63.06% (-106.93%, -19.18%)
Heterogeneity statistic Chi^2	1.125 (p=0.289)	0.107 (p=0.744)	0.538 (p=0.463)	13.021 (p<0.001)	7.936 (p=0.005)

CFB: Change from baseline
* No hysterectomy or endometrium ablation post treatment in the PEARL trials.
** No surgery post treatment in the PEARL trials.
Note: Treatment in the PEARL I and II trials was discontinued after week 13.

Table 20 Direct comparison: leuprorelin versus UPA

	Mean difference %-CFB Week 4	**Mean difference %-CFB Week 8**	**Mean difference %-CFB Week 12**	**Mean difference %-CFB Week 24-no hysterectomy ***	**Mean difference %-CFB Week 24-no surgery****
Leuprorelin vs. UPA (95% CI)	+31.14% (-52.49%, 114.77%)	-3.79% (-105.03%, 97.45%)	-1.50% (-71.05%, 68.05%)	+23.45% (-91.88%, 138.78%)	+14.12% (-114.80%, 143.04%)

Note: Treatment in the PEARL I and II trials was discontinued after week 13.

The company did not follow the ERG's suggestion of conducting a network meta-analysis (NMA) or ITC to compare relugolix CT with GnRHa. The company's justification for not performing an NMA was not considered satisfactory by the ERG: "*The only outcome used by the economic model that was informed by the indirect treatment comparison was MBL, which was subsequently used in the utility algorithm. In the majority of economic models where a network meta-analysis (NMA) is used to inform the model efficacy parameters, this is usually carried out on a small number of outcome measures deemed consistent or similar across studies in the network.*" In the absence of direct trial evidence or the opportunity to link relugolix CT and GnRHa through a common comparator, the ERG believes that a network would have been a more appropriate form of analysis as this would better represent the uncertainty, which exists due to the number of required comparisons and the difference in disease-specific characteristics between the PEARL I study and the LIBERTY and PEARL II trials. The ERG notes that the steps required to perform an NMA are similar to those undertaken by the company to perform the ITC.

The ERG also questioned why MBL volume was the only outcome for which the company attempted an ITC. While the ERG understands there may have been difficulties in comparing outcome measures, they notice that UFV/UV, haemoglobin levels, and health-related quality of life were reported in the LIBERTY and PEARL trials and could have been assessed using an ITC. It is also worth noting that time to MBL response, pain, UFV/UV, haemoglobin levels, and health-related quality of life were listed in both the NICE final scope and the company's decision problem and a comparison between relugolix CT and GnRHa was, therefore, expected. In particular, there is a lack of patient-reported outcomes measures (PROMs) in the CS because health-related quality of life measures were not assessed.

The ERG believes that a comprehensive summary of the clinical effectiveness of the technology is missing in the current CS. Similarly, the company's clarification response focused mainly on aspects related to the economic modelling rather than on aspects related to the clinical effectiveness of the technology.

3.5 Additional work on clinical effectiveness undertaken by the ERG

At clarification, the ERG queried by an NMA on MBL response had not been performed and consider performing this analysis themselves. However, performing an NMA using the currently available data would have required assumptions to be made such as approximating

the mean with median and the standard deviations with an adjustment of the interquartile range. Moreover, there are slight differences in the time points of the available outcome data. For these reasons, the ERG did not attempt the NMA.

Using the ITC results provided by the company, the ERG carried out ITCs comparing relugolix CT versus GnRHa. Results of these comparisons are presented in Table 21.

Table 21 ITC results: relugolix CT versus GnRHa

	Mean difference % - CFB (95% CI)
Week 4	-50.57 (-141.58, 40.44)
Week 8	8.32 (-96.50, 113.14)
Week 12	-9.23 (-84.46, 66.00)

The ERG agrees with the company's assumption that relugolix CT and GnRHa are equally effective for reducing MBL. However, all of the confidence intervals around the point estimates are wide and this observed uncertainty should be fed into the probabilistic analysis of the cost-effectiveness model (see Section 4.2.6).

3.6 Conclusions of the clinical effectiveness section

The company only presented the ITC results for MBL but did not attempt any ITC for other outcomes listed in either the NICE final scope or their decision problem. In particular, the company presented only a comparison between relugolix CT and UPA and a comparison between GnRHa and UPA but not a comparison between relugolix CT and GnRHa. The ERG believes the other outcomes in the scope could have been compared considering it is likely the company have access to data from the LIBERTY trials, which could be matched to the 13-week timepoint data in the PEARL trials.

The ERG has some concerns over the population of PEARL and LIBERTY trials as the participants in the PEARL trials were expected to receive surgery after 13 weeks while those in the LIBERTY trials appear unlikely to be receiving surgery. This suggests that two different populations were included in the respective trials. The PEARL I trial, which is required to link relugolix CT and GnRHa, appears to include participants who initially have higher MBL and uterine volume and lower haemoglobin levels.

The ERG strongly felt that an NMA should have been conducted by the company for MBL as well as for the other relevant outcomes. The ERG believes that an ITC is suitable when the required comparison can be made by linking two trials through a common comparator but given the evidence presented in the current CS, considers that a network would have been more appropriate. Nevertheless, the ERG agrees with the company's assertion that relugolix CT and GnRHa are equally effective in reducing MBL, even though the wide confidence intervals around the estimates of effect indicate some uncertainty.

The ERG has inspected the adverse events being reported in Table 29 of the CS and Tables 6-7 of the clarification response. The proportion of participants experiencing headaches, hot flushes, and abdominal pain were lower amongst patients receiving relugolix CT. The ERG is not concerned with any differences in serious adverse events or rates of adverse events.

Lastly, the ERG felt that the company did not place enough importance on the clinical effectiveness section of their submission and focused more on the cost-effectiveness section.

4 COST-EFFECTIVENESS

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

The company conducted a systematic literature review of cost-effectiveness analyses of pharmacological interventions used to treat fibroids for women who have failed conventional hormone therapy. A total of 63 records were identified, and 14 studies were included after screening and full-text review; 9 of which were unique economic evaluations. Full details of the cost-effectiveness review methods, including search strategies and selection criteria are provided in Appendix G of the company submission. Table 37 of the company submission summarises the identified studies.

The ERG has reviewed the company's search strategies and methodology and are satisfied that robust methods have been used to identify the literature. However, the ERG believes that the characteristics of the identified models (including modelled Markov states) should have been more clearly reported, and their usefulness for the current assessment critiqued. Therefore, the ERG provides further details of the model structures in Table 22 below, focusing on the Markov states included in models and their relevance to the current decision problem.

The ERG agrees with the company's assessment that none of the identified economic evaluations are directly relevant to the current decision problem, with all studies assessing the cost-effectiveness of UPA, often compared to GnRHa, and often using data from the PEARL studies. Half of the identified studies were abstracts of conference presentations and therefore provided limited information that might be useful for the development of the current model structure. Among the five published studies, three were Markov models, two of which defined Markov states according to health[17, 18] (based on bleeding and / or symptom control) prior to surgery, with one structure developed using states defined according to treatment received.[19] The ERG considers the definition of states for the economic model to be an important consideration and feels that further critique of these studies would have been useful in determining and justifying the most appropriate model structure for the assessment.

Table 22. Summary of cost-effectiveness model structures identified in the company literature review.

Study	Year	Intervention / comparators	Model type (e.g., Markov)	Modelled health states	ERG interpretation of relevance to decision problem
Badiani[17]	2018	UPA + surgery vs. Placebo + surgery (PEARL I)	Markov	• controlled bleeding • Uncontrolled bleeding • surgery	• Comparison not directly relevant, • useful for model structure
Choi [20]	2016	UPA vs. GnRH agonist prior to surgery [A]	Markov	NR	• Comparison not directly relevant • Insufficient detail on model structure
Lorenzovici[21]	2014	UPA vs. monitoring and UPA vs. hysterectomy [A]	Markov	• Mild excessive bleeding • Moderate excessive bleeding • Severe or persistent excessive bleeding • Myomectomy • Post myomectomy (mild-moderate bleed), • Post myomectomy (severe bleeding) • Hysterectomy • Post hysterectomy • Post menopause • Death	• Comparison not directly relevant, • Limited information potentially useful for model structure

Study	Year	Intervention / comparators	Model type (e.g., Markov)	Modelled health states	ERG interpretation of relevance to decision problem
Maratea[22]	2016	Repeated UPA vs. one-of pre surgical UPA	Simulation model	Not applicable	• Comparison not directly relevant, • Unlikely to be useful for model structure
Nagy	2012 [A18] and 2014[23]	UPA vs. monitoring and UPA vs. hysterectomy	Markov	• Mild or moderate excessive bleeding • Severe or persistent excessive bleeding • Myomectomy • Post myomectomy (mild-moderate bleed) • Post myomectomy (severe bleeding) • Hysterectomy • Post hysterectomy • Post menopause • Death	• Comparison not directly relevant, • useful for model structure
Paladio-Hernandez[24]	2015	UPA vs. GnRHa [A]	Decision tree	NR	• Comparison not directly relevant, • Unlikely to be useful for model structure
Paquete[25]	2016	UPA vs. Surgery [A]	Markov	Unclear (possibly states of on/ off treatment)	• Comparison not directly relevant, • Limited information potentially useful for model structure

Study	Year	Intervention / comparators	Model type (e.g., Markov)	Modelled health states	ERG interpretation of relevance to decision problem
Tsoi[26]	2015	UPA vs. GnRHa	Decision tree	Decision tree branches for: • Controlled bleeding (with / without hot flushes) • Uncontrolled bleeding (with / without hot flushes)	• Comparison not directly relevant, • Partially useful for model structure
Geale[19]	2017	UPA + surgery vs. BSC + surgery	Markov	Treatment states: • UPA • BSC • Surgery • Post-surgery • Death Health state utilities defined according to bleeding and pain outcomes.	• Comparison not directly relevant, • Useful for model structure

Abbreviations: BSC: best supportive care; GnRHa: Gonadotropin-releasing hormone analogue NR: Not reported; UPA: Ulipristal acetate

[A] Abstract only, limited details available.

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

4.2.1 NICE reference case checklist

Table 23 below provides the ERG assessment of the company submission against the NICE reference case.

Table 23 NICE reference case checklist

Element of health technology assessment	Reference case	ERG comment on company's submission
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	**Partly.** A two-step approach was followed to derive utilities for the treatment states in the model, where a) UFS-QoL data from LIBERTY were mapped to EQ-5D, and b) an OLS model was used to derive a utility function describing the impact of one-unit changes in MBL on mapped utilities. It is unclear to what extent MBL captures all direct health effects.
Perspective on costs	NHS and PSS	**Yes**. An NHS perspective has been adopted.
Type of economic evaluation	Cost-utility analysis with fully incremental analysis	**Yes**.
Time horizon	Long enough to reflect important differences in costs or outcomes between the technologies being compared	**Yes**. For an average age of 42, a life-time horizon is modelled in the base case with a scenario analysis to the average age of menopause (age 51). Note that the current model configuration would not allow sensitivity analyses on starting ages < 42 to reflect a life-time horizon.

Element of health technology assessment	Reference case	ERG comment on company's submission
Synthesis of evidence on health effects	Based on a systematic review	**Partly.** Whilst a systematic review was undertaken, the resultant indirect treatment comparisons for evidence of health effects between relugolix CT and GnRHa were limited to one outcome only (MBL). The company used results from the ITC to derive mean MBL for each treatment arm in the model but did not report these results (including measures of uncertainty around the treatment effects that could be incorporated into the PSA).
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.	**Partly.** Health effects were measured in QALYs. Whilst EQ-5D data were available from the LIBERTY study, indirect mapping and regression of MBL on QoL were used in the model because of a lack of sensitivity to measure the impact of patient symptoms on QoL, given inappropriate timing of questionnaires and a single day EQ-5D recall.
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	**Yes.** UFS-QoL data were reported directly from the LIBERTY trial, but the mapped values have not been reported. The ERG would have appreciated seeing the incremental effect of randomised treatment on the mapped utilities.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	**Yes.**

Element of health technology assessment	Reference case	ERG comment on company's submission
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	**Yes.**
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	**Partly.** Resource use required for routine monitoring was based on clinical expert input, but the ERG considers the resource use requirement to be an over-estimate of routine monitoring in UK clinical practice.
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	**Yes.** Though the ERG notes that the discount rate was not varied in sensitivity analyses.
PSS, personal social services; QALYs, quality-adjusted life years; EQ-5D, standardised instrument for use as a measure of health outcome.		

4.2.2 Model structure

The company has submitted a Markov cohort model developed in Microsoft® Excel to determine the cost-effectiveness of relugolix CT compared to GnRHa for the treatment of moderate to severe symptomatic fibroids in adults. The model captures the cost and QALY implications associated with the cohort's transition through a set of mutually exclusive "treatment" states in monthly cycles over a life-time horizon, informed by treatment discontinuation assumptions. The cohort enters the model in the "on treatment" relugolix CT or GnRHa health states. The cohort can then remain on treatment in any given cycle or can discontinue where they immediately enter the BSC state or can be scheduled for surgery. Once a treatment has been discontinued, a second course of pharmacological treatment is not allowed within the model structure. The proportion of those discontinuing that are scheduled for surgery immediately enter the "waiting time" state of assumed duration 15 months before progressing to the surgery state. The waiting time state is essentially an extension of the BSC state

where patients remain off active treatment whilst waiting for surgery. Entry to the 'waiting time' state, and hence scheduling for surgery is therefore modelled to be conditional on treatment discontinuation in the company base case analysis.

The surgery state is a tunnel state that patients remain in for one cycle. This state includes different types of surgery which are each explicitly modelled to describe the distribution of patients currently undergoing surgery by surgery type and to allow correct application of surgery related mortality risks and adverse events. Following surgery, patients move to a post-surgery state that is divided in two – reflecting patients who received hysterectomies and those who did not. Patients who did not receive hysterectomies can then transition to a second surgery state following the completion of further waiting time. For all women, resolution of fibroid symptoms is assumed to have occurred by the point of menopause (age 51), where the cohort all enter the 'menopause' state of the model and receive general population utilities and all-cause mortality risks. The cohort can enter the death state from any other model state according to age and sex-adjusted general population mortality risks. There is an added mortality risk applied from the surgery state to reflect a small additional risk of surgical mortality. The model structure is re-produced from the company submission in Figure 3 below.

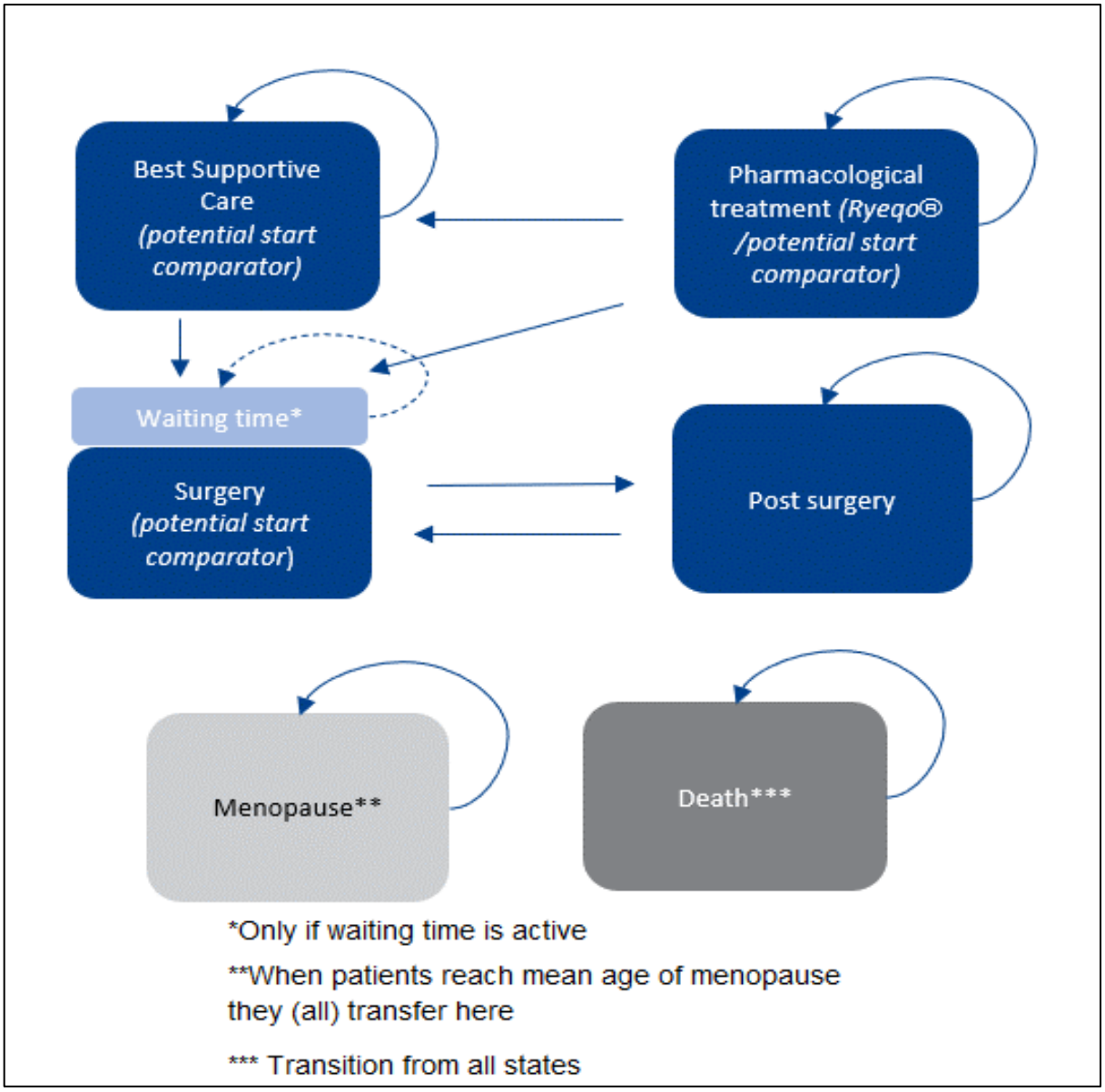


Figure 3 Model structure [reproduced from Figure 29, Document B of the CS].

Decision to model 'treatment' rather than 'health' states

The model structure is built to reflect the treatment pathways that might be experienced in clinical practice, with Markov states defined according to treatment received at any given time point "on-treatment: relugolix CT / GnRHa, off-treatment: BSC, waiting for surgery and surgery.

The ERG does not consider the company's decision to model "treatment" states rather than states defined by "health" outcomes to be sufficiently explained or justified in the submission. The ERG would have considered a more appropriate

model structure to be one, like that of Nagy et al 2014, where the cohort transition through a series of mutually exclusive health outcomes states.[23] Such states might be defined according to bleeding symptoms: such as 'mild', 'moderate' or 'severe' bleeding, or symptom control: 'uncontrolled', 'controlled'. In model states defined by 'health' outcomes, a proportion of the cohort in those states could be modelled to be 'on' or 'off' treatment, according to available treatment discontinuation data from LIBERTY for relugolix CT and from PEARL II / clinical expert opinion for GnRHa. Such an approach would have two key advantages, especially in the pre-surgical states, namely:

1) *MBL effectiveness data from the LIBERTY and PEARL II studies could be linked directly to treatment received, as opposed to the company's approach which applies intention to treat effectiveness (i.e., MBL) data to an 'on treatment' cohort. The approach likely under-estimates MBL and QALY gains in the 'on treatment' proportion of the cohort in both model arms. The overall direction of any bias though is unclear, and dependent on other modelling assumptions. The implications for the effectiveness and QALY gains are discussed in Section 4.2.7.*

2) *Resource use requirements in terms of patient management, investigations, examinations, and follow-up are linked to treatment received in the model. The ERG's clinical expert considers this to be inappropriate, because, in clinical practice, decisions about patient management are more likely to be based on clinical need, which is determined by whether a patient's symptoms are adequately controlled and not necessarily depending on whether they are 'on' or 'off' treatment. The ERG considers that a model based on 'health' states would enable application of more appropriate monitoring and symptom management assumptions. The implications for resource use and costs are discussed in Section 4.2.8.*

Pre-surgery waiting time state

All patients scheduled for surgery first enter a "waiting time" state, of the assumed duration of 15 months. Entry to the waiting time, and hence surgery health states is conditional on treatment discontinuation. The cohort is assumed to only enter the

'waiting time' state if the transition would occur before age 46, on the assumption that patients would not be listed for surgery within five years of menopause (age 51).

The ERG's concern with the company's approach is that, in clinical practice, patients would be unlikely to discontinue treatment before being listed for surgery. Similarly, patients would be unlikely to discontinue treatment whilst waiting for surgery. Indeed, the ERG's clinical expert advisor is of the view that it is advantageous for people to remain on treatment in preparation for surgery to ensure maximum fibroid shrinkage at the point of surgery to improve the chance of surgery success, and potentially even enabling surgeons to conduct surgery via less invasive routes. The ERG, therefore, does not consider the structural assumption to be appropriate, or evidence based.

The implication of the 'waiting time' state is to delay the time point of transition to the surgery state in both arms of the model. However, the combination of three modelling assumptions: A) that listing for surgery is conditional on treatment discontinuation and B) that the cohort can only be listed for surgery between the ages of 41 and 46 and C) given that the relugolix CT treated cohort remains on treatment for longer than GnRHa means that the impact of removing the waiting time state has a much greater relative impact on the relugolix CT arm of the model than the GnRHa arm. Removal of the 'waiting time' state, therefore, leads to a substantial increase in the ICER. The implications for state transitions are discussed in Section 4.2.6.

Surgery model states

The ERG notes that the premise of the company's value case is that longer treatment duration with relugolix CT can maintain adequate response for longer than GnRHa, thus preventing the need for surgery by allowing women to reach the age of menopause where symptoms tend to resolve naturally.

Whilst the ERG accepts that longer duration of a successful medical treatment may lead to some reduction in the need for surgery, there are no data presented by the company to indicate the magnitude of surgery reduction that might be achievable for relugolix CT compared to GnRHa. The company has provided scenario analyses removing the surgery states from the model. The ERG considers that this scenario may reflect the cost-effectiveness of relugolix CT versus GnRHa for the treatment of

fibroids solely in a group of women who will not receive surgery. The scenario analysis would also represent a conservative approach where no differences in surgery outcome would be achieved and based on the assumption that a decision to have surgery is based predominantly on patient preference, rather than whether medical treatment was discontinued or not. On balance, the ERG considers that some effect on surgery may be plausible, particularly in women who do not continue long-term off-licence use of GnRHa, but the magnitude of any effect on surgery reduction is unclear, not evidence-based, and highly uncertain given the available data to inform these transitions.

4.2.3 Population

The company state that their modelled population is informed by the pooled patient characteristics in the LIBERTY 1 and LIBERTY 2 studies of relugolix CT. This results in a model cohort starting age of 42. The company states that the modelled population is reflective of how relugolix CT would be used in UK clinical practice.

The ERG can confirm that the starting age of the model cohort is consistent with that of the pooled LIBERTY study populations. Most characteristics of the LIBERTY study appear to be a reasonable reflection of the population in which relugolix CT might be used in clinical practice, with two exceptions. The first is that the ERG's clinical expert confirms that the model starting age is appropriate but that some women may start treatment at a younger age, especially those who have had their families. There is likely to be substantial variability among the characteristics of the treated population in clinical practice. The second concern relates to treatment goals, and the role of surgery. The ERG notes that the goal of treatment in the LIBERTY study (relugolix CT) is substantially different from the goal of treatment in the PEARL II study used to inform the model comparator (GnRHa). Participants in the LIBERTY studies were not intended to receive surgery and indeed planned surgery was a trial exclusion criterion. In contrast, participants in the PEARL II study were all listed for surgery at baseline. The ERG's view is that the study populations are not comparable with respect to the role of surgery in the treatment pathway. The ERG is concerned that mixing the trial populations to parameterise the economic model without adequate consideration of these different treatment goals is an important limitation of the company's approach. Given that transitions to the surgery states are an important

driver of the ICER, the ERG would have considered it appropriate to model two groups of patients separately, according to their desire to have surgery:

A) *Group A: women who are listed for surgery who receive medical treatment to ensure maximum fibroid shrinkage pre-surgery to improve surgical outcomes (consistent with the population enrolled in the PEARL II study) and*
B) *Group B: women who do not wish or cannot receive surgery, who receive medical treatment to manage fibroid symptoms, such as to reduce blood loss (consistent with the population enrolled in the LIBERTY study).*

This distinction has important implications for the model structure and in particular the role of the surgery model health states. In group A, one could reasonably assume equivalence in transitions to surgery and the decision problem becomes one of cost-minimisation over short term (e.g., 3 months of treatment) prior to surgery. For group B, transitions to surgery may be much lower in both model arms, given that women have already expressed a preference to avoid surgery by initiating treatment with a long-term goal of symptom management.

4.2.4 Interventions and comparators

Intervention: relugolix CT

The intervention under assessment is relugolix CT, containing 40mg relugolix, 1mg estradiol (as hemihydrate) and 0.5mg norethisterone acetate. The drug is self-administered by the patient, orally, as one tablet taken daily. A Dexa-scan is recommended after 52 weeks of treatment to assess bone mineral density and osteoporosis risk. There are no specified treatment stopping rules in the marketing authorisation, other than to recommend cessation of treatment at menopause. The company model treatment to continue indefinitely unless discontinued.

Whilst the clinical community do not have experience of long-term treatment of their patients with relugolix CT, the ERG's clinical expert is satisfied that so long as bone mineral density is monitored through Dexa scans, its modelled usage, which is in line with the marketing authorisation, broadly reflects how relugolix CT would be intended for use in clinical practice, though as noted in Section 4.2.3, some patients

may receive treatment in preparation for surgery, which does not appear to be incorporated in the current model, given the additional 'waiting time' state. There are no modelled stopping rules, other than menopause (age 51), where all treatment is stopped on the assumption that fibroids will shrink without treatment at this point. The modelled cessation of treatment is also in line with the marketing authorisation for relugolix CT and its likely use in clinical practice.

Comparator: GnRHa

The company considers GnRHa to be the most appropriate comparator for use in the model. Six different types of GnRHa are included, based on the treatments currently licensed for use in the UK (goserelin, leuprorelin acetate, and triptorelin) as short acting monthly and long-acting 3 monthly formulations. GnRHa may be used within their licence for the treatment of moderate to severe fibroids up to 3-6 months, but the company's clinical expert opinion is that they are often used off-licence for longer in clinical practice, especially where there is a need to delay or avoid surgery. Long term use requires the addition of add-back HRT to reduce BMD loss.

The ERG agrees that GnRHa are an appropriate comparator for the cost-effectiveness model, as they are the most commonly used medical treatments in UK clinical practice in this setting. Other available medical treatments, such as those included in the NICE scope, target symptom management rather than the underlying fibroids. The ERG's clinical expert is also in agreement that longer term usage of GnRHa is common in clinical practice but given that its usage beyond six months is off-label, duration of treatment in UK clinical practice is likely to vary substantially. The ERG agrees with the company that all GnRHa would likely have similar effectiveness.[27] The ERG therefore considers it appropriate to select the GnRHa with the lowest treatment acquisition costs for calculation of the ICER because all other GnRHa will be dominated (less costly and of equal effectiveness) and thus excluded from the fully incremental analysis. Goserelin monthly has the lowest treatment acquisition cost, and the ERG considers this the most appropriate comparator against which to compare relugolix CT.

Whilst other treatments from the NICE scope have not been included directly as comparators, the ERG's clinical expert is of the view that their role in symptom

management may have an important role to play in best supportive care following treatment discontinuation. This is further addressed in Section 4.2.8.

4.2.5 Perspective, time horizon and discounting

The company submission used an NHS and personal social services (PSS) perspective for costs. The economic model includes functionality that would enable exploration of wider productivity and non-healthcare costs, but these have not been included in the assessment.

The ERG is satisfied that the costing perspective meets the requirements of the NICE reference case.[28]

The model time horizon runs for a maximum of 719 monthly cycles, up to a maximum age of 102 for a starting cohort of age 42. A shorter time horizon, of 9 years, from start age to an assumed average menopause age of 51 is explored in sensitivity analyses, after which point the incremental benefits of treatment are less clear.

In the case of this assessment, a time horizon up to the point of menopause may be sufficient to capture all the costs and benefits of treatment and could be considered as a scenario analysis. The ERG's clinical expert is of the view that post menopause, any incremental benefits of treatment would be difficult to measure with accuracy and the majority of additional health service resource use and quality of life benefit will be accrued prior to menopause. The ERG cautions that any amendments to the model starting age to explore, for example, treatment in younger age groups would not reflect a full lifetime horizon in the current model framework.

Costs and QALYs were discounted by 3.5% per annum in the model.

The ERG is satisfied that discounting has been correctly implemented within the company's economic model and that the base case discount rate applied is in accordance with the NICE reference case.[28] *However, the company has not provided the recommended scenario analyses that vary the discount rate between 0% and 6%*

for both costs and QALYs. The ERG provides scenario analyses that illustrate the impact of different discount rates on the ICER in Section 6.2.

4.2.6 Treatment effectiveness and extrapolation

The following LIBERTY (relugolix CT) and PEARL II (GnRHa) trial data are used to inform the economic model:

A) Treatment discontinuation over time. For relugolix CT, treatment discontinuation is based on the withdrawal rates from the LIBERTY 1-3 studies and the LIBERTY withdrawal study, but with modification to reflect clinical expert opinion that discontinuation in the trials over-estimates discontinuation that might be expected in clinical practice. For GnRHa, data from the PEARL II clinical trial up to three months are supplemented with clinical expert opinion regarding off-licence usage in the longer term.
B) Treatment effectiveness, in terms of menstrual blood loss (MBL) obtained from the LIBERTY studies for relugolix CT and via an indirect treatment comparison (ITC) to the comparator arm of the PEARL II study for GnRH analogues. MBL data are obtained from an ITT analysis of LIBERTY data and applied to an 'on treatment' cohort in the model.
C) Adverse events from LIBERTY 1 and LIBERTY 2 studies and PEARL II studies for relugolix CT and GnRHa respectively, with adverse events beyond trial follow up assumed to equal the rate in the follow up period for the duration of time on treatment.
D) UFS-QoL data mapped to EQ-5D and regressed on MBL to estimate time varying treatment specific health state utility values (See Section 4.2.7).

Summary of model transition probabilities

The ERG note that the company has not directly used transition matrices to govern progression through the model states, with health state occupancy instead determined according to time-varying treatment discontinuation data and assumptions. The ERG has approximated average implied transition matrices from the company base case analysis in Tables 24 and 25 below for relugolix CT and GnRHa respectively. The purpose of this information is to describe the model flow and the differences in health state occupancy over time. Cohort traces are provided in the company submission

Table 24 Summary of approximate transition probabilities among surviving health states (relugolix CT)

	Time (Month)	Treatment	BSC	Waiting for surgery 1	Surgery 1	Post-surgery 1	Waiting for surgery 2	Surgery 2	Post-surgery 2	Menopause
Treatment	Month 1-6	R	0.004	0.0033	-	-	-	-	-	
	Month 7-12	R	[redacted]	[redacted]	-	-	-	-	-	
	Month 13-24	R	[redacted]	[redacted]	-	-	-	-	-	
	Month 24 +	R	[redacted]	[redacted]	-	-	-	-	-	
BSC	All	-	R	0.005	-	-	-	-	-	
Waiting for surgery 1	All	-	-	-	1	-	-	-	-	Age< 51: 0 Age 51+: 1
Surgery 1	All	-	-	-	-	1	-	-	-	
Post-surgery 1	All	-	-	-	-	R	0.0172	-	-	
Waiting for surgery 2	All	-	-	-	-	-	-	1	-	
Surgery 2	All	-	-	-	-	-	-	-	1	
Post-surgery 2	All	-	-	-	-	-	-	-	1	
Menopause	All	-	-	-	-	-	-	-	-	1

Abbreviations: BSC: Best supportive care; R: Remainder

Notes: 1) Proportion transitioning into Surgery state are first on a 15-month waiting list; 2) Everyone transitions into Menopause state aged 51; 3)Patients can enter the Death state from any state according to the general population all-cause mortality; 4) The post-surgery state splits into two sub-states: post-surgery (hysterectomy) and post-surgery (non-hysterectomy), divided according to the proportion having hysterectomy in the model (58.2%); 5) Re-treatment with medical management is not possible. For example, the model does not allow patients to receive GnRHa if relugolix CT is unsuccessful; 6) Patients are not allowed to have more than two surgeries. Once patients enter the Post-surgery 2 state they cannot leave (unless they transition to the Death state) until they reach menopause.

Table 25 Summary of approximate transition probabilities among surviving health states (GnRHa)

	Time (Month)	Treatment	BSC	Waiting for surgery 1	Surgery 1	Post-surgery 1	Waiting for surgery 2	Surgery 2	Post-surgery 2	Menopause
Treatment	Month 1-6	Remainder	0.0105	0.0086	-	-	-	-	-	Age< 51: 0 Age 51+: 1
	Month 7-12	~0.905	~0.052	~0.043	-	-	-	-	-	
	Month 13-60	~0.994	~0.003	~0.003	-	-	-	-	-	
	Month 60+	0.998	0.001	0.001	-	-	-	-	-	
BSC	All	-	Remainder	0.005	-	-	-	-	-	
Waiting for surgery 1	All	-	-	-	1	-	-	-	-	
Surgery 1	All	-	-	-	-	1	-	-	-	
Post-surgery 1	All	-	-	-	-	R	0.0172	-	-	

	Time (Month)	Treatment	BSC	Waiting for surgery 1	Surgery 1	Post-surgery 1	Waiting for surgery 2	Surgery 2	Post-surgery 2	Menopause
Waiting for surgery 2	All	-	-	-	-	-	-	1	-	
Surgery 2	All	-	-	-	-	-	-	-	1	
Post-surgery 2	All	-	-	-	-	-	-	-	1	
Menopause	All	-	-	-	-	-	-	-	-	1

Abbreviations: BSC: Best supportive care; R: Remainder

Notes: 1) Proportion transitioning into Surgery state are first on a 15-month waiting list; 2) Everyone transitions into Menopause state aged 51; 3)Patients can enter the Death state from any state according to the general population all-cause mortality; 4) The post-surgery state splits into two sub-states: post-surgery (hysterectomy) and post-surgery (non-hysterectomy), divided according to the proportion having hysterectomy in the model (58.2%); 5) Re-treatment with medical management is not possible. For example, the model does not allow patients to receive GnRHa if relugolix CT is unsuccessful; 6) Patients are not allowed to have more than two surgeries. Once patients enter the Post-surgery 2 state they cannot leave (unless they transition to the Death state) until they reach menopause.

Treatment discontinuation – relugolix CT

Treatment discontinuation for relugolix CT was obtained from the LIBERTY 1-2 trials (pooled data for months 1-6), LIBERTY 3 study (months 7-12), and the LIBERTY withdrawal study (months 13-24). Clinical expert opinion obtained by the company from N=3 KOLs indicated that the number of patients discontinuing treatment in the LIBERTY studies exceeded what might be expected in UK clinical practice. The company base case model, therefore, assumes that patients discontinuing treatment in the LIBERTY studies for the following reasons would remain on treatment in UK clinical practice.

A) mild (e.g., mood swings) or non-drug-related adverse events,
B) protocol deviations and loss to follow up,
C) most patients that withdrew from the study,
D) some patients that withdrew due to lack of efficacy, given that the MBL measurement used in the trials would not be used in clinical practice and
E) patients withdrawing for several other unspecified reasons

Modified and unmodified discontinuation data are compared in Table 26 below.

Table 26 Relugolix CT modelled treatment discontinuation rates [reproduced from Tables 39 and 40, Document B of the CS].

	LIBERTY 1		LIBERTY 2		LIBERTY 3		LIBERTY withdrawal study	
	Unmodified (ERG preferred)	**Modified (Co. preferred)**	**Unmodified (ERG preferred)**	**Modified (Co. preferred)**	**Unmodified (ERG preferred)**	**Modified (Co. preferred)**	**Unmodified (ERG preferred)**	**Modified (Co. preferred)**
N	128	128	126[A]	125[A]	163	163	115	115
Discontinuation reason								
Adverse event	7	3	2	1	[redacted]	[redacted]	[redacted]	[redacted]
Protocol deviation	1	0	1	0	[redacted]	[redacted]	[redacted]	[redacted]
Lost to follow-up	1	0	4	0	[redacted]	[redacted]	[redacted]	[redacted]
Withdrawal by patient	10	1	13	1	[redacted]	[redacted]	[redacted]	[redacted]
Lack of efficacy	4	4	2	0	[redacted]	[redacted]	[redacted]	[redacted]
Pregnancy	0	0	0	0	[redacted]	[redacted]	[redacted]	[redacted]
Other	5	0	1	0	[redacted]	[redacted]	[redacted]	[redacted]
Total	**28**	**8**	**23**	**2**	[redacted]	[redacted]	[redacted]	[redacted]
% withdrawing	**22%**	**6%**	**18%**	**2%**	[redacted]	[redacted]	[redacted]	[redacted]
Cycle specific probabilities of discontinuation								
Months 1-6 [B]	**4.00%**	**0.72%**	**4.00%**	**0.72%**	-	-	-	-
Months 7-12	-	-	-	-	[redacted]	[redacted]	-	-
Month 13 onwards	-	-	-	-	-	-	[redacted]	[redacted]

[A] The ERG notes that the total number of patients in LIBERTY 2 (n=126) and for the modified withdrawal rates (n=125) do not match. The ERG assumes this is a typo.

[B] Data pooled across LIBERTY 1 and LIBERTY 2 studies.

The ERG notes that the company submission provided insufficient detail and explanation to justify the exact modifications applied to the LIBERTY study data for use in the model. The company mention that cases were reviewed to decide which discontinuers reflected clinical practice, but it is unclear how this was done, whether clinical experts were involved, and if so, how many, and how consensus was achieved. It was also unclear how decisions were reached regarding which discontinuers categorised as 'other' and 'patient withdrawal' were deemed transferable to UK clinical practice.

Whilst the ERG appreciates that some patients discontinuing treatment may do so because of trial processes, it is very difficult to accurately identify which discontinuers are non-generalisable. The ERG prefers the use of unmodified treatment relugolix CT discontinuation rates, as observed in the LIBERTY trials for the following reasons:

A) The ERG's clinical expert sees no strong evidence that the discontinuations are inappropriate for clinical practice. Whilst adverse events may appear mild, patients may still discontinue treatment for these reasons.
B) The data from the LIBERTY studies are the best available evidence on relugolix CT discontinuation over time,
C) GnRHa discontinuations from the PEARL II study were not modified. Modifying discontinuations for relugolix CT but not GnRHa may generate further bias
D) MBL data from the LIBERTY trials reflect the discontinuation as observed in the studies. Adapting the costs, without any corresponding adjustment to treatment benefit is inappropriate.

For all of these reasons, the ERG prefers the use of unmodified treatment discontinuation data.

Treatment discontinuation - GnRH analogues

Treatment discontinuation for GnRH analogues was informed using a combination of data from the PEARL II study and assumptions based on clinical expert opinion as follows:

- **Months 1-3**: Data from the PEARL II study show that, by 13 weeks of follow up, 6/101 (5.9%) of participants discontinued treatment. The company converted this to a monthly probability of treatment discontinuation of 1.91%
- **Months 4-6:** The monthly probability of treatment discontinuation was assumed equal to that observed in the PEARL II study up to week 13 (i.e., 1.91%). A scenario analysis assumed 6-monthly discontinuation rates equal to relugolix CT.
- **Months 7-119**: The company use expert opinion from N=7 KOLs to determine the proportion of patients that would remain on treatment at 1, 5 and 10 years, reflecting that GnRHa may be used off-licence, with add-back HRT beyond the current licence of 6-months. On average, the KOLs predicted that 43.2% (range: 5% to 80%), 13.6% (range: 0% to 55%) and 0.7% (range: 0 to 5%) would remain on treatment at 1, 5 and 10 years respectively. Monthly transition probabilities out of the GnRHa state are calculated using interpolation between these time points.
- **Month 120 onwards**: All patients are assumed to have discontinued treatment.

The ERG was unable to exactly reproduce the probability of discontinuing treatment on GnRHa (1.91%) given that the probabilities are hard coded in the model file rather than showing the underlying calculations. The ERG considers it important to embed all calculations within the model file for transparency to enable reproduction of data. However, the ERG is satisfied that any discrepancies are most likely due to rounding and would only have a negligible impact on the ICER.

The ERG notes that there is substantial variation in the KOL responses regarding long-term off-licence use of GnRHa beyond 6 months (See Table 44 of the company submission). The ERG's clinical expert opinion is that wide variation in UK clinical practice is to be expected, given that the use of GnRHa longer term is off-licence, and that the expert opinion sought by the company likely provides a plausible range. The ERG notes that the longer-term proportion of patients discontinuing treatment has important implications for costs and hence the ICER in the economic model. The ERG does not consider the company's base case approach of including a point estimate of

the mean across clinical experts to adequately reflect this uncertainty. In response to a clarification query (QB3), the company updated the PSA to incorporate uncertainty, assuming a standard error (SE) equal to 10% of the mean. The ERG is not convinced that the approach taken adequately captures the uncertainty, given that a standard error could have been calculated using the available KOL responses. The ERG preferred probabilistic analysis, therefore, incorporates standard errors obtained from the KOL data provided by the company and further deterministic analyses explore using the minimum and maximum values of the ranges provided. The company preferred standard errors are 4.32%, 1.36%, and 0.07% for the proportion on treatment at 1, 5, and 10 years respectively. In contrast, the ERG preferred standard errors are 12.18%, 7.38% and 0.71%.

Treatment discontinuation – Relugolix CT versus GnRH

Treatment discontinuation for relugolix CT and GnRHa under different assumptions is depicted in Figure 4. The company's base case assumes that the modified withdrawal rates from the LIBERTY trials are applied, but the ERG prefers unmodified data as described above. The ERG and company preferred treatment discontinuation assumptions are aligned; however, the graph shows the impact of applying the minimum and maximum proportions remaining on treatment as per the KOL input sought by the company. If GnRHa was used strictly within its licence, then all patients would discontinue at 6 months. The large differences in the areas between the curves illustrate the substantial variation when applying alternative plausible assumptions. The impact of this variation on the ICER is explored by the ERG in further scenario analyses (See chapter 6). The ERG notes that treatments which are discontinued earlier in the model are more likely to be cost-effective. This is likely due to savings in treatment acquisition costs, which are proportionally greater than the reductions in treatment benefit, especially given that the company's base case model assumes costly monitoring for BSC and general population utilities for a successful surgery. Further elaboration from the company regarding the face validity of these findings would be useful.

Figure 4 Treatment discontinuation over time

Transition to BSC or surgery following treatment discontinuation

Transition to surgery is conditional on treatment discontinuation from relugolix CT or GnRHa. The proportion of those who discontinue that immediately transition to surgery (and thus enter the waiting time state) is informed by the PEARL II study which reported that 45.1% of GnRHa patients required surgery at 13 weeks of follow up. The remaining 54.9% had planned surgery cancelled in the PEARL II study but were assumed to transition to BSC in the company's model. Therefore, the company base case assumes that that 45.1% and 54.9% of discontinuers in each model cycle transition into the surgery and BSC states, respectively. The resultant monthly transitions to surgery (i.e., first entering the waiting state) were 0.33%, and [redacted]at ≤6 months, 7-12 months, and 13 months onwards respectively for relugolix CT and 0.86% for GnRHa. The remainder of discontinuers transition to BSC as follows: 0.40%, 0.81%, 0.50% at ≤6 months, 7-12 months, and 13 months onward for relugolix CT, respectively, and 1.05% for GnRHa.

The ERG is concerned that the long-term transitions into the surgery state are not evidence-based and that the use of data from PEARL II is inappropriate. The data from the PEARL II study reflect the proportion of patients (45.1%) in that study who did not have a planned surgery cancelled by week 13 of follow-up. The ERG does not consider these data to be transferrable to the modelled cohort, who did not wish to or were unable to have surgery at the point of medical treatment initiation. The company assumed that this proportion (assumed to have surgery) would be applied to the proportion of women who discontinued pharmacological treatment in each cycle (where discontinuation is informed by withdrawal data from the LIBERTY studies and PEARL II respectively). However, this decision appears to be arbitrarily chosen without any appropriate justification. The ERG considers that the company's approach may therefore substantially over-estimate the proportion of the modelled cohort that enters the surgery states after treatment discontinuation. Furthermore, it is unclear what proportion of people would receive surgery in the relugolix CT arm of the model. It is feasible that it may be less than GnRHa given a longer duration of treatment under the company's base case assumptions. However, any proportion would be hypothetical and not evidence based as rates of surgery were not collected as an outcome from the LIBERTY studies.

The ERG does not consider the company's approach to be plausible. The ERG is concerned that the assumptions used in the model generate results that are inconsistent with the quoted data and it is unclear how accurately they may reflect transition to surgery in UK clinical practice. For example, the model predicts that at 3 months 94% of the GnRHa cohort are on treatment, 3% on BSC, 3% waiting for surgery, with 0% receiving surgery. This contrasts with data from PEARL II where almost 45.1% had surgery immediately after the end of treatment. This mismatch illustrates why it is inappropriate to use data from PEARL II study, from a subgroup who were listed for surgery, to populate the risk of surgery in a different group who were not initially intended to receive surgery. As discussed in Section 4.2.3 above, the ERG queries whether it may be appropriate to consider a separate subgroup analysis in a population of people for whom surgery is intended and medical management is used to prepare for surgery. In this short-term treatment (3 months) scenario, the data from the PEARL II study may be more appropriate to enable a comparison of relugolix CT versus GnRHa. Chapter 6 shows the impact on the ICER of the ERG's exploratory analysis around the potential cost-effectiveness of relugolix CT vs. GnRHa in this setting.

In summary, the ERG accepts that some patients may transition to surgery if symptoms are not controlled whilst on medical treatment. Whilst it is plausible that the proportion would be lower for medical treatments that enable longer treatment duration, the rates of transition to surgery are highly uncertain and the chosen sources for the company's base case analysis are likely to generate an overestimate. The ERG believes that the company should have conducted a more thorough review of the literature to identify rates of surgery in a population for whom surgery was not originally intended. Such data would more closely match the setting in which the company appears to be positioning relugolix CT.

In addition to the immediate transition to surgery (waiting list state) for discontinuers, the model also applies a background risk of transition to surgery from the BSC state. The risk is obtained from the PREMYA study, a cohort of 1139 patients, 142 of whom had previously received UF surgery with an average time to surgery of 26.6 months. This resulted in a monthly transition probability of 0.5%.

The ERG considers the calculation approach applied to be reasonable but are concerned that the application of a further transition to the surgery state from BSC may partially double count some of the surgery transitions following treatment discontinuation. An alternative approach would have been to apply the 0.5% monthly transition to surgery for both treatment discontinuers and those entering surgery from BSC.

Waiting time duration prior to surgery

Once a patient has been listed for surgery, they enter the waiting time state for 15 months prior to receiving surgery. KOL advice sought by the company indicated that considering the covid-19 pandemic waiting times for surgery are significantly longer than pre-pandemic. Five expert opinions were obtained, and the company took the average duration from the 5 responses which ranged from 9 to 18 months.

At the clarification stage, the ERG asked the company to provide an estimated waiting time in a world without covid-19, highlighting that the pandemic and its implications on waiting times would not apply indefinitely. The company referred to their scenario removing waiting time altogether but did not provide an estimate of likely waiting times. The ERG notes that NHS England guarantees an 18-week referral time period for non-urgent treatments. As noted in Section 4.2.2, the ERG prefers the removal of the waiting time state from the model as it does not reflect how patients are managed in clinical practice. However, even if waiting time was considered appropriate, the ERG considers an average waiting time of 5 months to be a more appropriate representation of how services might be delivered in the future.[29]

Surgery outcomes (transitions to the post-surgery states)

Surgical outcomes are dependent on the type of surgery received. The proportion of patients that receives hysterectomy have one surgery only, after which point they are assumed cured. The proportion having other surgeries (myomectomy, UAE, and MRgFUS) may have up to two surgeries. Table 36 of the CS details the pre-surgery rates, obtained from Gupta et al. 2014 and Gorny et al 2017 resulting in a monthly chance of re-surgery of 1.72%.[29, 30] The proportion of patients that are assumed to be cured after having surgery was calculated by converting the annual risk of re-surgery (20.60%) to a 10-year probability (where 10 years is based on the maximum time you

can remain on GnRH treatment). The resulting proportion of patients that were cured after surgery was 12.52% while 87.48% [1-EXP(-monthly rate of 1.73%*120 cycles)] were assumed to have a second surgery.

The ERG considers this to be a substantial overestimate of the risk of re-surgery. In general, the ERG queries the appropriateness of allowing more than one round of surgery given that listing for surgery is only assumed to occur between ages 42 (model start age) and 46 (five years before menopause). As a result, the proportion of the cohort entering the second surgery states is very small and amendments to these parameters have only a negligible impact on the ICER. The ERG explores a scenario where only one round of surgery is allowed within the model structure (i.e., pre-surgery rates assumed = 0%).

Clinical-effectiveness parameters in the model: menstrual blood loss (MBL) volume

MBL volume was the main clinical outcome from the LIBERTY and PEARL II studies used in the economic model and is used to estimate utilities for the relugolix CT, GnRHa, and BSC states of the model. Data are obtained directly from the LIBERTY 1 and LIBERTY 2 trials (up to month 12) for relugolix CT and BSC (placebo arm of LIBERTY studies), and the company's ITC for GnRH analogues. Data from the last MBL measurement time point (week 52 for relugolix CT, week 28 for BSC (placebo), and week 12 for GnRHa) were assumed to be carried forward for the remainder of the patient's time on treatment. Table 56 in the CS reports the time-varying MBL data applied in the model, reproduced graphically in Figure 5. The implications of using MBL data to inform QALY gains are critiqued in Section 4.2.7.

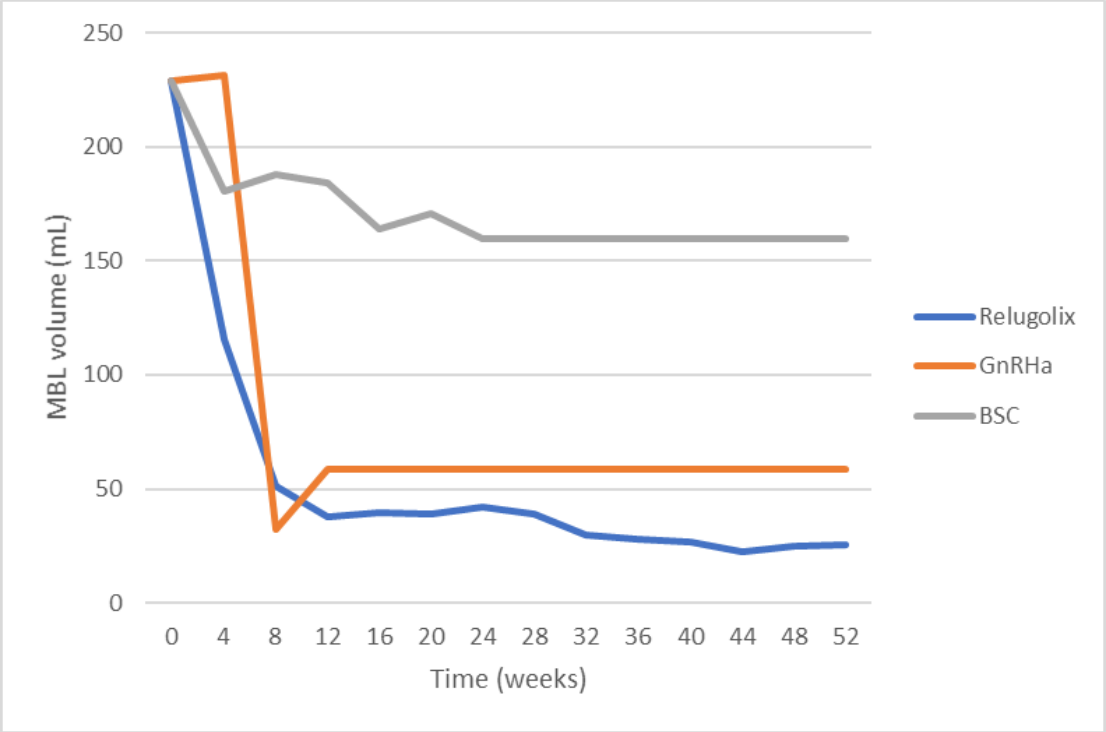


Figure 5 MBL volume over time

The ERG has several concerns regarding the data and assumptions used to integrate MBL data into the economic model:

1) *The ERG's full critique of the company's ITC methodology, and in particular concerns about limited reported data can be found in Sections 3.4 and 3.5 of this report. The points raised in this critique are also relevant considerations for the economic model. The ERG would have preferred to see more clinical outcomes included in the ITC to determine whether the model incorporates sufficient information on patient benefit from which to derive all impacts on quality of life and hence QALY gains.*

2) *Assuming that MBL data are sufficient, the ERG's main concern is that the company did not provide details of the results of the ITC of MBL for relugolix CT versus GnRHa within their submission. The ERG attempted to re-run the company's ITC and was able to generate similar data to those reported in Table 56 of the CS and included in the economic model. The ERG is therefore satisfied that the mean MBL data are indeed sourced from the ITC. However, importantly, no measure of uncertainty surrounding the estimated ITC*

treatment effects was reported or included within the economic model. The ERG's replication of the ITC indicates wide confidence intervals for the comparison of relugolix CT vs GnRHa and hence substantial uncertainty which has not been considered in the economic model. The ERG considers this to be an important omission and one that results in substantial underestimation of the uncertainty surrounding the ICER derived from the company base case probabilistic analysis. The ERG, therefore, uses standard errors derived from our reproduction of the ITC within an updated probabilistic analysis.

3) *Similarly, for BSC MBL, the company include a fixed parameter from what appears to be a pooled analysis of the LIBERTY 1 and* LIBERTY *2 studies. However, again, no estimate of uncertainty has been provided surrounding the pooled MBL treatment effect. The ERG explores applying an approximated standard error obtained from the LIBERTY 1 study, obtained from Table 18 of the CS to incorporate some uncertainty into the probabilistic analysis.*

4) *For the proportion of the cohort in both model arms that discontinue relugolix CT or GnRH analogues and progress to BSC, an immediate increase in MBL is assumed. The ERG's clinical expert considers this to be unreasonable as BSC in clinical practice may still maintain lighter blood loss. It may also take some time for the blood loss levels to revert to placebo following discontinuation. Because a higher proportion of patients come of treatment with GnRH analogues and therefore incur the blood loss levels of someone on BSC (close to pre-GnRHa / relugolix CT levels), any bias would likely be in favour of relugolix CT. Whilst this is an issue of uncertainty that should be considered, the ERG does not have sufficient data to provide a more robust set of assumptions in the model.*

5) *ITT analysis results were used to generate MBL data applied in the economic model. This approach contradicts the model structure which is defined according to treatment received health states. Whilst the ERG would prefer a model structured around 'health' states (see Section 4.2.2), an alternative approach may be to provide a per-protocol analysis of MBL data to apply to*

the 'on treatment' cohort in the current model structure. The implication is that MBL may be underestimated in the 'on treatment' cohorts of both arms of the model. The net impact of any bias depends on the preferred treatment discontinuation assumptions, but in the company's base case analysis any bias is likely to favour GnRHa.

In summary, the ERG would prefer a model structure built on 'health' states that incorporates MBL obtained from the ITC of relugolix CT vs. GnRHa with appropriate standard errors to enable a full assessment of uncertainty in the probabilistic analyses.

Adverse events

Treatment-related adverse events that occurred in ≥5% of patients in the trials (LIBERTY 1-2 and PEARL II) were included in the model (see Table 50 in the CS). The ERG sought further clarification from the company regarding the incorporation of adverse event data in the model. The ERG queries, company clarification and ERG comments on the response are summarised in Table 27.

Whilst there are some uncertainties and the ERG notes that longer-term treatment-related AE data may subsequently become available from the longer-term LIBERTY studies, the ERG is satisfied with the company responses and agree that any impact on the ICER of amending the AEs included in the model is likely to be negligible.

Table 27 Summary of the issues surrounding treatment related adverse events

ERG query	Company response at clarification	ERG comments
Unclear why the company only use adverse event rates from LIBERTY 1-2 and not LIBERTY 3 or the withdrawal studies.	Adverse event data were not available from LIBERTY 3/withdrawal study and therefore could not be used in the economic model.	The ERG is satisfied with the company's clarification response and notes that adverse event rates are not a major driver of the ICER.
Unclear why the company have not included all adverse events in the model and not just those occurring in 5% or more of patients.	1) A total of 35 adverse events were reported in LIBERTY 1 and LIBERTY 2 and including all those adverse events in the economic model would take considerable effort with little impact on the ICER. 2) This would be a biased comparison because of the longer-term data available from LIBERTY compared to GnRHa with 3 months of data from PEARL II. Also, it would be an unfair comparison given the tolerance issues of GnRHa in the longer term. Therefore, extrapolating PEARL II adverse event data for GnRHa and using longer-term adverse event data for relugolix CT might not be appropriate.	The ERG would have preferred to include all adverse events in the trials but appreciates that the impact on the ICER is negligible. The ERG is satisfied that the company's approach to modelling adverse events, whilst not ideal, is sufficient for decision making.

The model also includes adverse events associated with surgery for the proportion of the cohort that enters the surgery health states. Monthly probabilities of short-term surgery-related adverse events were obtained from three studies (Brummer et al. 2011, Manyonda et al. 2012 and Gorny et al. 2011). Details are reported in Table 51 of the CS. The incidence of long-term adverse events related to hysterectomy are reported in Table 52 of the CS.

The ERG noted that no justification was provided for the choice of sources used to obtain short-term adverse events for surgery or their applicability to the modelled population. Whilst further information and a more systematic approach to identifying adverse events would have been preferable, the ERG is satisfied that removing both the short-term and long-term surgery-related adverse events only has a minor impact on the ICER when applied to the company base case analysis. However, in any scenarios in which incremental QALYs are smaller for relugolix CT compared to GnRHa, decisions about these parameters may become more important.

Survival and probability of transition to death state

The company used general population, sex-specific all-cause mortality rates from national life tables to inform transition to the death state in the model. For the proportion of the cohort receiving surgery an excess mortality risk was applied to reflect the risk of surgical mortality. The additional risk was obtained from Settnes et al. 2020, a Danish cohort study.[34] The surgery-specific mortality risks are reported in Table 53 of the CS.

No details were provided regarding how the chosen source was identified or whether a UK source was available. However, the ERG notes that there are minimal incremental life-year gains in the company's base case analysis. Therefore, the impact of surgical mortality only has negligible impact on the ICER.

4.2.7 Health-related quality of life

Section 4.2.6 describes life year gains for relugolix CT in the economic model achieved via lower rates of surgery, and hence a lower overall risk of surgical mortality. However, the predominant driver of quality-adjusted life years (QALYs) within the economic model is through assumed gains in quality of life (utilities) for

relugolix CT compared to GnRHa. There are several routes to utility gain for relugolix CT within the company's economic model:

1) Treatment arm specific health state utility gains associated with lower MBL whilst on treatment with either relugolix CT or GnRHa compared to BSC.
2) Utility decrements due to anxiety and depression associated with being placed on a waiting list for surgery.
3) Disutilities associated with treatment-related adverse events and
4) Gains in utility associated with successful surgery (assumed equal to the general population) offset by utilities equal to the BSC state for unsuccessful surgery, disutilities associated with surgical adverse events, and loss of uterus following hysterectomy applied up to the point of menopause.
5) After menopause (model age 51), the whole cohort receives general population age, but not sex-adjusted utilities.

The ERG's critique of these issues is presented in the following sections.

Treatment state utility values

The model includes treatment-specific utilities that are informed by MBL from the relugolix CT and BSC arms of the LIBERTY studies, and for GnRHa via an ITC with the PEARL II study. Three measures of QoL were included in the LIBERTY studies: EQ-5D-5L, Uterine fibroid symptom and quality of life (UFS-QoL), and patient global assessment (PGA). The LIBERTY studies demonstrated significantly higher improvements in UFS - QoL from baseline for relugolix CT, compared to placebo (BSC), but there was no evidence of any differences between the groups in terms of EQ-5D-5L utilities, with little differences between baseline and follow up in either arm of the trial. The company highlight two concerns that limit the potential for EQ-5D-5L data from the LIBERTY studies to adequately capture QoL benefits of relugolix CT. The first is that EQ-5D data were only collected at baseline and once over follow-up, at 24 weeks. The second is that a recall point of "today" for completing the EQ-5D-5L would likely have failed to capture the QoL implications for women unless the questionnaire happened to be completed during menstruation, which the company state was rare. The company argue therefore that UFS-QoL, which was administered at baseline and twice over follow-up at weeks 12 and 24 and

had a recall time covering the whole follow-up time frame, is a more appropriate measure of QoL.

The ERG agrees that the available EQ-5D-5L data are likely to be insufficient to capture and QoL benefits of treatment. The ERG's main concern with the use of EQ-5D-5L in this context relates to how the instrument was used in the trial, rather than concerns with the instrument's validity per se. The ERG notes that the company could have administered EQ-5D-5L more frequently in their study and could have asked respondents for a mix of responses both during menstruation and at other points in their menstrual cycle. Such an approach would have provided a much richer dataset that would likely have been sufficiently sensitive to measure QALY gains directly in the trial. The company claim that the study visits where EQ-5D-5L was completed rarely occurred during menstruation, but no evidence to support this claim has been provided. Given how EQ-5D-5L was administered in the trial, the ERG generally agrees that UFS-QoL may be a more appropriate measurement tool.

Due to a lack of a valuation tariff for UFS-QoL that would allow estimation of disease-specific QALYs, the company use an unpublished algorithm from Rowen et al to map from the UFS-QoL to EQ-5D-3L.

The ERG is generally satisfied that the underlying mapping process is reasonable, and notes that predicted utilities from the algorithm are generally higher than those of EQ-5D-3L, particularly for more severe health states. The ERG is satisfied that the mapping algorithm may give a conservative estimate of utility decrements associated with uterine fibroids. However, the ERG would have liked to see an estimate of the treatment effect of relugolix CT on mapped EQ-5D values from the LIBERTY studies. This would have helped to validate the company's argument and the ERG opinion that mapped utilities are an appropriate approach to generating treatment state utilities for the economic model.

The company then use a further OLS linear additive regression model as a utility function to predict treatment state utility values based on MBL and baseline age. The model was fitted to the LIBERTY trial data. The following utility function is applied to MBL data for relugolix CT, GnRHa and BSC in the company's base case analysis:

$$EQ - 5D_{mapped} = \alpha + \beta_1 MBL\ Volume + \beta_2\ Baseline\ Age + \varepsilon$$

The resultant utilities at each MBL measurement timepoint for relugolix CT, BSC (placebo) and GnRHa are detailed in Table 57 of the company submission.

The ERG raises several concerns with the company's approach to the estimation of treatment state utility values using their OLS regression.

The first issue is that the company has provided insufficient detail regarding the process of specifying the appropriate utility function, including choices regarding the included explanatory variables and functional form. For example, it is unclear whether non-linearities for age and MBL were explored, for example using squared terms in the OLS model. The company refers to a conference presentation where PBAC bleeding and VAS pain scores were used to directly predict EQ-5D. However, the utility function used in that study is not consistent with the one used in the current submission. The ERG is therefore not satisfied that sufficient information has been provided in the company submission on which to determine the most appropriate utility function

The second issue is that the original company submission did not include standard errors from the OLS regression model, and it was therefore not possible to incorporate the information into the probabilistic analyses. The company raised a concern that SEs from OLS models may be biased due to the repeated measures nature of the UFS-QoL and MBL data. The ERG suggested a repeated measures model at clarification. The company subsequently provided further details from both the OLS and repeated measures models that would enable the incorporation of uncertainty into the probabilistic analyses. The available utility function coefficients and standard errors are compared in Table 28. The ERG notes that the co-efficient on MBL in the repeated measures model is somewhat higher than in the OLS model. However, the most appropriate specification for the utility function remains unclear. In the absence of a full exploration of the advantages and disadvantages of different approaches, the ERG prefers the repeated measures model because it allows more appropriate exploration of uncertainty. The implication of applying the repeated

measures model is that there is a slightly higher reduction in utility for every one-unit increase in MBL compared to the company preferred OLS model. This leads to lower QALYs in both arms of the model, slightly higher incremental QALY gains for relugolix CT and hence a lower ICER compared to the company preferred base case model.

Table 28 Comparison of different utility functions used to populate the economic model

Model parameter	Company base case utility function (OLS)		Company scenario analysis utility function post clarification (repeated measures model)	
	Mean	SE	Mean	SE
Intercept	0.69568	0.02999	0.7035	0.04196
MBL volume (dL)	-0.03877	0.00238	-0.0593 [A]	0.00350 [A]
Age at baseline (Years)	0.00296	0.0007	0.003	0.0001

[A] Note that the numbers reported for MBL volume refer to the company's corrected clarification response (post FAC)

Abbreviations: dL: decilitre; OLS: Ordinary least squares; SE: standard error

Utility in the 'waiting time' and 'surgery' states.

The proportion of the cohort in the 'waiting time' state prior to a first or second surgery are assumed to have the same utilitiy as those on BSC. The justification is that people who are listed for surgery have experienced treatment failure. A further disutility of -0.01 is added to reflect concern or worry among people listed for surgery.

The ERG does not consider the inclusion of a waiting time state to be appropriate (See section 4.2.2.). Even if a 'waiting time' state were included, the ERG disagrees that an additional disutility for anxiety should be applied. The source stated by the company does not reflect a population of people waiting for surgery and is instead a disutility for patients suffering from anxiety. These health states are not comparable.

The ERG sees no evidence that a disutility should be applied during waiting time and an argument could equally be made that people who are listed for surgery may gain positive utility from the anticipation of having a successful resolution of theiry symptoms from surgery.

The utility in the surgery state is calculated as an average of the general population and BSC utilities weighted according to the proportion cured (12.52%) or not cured (87.48%) respectively. A further disutility decrement was then applied to reflect a disutility associated with surgery as outlined in Table 59 of the company submission. A further annual disutility of -0.18 associated with loss of uterus was converted to a monthly disutility and applied in each model cycle to the proportion of the cohort receiving hysterectomy up until the point of menopause.

The ERG considers the company's approach to applying different utilities according to surgical outcome to be reasonable, but as noted in Section 4.2.6, the ERG considers the surgery cure rate to be rather low, and it may be plausible that a larger proportion of the cohort who enter the surgery state may achieve the general population utility than that modelled by the company. The ERG is also concerned that applying multiple disutilities in addition to this may risk double counting. For example, the company has not provided any justification that the disutility of loss of uterus applied in the hysterectomy state is not at least partially captured in the disutilities reported in Sculpher et al (Table 58 and 59 of the company submission).

Furthermore, the ERG notes that the utility function applied to active treatment and BSC underestimates the utility of an age and sex matched UK general population cohort when MBL is low. This means that the incremental QALY gains achieved with progression from active treatment or BSC to a successful surgery, where general population utilities are applied, may be over-estimated. The ERG explores the impact of uncertainty surrounding this assumption in scenario analyses.

Disutilities associated with treatment-related and surgery-related adverse events

Disutilities are also applied to treatment-related adverse events in the model. In response to clarification queries the company provided further detail on the disutility sources applied in the model (see the company's clarification response – B9).

The ERG raised a concern at clarification that the approach used to identify adverse event disutilities did not appear to be systematic. However, following the company's clarification response (B9) providing details of utility measures and value sets applied from the sourced studies, the ERG is now satisfied that the disutilities of adverse applied in the model are reasonable. The ERG also notes that these disutilities are not an important driver of cost-effectiveness results.

UK general population utilities - applied in the menopause state

The company has applied UK general population age-adjusted utility norms, as published in Szende et al., based on the UK-TTO value set.[35] General population utility was used as the starting point for application of all utility decrements incurred after the point of menopause (age 51) and were applied regardless of the experience of surgical procedures, or loss of uterus.

The ERG is satisfied that general population utilities have been appropriately incorporated into the by mode age band. However, the ERG would have considered it more appropriate to use the female-specific general population value set for this population. The ERG accepts however that the impact of changing from the full population value set to a female-specific value set has only a minimal impact on the ICER. The ERG and company preferred value sets are compared in Table 29 below for completeness.

Table 29: General population EQ-5D utility weights used in the model

Age band	Company preferred approach (full population)	ERG preferred approach (female only)	Source
18-24	0.940	0.943	Szende et al.[35]
25-34	0.927	0.925	
35-44	0.911	0.909	
45-54	0.847	0.849	
55-64	0.799	0.815	
65-74	0.779	0.777	
75+	0.726	0.712	

4.2.8 Resources and costs

Treatment acquisition costs (relugolix CT and GnRHa)

The treatment acquisition costs for relugolix CT is £72 at the list price, for a 28-pack of 40 mg/1 mg/0.5 mg tablets. At one tablet per day, this results in a monthly (30.5 days) cycle cost of £78.43.

The treatment acquisition cost for GnRH analogues was obtained from the NHS drug tariff (2021).[36] A total of 4 types of GnRH analogues (across 4 brands) were included. GnRH analogues are given as injections, one injection per month for the short-acting formulations and once every 3 months for the long-acting formulations. The costs are provided in Table 64 and 65 in the CS. The monthly formulations for leuprorelin acetate, triptorelin, and goserelin are priced at £75.24, £72.32 (weighted average of two brands obtained from BNF based on Prescription Cost Analysis data from 2017/18), and £70, respectively. The 3-monthly formulations for leuprorelin acetate, triptorelin, and goserelin are priced at £225.72, £207, and £235, respectively.

The ERG is not convinced that the company's decision to use a weighted average approach is appropriate and would have preferred the use of the lowest available cost for Triptorelin (monthly formulation), which is (£69). However, the impact on the ICER for relugolix CT versus triptorelin is minimal.

Two add-back therapies were included in the analysis. The company assumed 50% would be on tibolone (list price: £7.44; monthly cost: £8.10) and 50% on raloxifene (list price: £5.65; monthly cycle cost: £6.15), based on the BNF (2021). The estimated monthly cost is an average of the two: £7.13.

Although the 50:50 split is an assumption, varying this proportion on each add-back therapy has little impact on the cost-effectiveness results.

Best supportive care treatment costs

For treating persistent symptoms (pain and blood loss), a proportion of patients (informed by the LIBERTY 3 study for relugolix CT and BSC, and PEARL II for GnRH analogues) are assumed to be taking concomitant medications. NSAIDs (200mg ibuprofen) for pain and iron supplements (ferrous sulfate 200mg tablets) for

blood loss. Tables 76-79 in the CS present the concomitant dose assumptions, medication costs, proportion on each type of concomitant medication, and usage mg (per month). The resulting monthly cost of concomitant medication for relugolix CT, GnRHa, and BSC are £3.73, £1.83, and £4.25, respectively.

The ERG has 2 comments related to the concomitant medications used in the economic model:

1) *It is unclear to the ERG why only LIBERTY 3 was used to inform concomitant medication use for relugolix CT and BSC and not the other LIBERTY studies. The company has also provided insufficient information within their submission to clarify exactly what treatments were provided as BSC in LIBERTY 1 and 2, and how these reflect the lack of active treatments included as BSC in the economic model. Ideally, the ERG would like to see evidence that the medications taken in the trials are consistent with those incorporated into the economic model.*

2) *The ERG's clinical expert suggests that treatments following discontinuation of relugolix CT or GnRHa might include hormonal treatments or contraceptives to treat patient's symptoms and manage MBL. However, the company's model assumes no such treatments are included in BSC, including only iron supplements and ibuprofen which the ERG's clinical expert considers to be insufficient for treating fibroids in this patient group. The ERG notes that the company has provided insufficient details within their submission regarding the BSC treatments used in the LIBERTY studies, but it is unlikely that they reflect BSC in UK clinical practice. This raises an uncertainty for decision making. Whilst the ERG considers the costs of BSC to be under-estimated, it is likely that the benefits are also underestimated. Adjusting costs of BSC to better reflect UK clinical practice could be considered as a scenario analysis but doing so would generate further bias because it is unclear how the associated benefits should be adjusted.*

Routine monitoring and examination costs

Other treatment-related costs include an initial gynecologist consultation, GnRHa administration by a nurse, routine monitoring, and examinations. Full details of the

company's monitoring resource use can be found in Tables 68 to 71 of the company submission.

The ERG agrees with the company that a one-off and annual Dexa-scan would be required as part of UK clinical practice use of relugolix CT and GnRHa respectively whilst on treatment to monitor bone mineral density. However, the ERG disagrees with the company's base case assumption that patients receiving BSC would receive annual scans but that they would not have an associated gynaecologist consultation. The ERG also disagrees that patients would routinely receive six-monthly appointments either on or off treatment. The ERG, therefore, considers it more appropriate to assume a one-off gynaecologist consultation and scan, with these usually occurring about 3-4 months after treatment initiation to monitor patient progress and develop a longer-term care plan. Whilst there is uncertainty surrounding the type of imaging that might be used and this is likely to vary across UK clinical practice, the ERG considers one scan to be sufficient, and has applied the company's weighting assumptions as follows: Ultrasound (1/1.45); hysteroscopy (0.25/1.45) and MRI (0.20/1.45). The ERG prefers a scenario where this resource use would be incurred again if there was a major change in the patient's circumstances (for example discontinuing treatment). The ERG, therefore, applies the same resource use after entry to the BSC state in the model. The ERG's clinical expert advice is that patients would be monitored in secondary care and so it is not necessary to include the costs of 3-monthly GP consultation for patients on BSC because they would be managed through a one-off consultation with a gynaecologist instead.

The company's and ERG's preferred base cases on the resource use assumptions, a summary of the ERG's comments on the company's approach, and justification for the ERG's alternative approach are provided in Table 30.

Table 30 Resource use assumptions – admin, routine check-ups, and examinations

Resource category	Company preferred resource use (frequency)			ERG preferred resource use (frequency)			Company obtained source	*ERG comment*
	Relugolix CT	GnRHa	BSC	Relugolix CT	GnRHa	BSC		
Admin costs								
Initial gynecologist visit before starting treatment	Once only	Once only	0	0	0	0	Gynaecology, Non-Admitted Face-to-Face Attendance, Follow-up, Consultant Led, NHS reference costs 2019-20. Currency code: WF01A (NHS England, 2021)[37]	*The ERG considers the company's stated approach to be appropriate. However, the one-off cost of a visit to the gynaecologist to initiate the first treatment with relugolix CT or GnRHa does not seem to be applied in the model. However, because this is applied to both arms before treatment commences, there is no impact on the ICER, and the ERG do not consider this issue further.*
Nurse administration of GnRH agonists	0	Once per treatment	0	0	Once per treatment	0	Calculated as 10 minutes of practice nurse time (Curtis and Burns 2020)[38]	*The ERG considers the company approach to be appropriate.*

Resource category	**Company preferred resource use (frequency)**			**ERG preferred resource use (frequency)**			**Company obtained source**	***ERG comment***
	Relugolix CT	**GnRHa**	**BSC**	**Relugolix CT**	**GnRHa**	**BSC**		
Routine monitoring								
Gynaecologist consultation	6-monthly	6-monthly	None	Once only	Once only	Once only	3 KOLs; Gynaecology consultant Non-Admitted Face-to-Face Attendance, Follow-up (NHS England, 2021)[37]	*The ERG's clinical expert suggests a review once after 3-4 months after starting treatment. A visit to the gynecologist would be triggered if symptoms were not controlled. MBL volume for relugolix CT and GnRHa would suggest symptom control (with regards to blood loss) and therefore regular gynecologist visits may not be required. Therefore, the ERG prefers to assume a one-off visit to the gynaecologist to monitor patient progress and develop a longer-term care plan (applied to relugolix CT, GnRHa and BSC states)*

Resource category	**Company preferred resource use (frequency)**			**ERG preferred resource use (frequency)**			**Company obtained source**	***ERG comment***
	Relugolix CT	**GnRHa**	**BSC**	**Relugolix CT**	**GnRHa**	**BSC**		
GP visits	0	0	3-monthly	0	0	0	3 KOLs; Per surgery consultation lasting 9.22 minutes, PSSRU 2020 (Curtis and Burns 2020)[38]	*The ERG's clinical expert suggests that patients would not have regular 3 monthly visits to the GP. A visit would be triggered only if patients experienced poor symptom control. See comment above.*
Examinations								
DEXA scan	Once after the first year	Annual	0	Once after the first year	Annual	0	3 KOLs; Outpatient DEXA scan, Currency code: RD50Z (NHS England, 2021)[37]	*The ERG considers the company approach to be appropriate. A DEXA scan may also be considered before commencing treatment on both relugolix CT and GnRHa. However, as these would apply to both arms of the model, there is no impact on the ICER.*
Ultrasound	Annual (100%)	Annual (100%)	Annual (100%)	Once *(67%)*	Once *(67%)*	Once *(67%)*	3 KOLs; Transvaginal Ultrasound, Currency code: MA36Z (NHS England, 2021)[37]	*The ERG's clinical expert suggests a scan and review would be conducted after 3-4 months to consider treatment options and long-term plan going forward. The ERG considers one scan*

Resource category	**Company preferred resource use (frequency)**			**ERG preferred resource use (frequency)**			**Company obtained source**	***ERG comment***
	Relugolix CT	**GnRHa**	**BSC**	**Relugolix CT**	**GnRHa**	**BSC**		
Hysteroscopy	Annual (25%)	Annual (25%)	Annual (25%)	Once *(17%)*	Once *(17%)*	Once *(17%)*	3 KOLs; Diagnostic Hysteroscopy, Currency code: MA31Z (NHS England, 2021)[37]	*per patient to be sufficient, weighted according to the company resource assumptions:* *Ultrasound: (1/1.5 = 67%)* *Hysteroscopy: (0.25/1.5 = 17%)* *MRI: (0.25 / 1.5 = 17%)*
MRI[A]	Annual (25%)	Annual (25%)	Annual (25%)	Once *(17%)*	Once *(17%)*	Once *(17%)*	3 KOLs; MRI, Outpatient procedures, Currency code: DIM004 (NHS England, 2021)[37]	
Full blood count	Annual	Annual	Annual	Once only	Once only	Once only	3 KOLs; Haematology, Currency code: DAPS05 (NHS England, 2021)[37]	*ERG's clinical expert does not consider routine investigations for patients that have their symptoms under control. Instead, a review meeting is expected with a gynecologist that would trigger a full blood count measure.*

Abbreviations: BSC: Best supportive care; DEXA: Dual-energy X-ray absorptiometry scan, GnRHa: Gonadotropin-releasing hormone analogue, KOL: key opinion leader, MRI: magnetic resonance imaging

[A] Table 70 in the CS Document B reports 20% having an MRI whereas the model assumes 25%. The ERG assumes the model value is correct.

Surgery-related costs

The proportion of patients receiving each type of surgery was based on HES 2013 data and Carls et al. 2008.[32, 33] The HES 2013 data was used instead of the HES 2019/20 data because it contained more details on the proportions having each type of surgery. It provided the proportion having hysterectomy, abdominal myomectomy, laparoscopic/vaginal myomectomy and UAE. To further disaggregate the data, Carls et al. was used to inform the proportions having the different types of hysterectomy and myomectomy (laparoscopic / vaginal).[33] The distribution of patients having the first and second surgery is reported in Tables 48 and 49 of the company's submission. The surgery-related costs are provided in Table 73 in the CS. All surgery-related costs are a weighted average of elective, day case, and outpatient unit costs.

The ERG is concerned that the proportions having each type of surgery would be out of date considering the older data sources used by the company. The ERG prefers the use of the most up-to-date data sources where possible to inform model inputs however understands that the company have obtained the older HES data to obtain a more granular level of detail. The ERG will conduct a scenario analysis using an alternative source for informing the proportion on each surgery option (Strong et al. 2020). The Strong et al. study includes UK (London) hospital data from 2015-2018. The proportion on each surgery option was (re-weighted according to the surgery options included in the model): abdominal hysterectomy: 2% (company: 43.36%), laparoscopic hysterectomy: 27% (company: 6.36%), vaginal hysterectomy: 0% (company: 8.48%), abdominal myectomy 27% (company: 8.51%), laparoscopic myectomy: 43% (company: 8.24%), vaginal myectomy: 0% (company: 17.23%), UAE: 0% (company: 4.82%) and MRgFUS: 0% (company: 3%).

Adverse event-related costs

Treatment-related adverse event unit costs are reported in Table 74 in the CS. The unit costs are obtained from PSSRU 2020 (GP appointment)[38] and BNF 2021 (Metoclopramide).[39]

The company submitted table presents unrelated treatment-related adverse events that do not match the treatment-related adverse events in Table 50 of the CS (e.g., it included acne and anxiety as treatment-related adverse events). On further inspection of the model file, these unrelated adverse events were not applied. The costs associated with treatment-related

adverse events are listed below which the ERG has obtained from the company submitted model file.

Surgery-related adverse event unit costs are reported in Table 75 of the CS. Different sources were used to inform the surgery related adverse events rates: hysterectomy (Brummer et al. 2011),[40] myomectomy (Manyonda et al. 2012),[41] uterine artery embolization (Mayonda et al. 2012),[41] and MR-guided focused ultrasound (Gorny et al. 2011).[31]

No justification was provided by the company for the chosen HRG codes – some were for non-elective long stay and some for a non-elective short stay. This is not likely a driver of the ICER because of the small probability of having a surgery-related adverse event (see Table 31).

Table 31 Summary of adverse event rates and costs included in the economic model

	Unit cost (£)	Source / HRG code	Rates					
Treatment-related adverse events								
			Relugolix CT	*GnRHa*	*BSC*			
Cough	0	-	0.00%	0.00%	0.73%			
Upper respiratory tract infection	39.23	PSSRU 2020	0.00%	0.00%	0.66%			
Headache	0	-	1.72%	1.92%	2.38%			
Hot flush	0	-	1.44%	7.81%	1.01%			
Anaemia	39.23	PSSRU 2020	0.00%	0.00%	0.93%			
Insomnia	39.23	PSSRU 2020	0.00%	1.20%	0.00%			
Hypertension	39.23	PSSRU 2020	0.81%	0.00%	0.00%			
Nausea	0.97	BNF 2021	0.00%	0.00%	1.07%			
Surgery-related adverse events								
			Abdominal hysterectomy	**Laparoscopic hysterectomy**	**Vaginal hysterectomy**	**Myomectomy**	**UAE**	**MRgFUS**
Bowel obstruction	5 748.41	WH07C, non-elective long stay	0.00%	0.00%	0.00%	1.37%	0.00%	0.00%
Febrile event	2 103.38	WH07D, non-elective short stay	2.50%	1.40%	0.90%	0.00%	0.00%	0.00%
Fibroid expulsion	5 748.41	WH07C, Non-elective long stay	0.00%	0.00%	0.00%	0.00%	1.35%	0.00%
Groin haematoma	2 103.38	WH07D, Non-elective short stay	0.00%	0.00%	0.00%	0.00%	2.70%	0.00%
Haemorrhage	3 640.02	WH07C, Non-elective short stay	8.30%	5.70%	4.40%	1.37%	0.00%	0.00%
Ileus	0	Assumption.	0.00%	0.00%	0.00%	1.37%	0.00%	0.00%
Pelvic infection, haematoma or abscess	2 103.38	WH07D, Non-elective short stay	0.80%	3.20%	2.20%	0.00%	0.00%	0.00%
Pneumonia	2 103.38	WH07D, Non-elective short stay	0.00%	0.00%	0.00%	1.37%	0.00%	0.00%

Post embolisation syndrome	3 640.02	WH07C, Non-elective short stay	0.00%	0.00%	0.00%	0.00%	8.11%	0.00%
Pulmonary embolus	3 640.02	WH07C, Non-elective short stay	0.00%	0.00%	0.00%	1.37%	0.00%	0.00%
Sepsis	5 748.41	WH07A, Non-elective long stay	0.00%	0.00%	0.00%	1.37%	1.35%	0.00%
UTI	2 103.38	WH07D, Non-elective short stay	2.20%	0.70%	1.50%	10.96%	0.00%	0.00%
Urticaria	0	Assumption	0.00%	0.00%	0.00%	0.00%	1.35%	0.80%
Wound infection	2 103.38	WH07D, Non-elective short stay	2.40%	1.50%	0.90%	0.00%	0.00%	0.00%
Abdominal oedema	0	Assumption	0.00%	0.00%	0.00%	0.00%	0.00%	17.70%
Pain	0	Assumption	0.00%	0.00%	0.00%	0.00%	0.00%	3.80%

Abbreviations: BSC: Best supportive care; GnRHa: Gonadotropin-releasing hormone analogue; UAE: Urinary artery embolization; UTI: urinary tract infection.

5 COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results

The company have provided an updated economic model and set of cost-effectiveness analyses in response to clarification queries, correcting a minor error identified by the ERG in the model calculations of life year gains from the 'waiting time' state. All analyses and model results reported in Chapters 5 and 6 therefore refer to the company's updated economic model.

5.1.1 Determinants of cost-effectiveness - QALYs

QALY gains for each treatment arm across model health states are provided in Table 113, appendix K of the company submission. There are two main drivers of QALY gains in the model, as follows:

a) Patients in the relugolix CT arm of the model spend longer on active treatment as opposed to BSC, compared to GnRHa and thus accrue higher QALY gains in the pre-surgical states through lower MBL and higher utilities. It is important to note that the effectiveness assumptions surrounding both, time on treatment and MBL whilst on treatment, are subject to several assumptions and are highly uncertain (see full critique in Section 4.2.7).

b) The model predicts that the number of surgeries in the relugolix CT arm of the model is approximately half of that in the GnHRa arm. Surgery impacts on QALYs by leading to utility gain associated with successful surgery (general population utilities) compared to unsuccessful surgery (BSC utilities), with the ERG noting that different utility calculation approaches may over-estimate the utility gain of surgery success). These gains are offset through the application of a disutility in the surgery waiting state, surgical adverse event disutilities, and disutility associated with loss of uterus up to the point of menopause for patients receiving hysterectomy. There is also a slightly higher overall life-year gain for relugolix CT compared to GnRHa, due to a lower proportion of the cohort incurring a small surgical-related mortality risk.

The ERG considers the differences in the proportion of the model arms progressing to surgery to be highly uncertain and based on strong assumptions about the applicability of the PEARL II trial data to the relugolix CT treated cohort (See full critique in Section 4.2.8).

5.1.2 Determinants of cost-effectiveness - Costs

Table 114, appendix K of the company submission details the drivers of costs in the model. Treatment acquisition costs with relugolix CT are substantially higher than GnRHa, primarily due to the longer time on treatment. The additional cumulative treatment acquisition costs are partially offset by reductions to time spent in the BSC state where routine examination costs are applied. They are also offset by a lower proportion of the cohort entering the surgery states where they incur the costs of the first surgery, costs of surgical complications, and revision surgery up to the age of menopause.

5.1.3 Company deterministic and probabilistic base case ICER

The company's economic model is structured to provide separate results for six different GnRHa products (goserelin, triptorelin, leuprorelin), either monthly or 3-monthly. The company assume that all GnRHa are equally effective, and the ERG considers this assumption to be appropriate. Based on this assumption, the company have provided a fully incremental analysis where the lowest cost GnRHa dominates all other GnRHa treatments in all cases. In the company's analyses, goserelin monthly is the lowest cost comparator, and is, therefore, the most appropriate comparator for the ICER calculation. Fully incremental analyses are provided in the company's submission but based on this assumption, and the ERG's agreement with its validity, the ERG considers a pairwise comparison between relugolix CT and GnRHa to be sufficient for decision making in the context of the quoted list prices for all comparators. The company's preferred base case deterministic and probabilistic ICERs are re-produced in Table 32.

Table 32 Company base case deterministic and probabilistic ICERs [reproduced from Tables 2 and 3 of the company's revised cost-effectiveness analyses in response to clarification queries]

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Company base case analysis (deterministic)							
Goserelin monthly	7,742	21.525	16.530	-	-	-	-
Relugolix CT	9,854	21.525	16.894	2,112	0.000	0.364	5,796
Company base case analysis (probabilistic)							
Goserelin monthly	7,729	--	16.529	-	-	-	-
Relugolix CT	9,850	--	16.894	2,120	--	0.365	5,808

Scatter plots and CEACs from the company base case probabilistic analysis are reproduced in Figures 6 and 7 below.

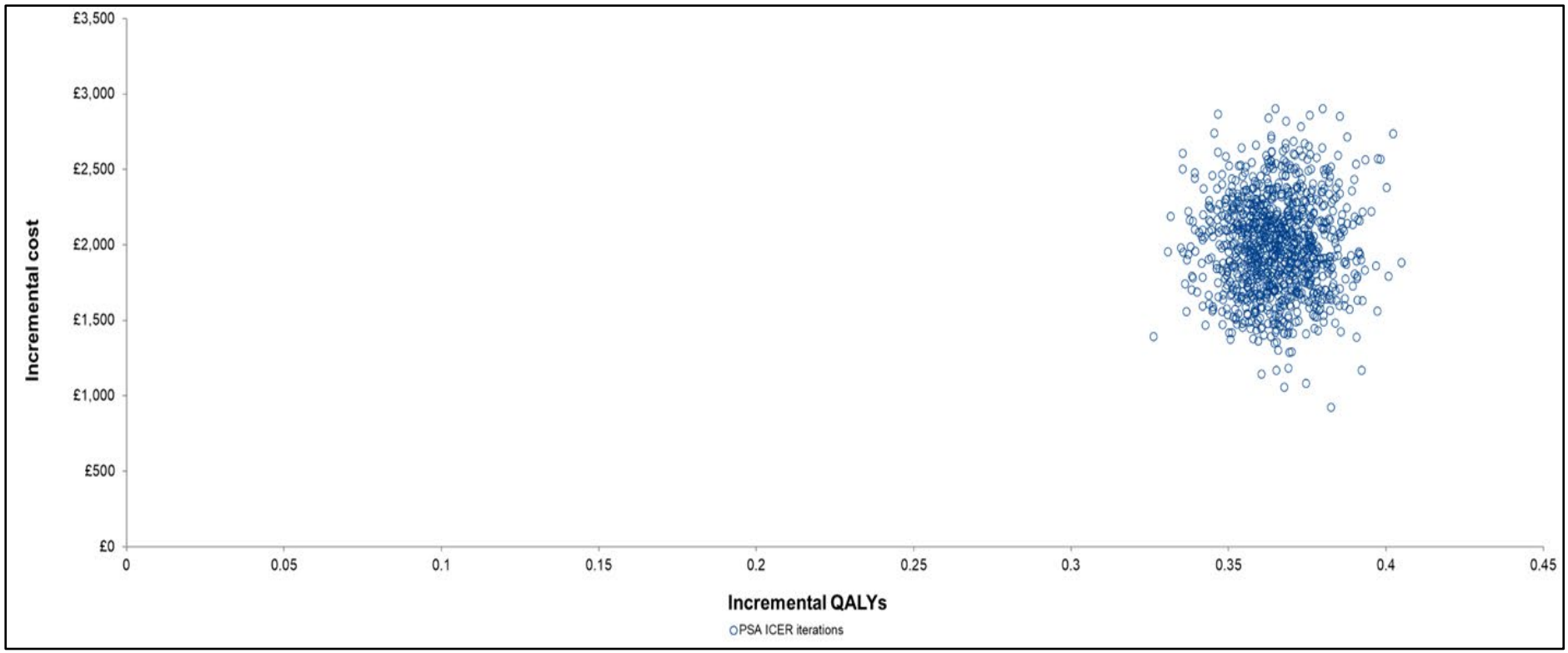


Figure 6 Company probabilistic analysis (scatter plot) - relugolix CT versus GnRHa [reproduced from Figure 1 of the company revised analysis following clarification queries]

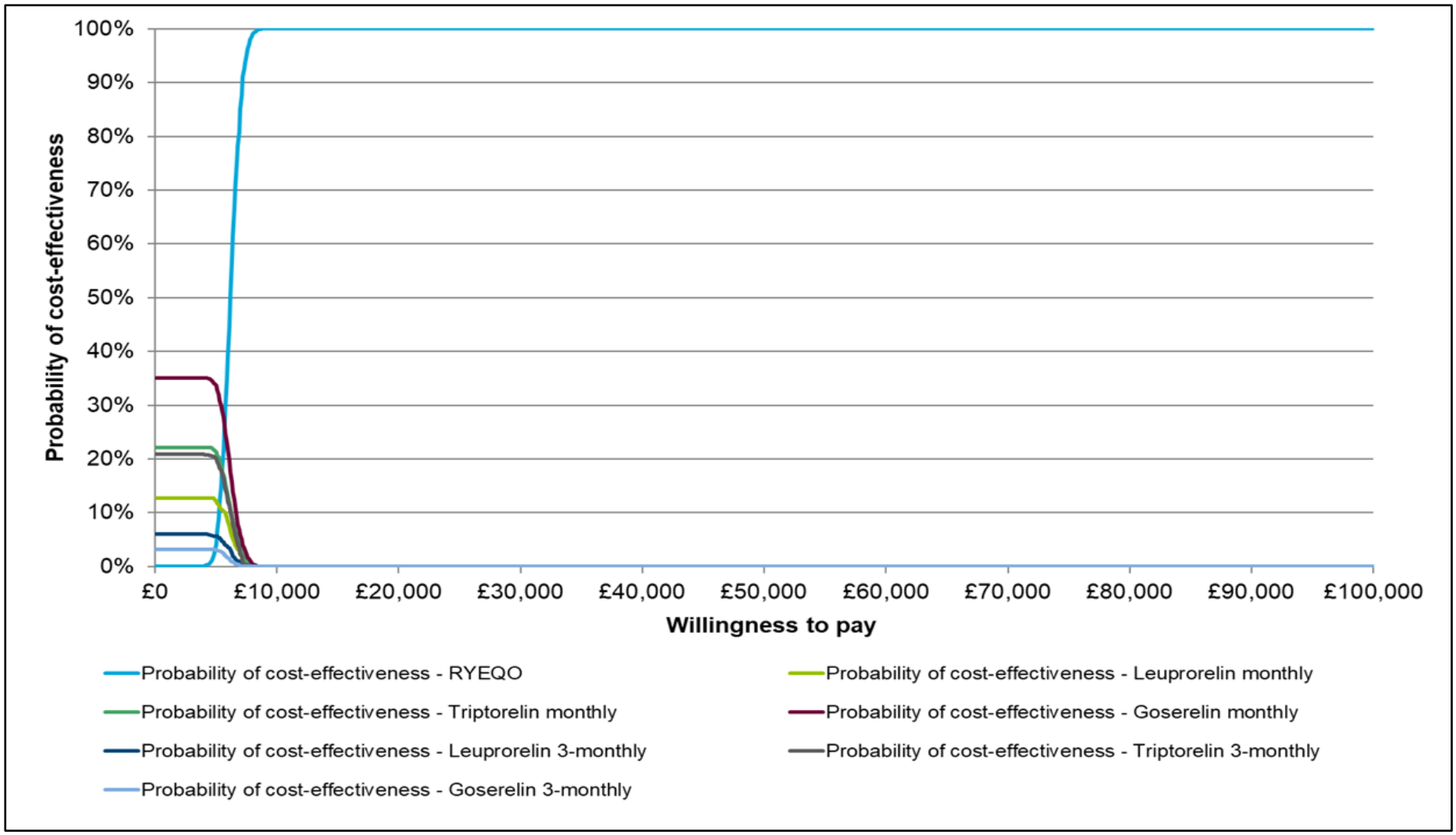


Figure 7 Company probabilistic analysis (CEAC) – all treatments [reproduced from Figure 2 of the company revised analysis following clarification queries]

The ERG has reviewed the company's approach to sampling from distributions for the PSA. The ERG agrees that the company has incorporated multiple parameters within their PSA, and that in general these included parameters are sampled from appropriate distributions (e.g., gamma distributions for costs). However, the ERG has several concerns around the importance of parameters that were not included in the company's PSA.

- *The most important excluded parameters from the PSA are the estimated differences in MBL between relugolix CT and best supportive care (from the LIBERTY trials) and the estimated differences between relugolix CT and GnRHa (from the ITC). The company failed to provide any estimates of uncertainty surrounding these effect sizes for use in the economic model, the ERG was able to recreate the company's ITC, including standard errors around the treatment effect. The ERG also approximates the standard error around the treatment effect from the LIBERTY studies for relugolix CT vs. BSC.*

- *Uncertainty surrounding the regression coefficients used to predict the impact of MBL on EQ-5D (mapped from UFS-QoL) was not incorporated in the original PSA. The ERG notes the company's concern that standard errors from their chosen OLS utility function may be biased because MBL is a repeated measures outcome. Following an ERG request, the company provided the results of a repeated measures model, including standard errors on estimated coefficients for both the OLS and repeated measures utility functions.*

- *Uncertainty surrounding KOL estimates of GnRHa discontinuation beyond six months of treatment was also excluded in the company base case probabilistic analysis. Despite some attempts to integrate this after clarification queries, the ERG still considers the magnitude of uncertainty assumed by the company (SE = 10% of mean) to be underestimated. The ERG prefers to calculate standard errors from available data provided in Table 44 of the CS across 7 KOLs.*

The impact of all these uncertainties on both the company and ERG preferred base case ICERs is illustrated in Section 6.

5.2 Company's sensitivity analyses

The company also provide a tornado diagram illustrating the impact of varying the most important model parameters by +/- 20% of their base case values on the ICER. The results are reproduced in Figure 8 below.

However, as described above, the ERG is of the view that the most important parameter drivers of cost-effectiveness have been excluded from the deterministic scenario analyses also. Therefore, the ERG is of the view that the deterministic scenario analyses and tornado plots substantially under-estimate the overall uncertainty surrounding cost-effectiveness in the company's economic model.

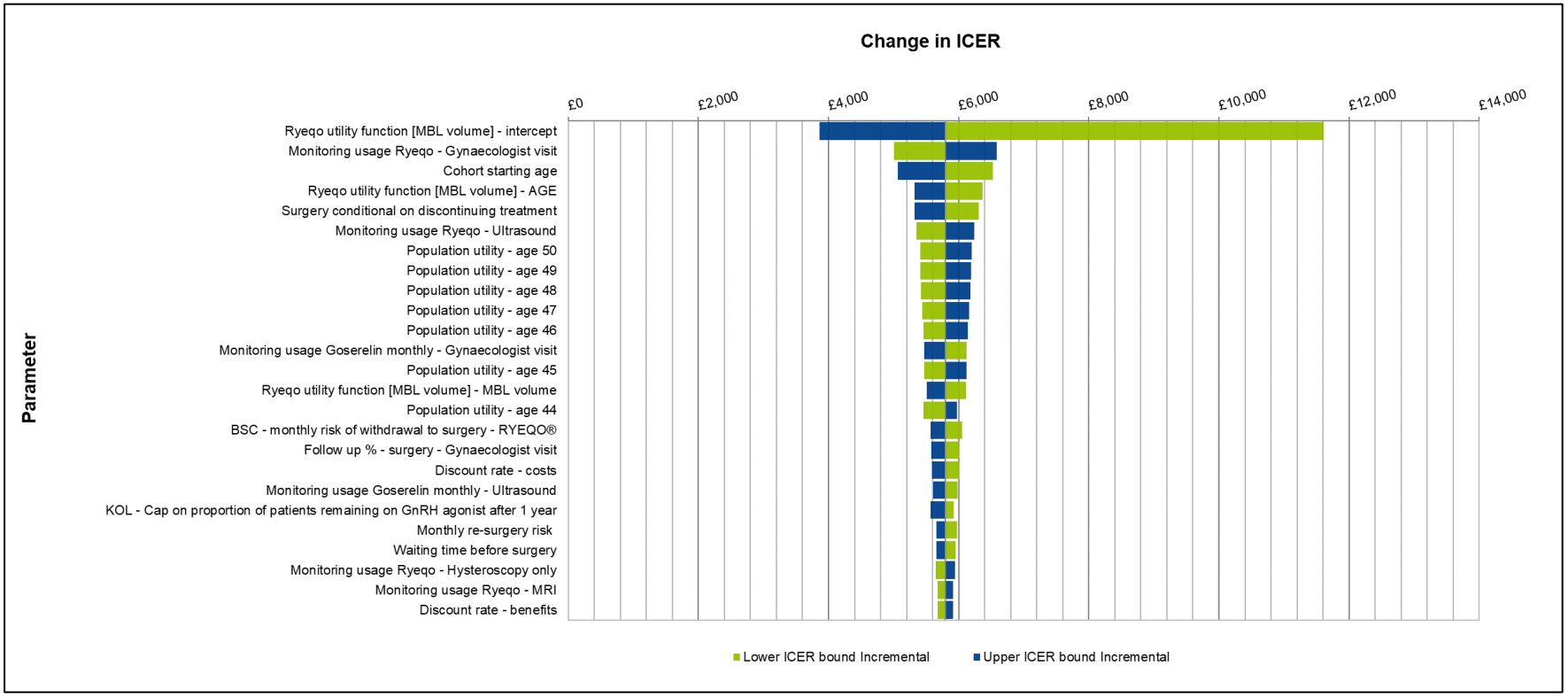


Figure 8 Tornado diagram of one-way sensitivity analyses [reproduced from Figure 3 of the company's clarification response]

The company conducted a total of 14 scenario analyses, varying assumptions around inclusion / exclusion of model states (surgery and waiting time), treatment discontinuation, utilities, and costs. The results of these scenario analyses are re-produced from the company's clarification response in Table 33.

The ERG notes that there is considerable uncertainty surrounding the base case ICER, with the company's one-way scenario analyses generating ICERs up to £15,978 per QALY gained when surgery states are removed from the model. Whilst the ICER under these one-way scenario analyses remains under £20,000 per QALY gained, the ERG notes that plausible optimistic and pessimistic combinations of assumptions and data inputs for relugolix CT would likely demonstrate much wider ICER ranges. The ERG conducts several additional scenario analyses in Chapter 6 to further illustrate the impact of uncertainty surrounding modelling assumptions and data inputs on the base case ICER.

Table 33 Company scenario analyses [reproduced from Table 5 of the company updated analyses following clarification queries]

Structural assumption	Base case scenario	Other scenarios considered	Comparator	Incremental costs	Incremental QALYs	ICER (relugolix CT vs GnRHa)
Base case			**Goserelin monthly**	**£2,112**	**0.364**	**£5,796**
			Triptorelin 3-monthly	**£2,057**	**0.364**	**£5,645**
Modelling of treatment withdrawal in GnRH agonist arm	Withdrawal rates estimated from GnRH agonist arm of PEARL II for the first 6 months and from KOL expert opinion after the first 6 months	Withdrawal for GnRH agonist assumed equal to the modelled withdrawal rates for relugolix CT for the first 6 months of treatment and from KOL expert opinion after the first 6 months	Goserelin monthly	£2,067	0.362	£5,706
			Triptorelin 3-monthly	£2,013	0.362	£5,556
Modelling of adverse events	Adverse events for relugolix CT informed by LIBERTY studies. Adverse events for GnRH agonist informed by PEARL II	Assume identical adverse event profile for relugolix CT and GnRH agonists	Goserelin monthly	£2,116	0.354	£5,982
			Triptorelin 3-monthly	£2,061	0.354	£5,827
MBL volume input for utility algorithm	MBL volume for GnRH agonists derived from ITC	Mean MBL in the GnRH agonist arms assumed the same as relugolix CT for the utility algorithm	Goserelin monthly	£2,112	0.340	£6,212
			Triptorelin 3-monthly	£2,057	0.340	£6,050
			Goserelin monthly	£2,052	0.364	£5,632

Structural assumption	Base case scenario	Other scenarios considered	Comparator	Incremental costs	Incremental QALYs	ICER (relugolix CT vs GnRHa)
Concomitant medication usage	Informed by proportions in LIBERTY 3 for relugolix CT arm and PEARL II for GnRH agonist arm	Assumed equal for relugolix CT and GnRH agonist arms	Triptorelin 3-monthly	£1,995	0.364	£5,475
Induction period of short-acting GnRH agonist required before receiving long-acting GnRH agonist	Yes	No	Goserelin monthly	£2,112	0.364	£5,796
			Triptorelin 3-monthly	£2,177	0.364	£5,974
Duration of short-acting GnRH agonist required before receiving long-acting GnRH agonist	3 months	1 month	Goserelin monthly	£2,112	0.364	£5,796
			Triptorelin 3-monthly	£2,062	0.364	£5,659
Inclusion of surgery health states	Included	Excluded	Goserelin monthly	£3,070	0.194	£15,798
			Triptorelin 3-monthly	£3,016	0.194	£15,516
Referral to surgery upon discontinuation of treatment	No referrals within 5 years of menopause	Referrals possible up until menopause (51 years of age)	Goserelin monthly	£2,203	0.344	£6,403
			Triptorelin 3-monthly	£2,148	0.344	£6,243
Waiting time before surgery	15 months	6 months	Goserelin monthly	£1,993	0.223	£8,947
			Triptorelin 3-monthly	£1,938	0.223	£8,700
Waiting time before surgery	15 months	12 months	Goserelin monthly	£2,099	0.353	£5,954
			Triptorelin 3-monthly	£2,044	0.353	£5,798

Structural assumption	Base case scenario	Other scenarios considered	Comparator	Incremental costs	Incremental QALYs	ICER (relugolix CT vs GnRHa)
GnRH agonist and HRT dose intensity	100%	50%	Goserelin monthly	£3,064	0.364	£8,409
			Triptorelin 3-monthly	£3,036	0.364	£8,331
Add-back therapy costs and effect on AEs for GnRH agonist	Included	Excluded	Goserelin monthly	£2,288	0.380	£6,019
			Triptorelin 3-monthly	£2,233	0.380	£5,875
GnRH agonist treatment duration and inclusion of add-back therapy	Cap on % remaining on treatment at multiple periods based on KOL opinion; add-back therapy included	Fixed maximum duration of 6 months as per SmPC, add-back therapy costs and effect on AEs excluded	Goserelin monthly	£3,362	0.497	£6,766
			Triptorelin 3-monthly	£3,354	0.497	£6,749
GnRH agonist treatment duration (including add-back)	Cap on % remaining on treatment at multiple periods based on KOL opinion	Fixed maximum duration of 12 months; PEARL II withdrawal rates applied throughout	Goserelin monthly	£2,960	0.488	£6,070
			Triptorelin 3-monthly	£2,949	0.488	£6,047

Abbreviations: AE: Adverse events; GnRHa: Gonadotropin-releasing hormone analogue HRT: Hormone replacement therapy; ICER: Incremental cost-effectiveness ratio; ITC: Indirect treatment comparison; KOL: Key opinion leader; QALY: quality adjusted life year

5.3 Model validation and face validity check

The ERG has quality assessed the model against the black-box checklist described by Tappenden and Chilcott 2014[42] and through additional face validity and a random selection of formulae checks in cells on the model trace. The findings of the ERG checks are provided in Table 34. Checks were applied to the company's updated economic model supplied in response to clarification queries, which corrected errors identified in the ERG's initial face validity checks. Those initial errors have been corrected by the company in response to clarification queries and are not discussed further here. The following issues were identified after completion of the updated model face validity check:

Table 34 'Black box' verification checks conducted on the company base case model

Model component	Model test	Unequivocal criterion for verification	Issues identified / ERG comment
Clinical trajectory	Set relative treatment effect (odds ratios, relative risks, or hazard ratios) parameter(s) to 1.0 (including adverse events)	All treatments produce equal estimates of total LYGs and total QALYs	No issues
	Sum expected health state populations at any model time-point (state transition models)	Total probability equals 1.0	No issues
QALY estimation	Set all health utility for living states parameters to 1.0	QALY gains equal LYGs	No issues
	Set QALY discount rate to 0	Discounted QALYs = undiscounted QALYs for all treatments	No issues
	Set QALY discount rate equal to very large number	QALY gain after time 0 tend towards zero	No issues
Cost estimation	Set intervention costs to 0	ICER is reduced*	No issues
	Increase intervention cost	ICER is increased*	No issues

Model component	Model test	Unequivocal criterion for verification	Issues identified / ERG comment
	Set cost discount rate to 0	Discounted costs = undiscounted costs for all treatments	The undiscounted total admin costs for GnRHa was picking up the adverse event costs and not the admin costs as it should (cell Q6 onwards in the 'Totals GnRH1-6' sheets). Correcting this error resulted in the discounted and undiscounted costs to equalize when discount rate is set at 0%. This does not affect any analyses conducted because the discounted values are calculated separately and are the ones used for the calculation of the ICERs.
	Set cost discount rate equal to very large number	Costs after time 0 tend towards zero	No issues
Input parameters	Produce n samples of model parameter m	Range of sampled parameter values does not violate characteristics of statistical distribution used to describe parameter (e.g., samples from beta distribution lie in range 0\x \1, samples from lognormal distribution lie in range x [0, etc.)	No issues
General	Set all treatment-specific parameters equal for all treatment groups	Costs and QALYs equal for all treatments	A minor issue was identified during FAC stage where the model traces did not completely capture all input parameters. This did not affect company analyses but impacted on subsequent ERG scenario analyses (Chapter 6)

Model component	Model test	Unequivocal criterion for verification	Issues identified / ERG comment
	Amend value of each individual model parameter*	ICER is changed	A selection of parameters was amended, and no issues identified. However, because the model parameters (live values in the 'Parameters' sheet of the model file) are not always active, it was cumbersome to identify which cell, for each model parameter, was being used in the model.
	Switch all treatment-specific parameter values*	QALYs and costs for each option should be switched	Partly attempted. ERG managed to get close to a full switch in treatment-specific parameters. Because the model file is not always flexible enough to switch the parameters, e.g. for the different cost inputs for relugolix and GnRHa, it is difficult to switch the parameters when e.g. the cost items are not the same.
ICER incremental cost-effectiveness ratio, LYG life-years gained, QALY quality-adjusted life-year * Note this assumes that the parameter is part of the total cost function and/or total QALY function			

6 EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

6.1 Exploratory and sensitivity analyses undertaken by the ERG

The ERG has undertaken several further exploratory and sensitivity analyses to illustrate the impact of variation in different plausible assumptions on the ICER. Table 35 describes each of the analyses undertaken, together with a justification for each.

Table 35 ERG justification for additional exploratory and sensitivity analysis

Analysis number	Parameter/ Analysis	Company base case assumptions	ERG preferred / exploratory analysis	Justification for ERG's assumption	ERG report section
	Model structure				
1 & 2	Appropriateness of the model 'waiting time' state	Company assumes that listing for surgery is conditional on treatment discontinuation. The proportion of the cohort that discontinue who are listed for surgery therefore enter a waiting time state of average duration 15 months following treatment discontinuation, prior to receiving surgery	**ERG preferred scenario:** Remove waiting time health state. **ERG exploratory scenario:** Reduce duration of waiting time state to 5 months	ERG clinical expert opinion is that patients will remain on treatment whilst waiting for surgery because pre-operative treatment is desirable to ensure optimal surgical outcomes. Exploratory analysis reducing waiting time is intended to reflect likely target waiting times post covid-19 pandemic.	4.2.2
3	Number of potential surgeries in the treatment pathway	Company assumes a maximum of two surgical procedures (one for hysterectomy)	**ERG exploratory analysis:** Assume only one round of surgery would be undertaken.	The ERG analysis explores the impact on the ICER of assuming multiple surgeries would not be conducted close to menopause in UK clinical practice.	4.2.2

Analysis number	Parameter/ Analysis	Company base case assumptions	ERG preferred / exploratory analysis	Justification for ERG's assumption	ERG report section
	Clinical effectiveness & transition probabilities				
4	Treatment discontinuation assumptions on relugolix CT	Applies a modification of treatment discontinuation data from the LIBERTY studies based on clinical expert opinion that the trial over-estimates discontinuation that might be observed in UK clinical practice	**ERG preferred scenario:** Apply unmodified treatment discontinuation rates from the LIBERTY studies	The company's approach is subjective, inconsistent with GnRHa data from the PEARL II comparator, and reduces the costs required to deliver the MBL treatment benefit	4.2.6
5, 6 & 7	Treatment discontinuation for GnRHa	Proportion remaining on treatment based on the <u>average</u> response from N=7 KOLs: 43.2%, 13.6% and 0.7% would remain on treatment at 1, 5 and 10 years respectively	**ERG exploratory analyses:** Varying the proportion on treatment between the minimum and maximum estimates provided by KOLs: range: 5% to 80% at 1 year; 0% to 55% at 5	To explore the impact of this highly uncertain parameter on the ICER	4.2.6

Analysis number	Parameter/ Analysis	Company base case assumptions	ERG preferred / exploratory analysis	Justification for ERG's assumption	ERG report section
			years and 0 to 5% at 10 years. Stopping all GnRHa treatment at 6 months reflecting its use within licence.		
8	Source for surgery rates	Based on the PEARL II study	**ERG exploratory analysis:** Using data from Strong et al.	To illustrate the impact of varying the rate of surgery on the ICER using alternative published sources.	4.2.6
	Utilities				
9	Utility function used to describe the impact of MBL on utility	Company base case uses a linear additive OLS regression model	**ERG preferred scenario:** The ERG prefers to use the repeated measures model provided by the company post – clarification (version corrected post FAC)	Whilst a complete assessment of the advantages and disadvantages of alternative utility functions has not been provided by the company, the ERG view is that the repeated measrues model more closely approximates general population	4.2.7

Analysis number	Parameter/ Analysis	Company base case assumptions	ERG preferred / exploratory analysis	Justification for ERG's assumption	ERG report section
				utility for low MBL and allows estimation of unbiased standard errors for the probabilistic analysis.	
10	Disutility associated with anxiety from waiting for surgery	The company applies a disutility of -0.01 for each month the cohort is in the 'waiting' state to reflect potential anxiety whilst on the waiting list for surgery	**ERG preferred scenario:** To exclude waiting state completely, but even if waiting state is included, the ERG prefers to remove the disutility.	The ERG does not consider the utility source to be generalisable to a population on the waiting list for surgery. Furthermore, there may be positive utility associated with anticipation of a resolution of symptoms.	4.2.7
11	UK general population utility norms	The company apply age-adjusted general population norms for the whole population	**ERG preferred scenario:** General population age and sex-adjusted norms (female)	The ERG considers female-specific general population norms to be more appropriate in the context of the decision problem for this assessment	4.2.7
12	Relugolix CT, GnRHa and BSC utilities	Calculated directly from utility function	**ERG exploratory scenario:** Applied as a decrement to general population	To improve consistency between the application of utilities in treatment states and surgery states	4.2.7

Analysis number	Parameter/ Analysis	Company base case assumptions	ERG preferred / exploratory analysis	Justification for ERG's assumption	ERG report section
	Resource use and costs				
12	Routine monitoring in clinical practice	In addition to dexa-scans, the company include the following routine monitoring resource use: Relugolix CT & GnRHa receive six-monthly gynaecologist consultations and annual scans BSC: receive no gynaecologist consultations, but the same annual scans as the on treatment cohort.	**ERG preferred scenario:** The ERG agrees with the company's modelled use of dexa scans (once only for relugolix CT and annual for GnRHa). The ERG prefers a one-off consultation and scan every time treatment is changed (i.e. 3-4 months after starting relugolix CT / GnRHa and again after treatment discontinuation)	The ERG's preferred assumptions are more likely to reflect patient monitoring in UK clinical practice, where consultations and scans are triggered by patient's symptom control rather than the treatment they receive.	4.2.8
	Scenarios to explore the impact of methodological uncertainty				
13 & 14	Discount rates	Costs: 3.5% per annum QALYs: 3.5% per annum	**ERG exploratory analyses** Discount rate for	Scenario analyses to comply with the NICE reference case[28]	4.2.5

Analysis number	Parameter/ Analysis	Company base case assumptions	ERG preferred / exploratory analysis	Justification for ERG's assumption	ERG report section
			costs and QALYs varied between 0%-6%		
15	Time horizon	Lifetime horizon	**ERG exploratory analysis** Setting maximum time horizon of 9 years (from age 43 to 51).	The justification for this scenario is that all relevant costs and QALYs will most likely have been incurred by menopause (average age of 51).	4.2.5
	Scenarios to explore the impact of treating different subgroups				
16 & 17	Modelled population	Model cohort appears to be structured around the LIBERTY study population, with the intention of using medical treatment to avoid surgery among those who do not wish to have surgery	**Two ERG exploratory analyses:** Removing surgery states from the model to reflect approximate cost-effectiveness of long-term medical management when surgery is not an option.	The ERG provides scenario analyses to help understand the potential drivers of cost-effectiveness in different subgroups. The analyses also seek to illustrate the potential magnitude of bias associated with using surgery rates from the PEARL II study (where surgery was a trial inclusion criterion) to estimate transitions to surgery for relugolix	4.2.3

Analysis number	Parameter/ Analysis	Company base case assumptions	ERG preferred / exploratory analysis	Justification for ERG's assumption	ERG report section
			Apply a 3 month course of treatment for all in preparation for surgery to reflect approximate cost-effectiveness of short term medical management pre-surgery to optimize surgical outcomes (assumes equal effectiveness as per the limited available ITC data).	CT (where being listed for surgery was an exclusion criterion for the LIBERTY studies)	

Abbreviations: ITC: Indirect treatment comparison; MBL: Menstrual blood loss; QALY: Quality adjusted life year;

6.2 Impact on the ICER of additional clinical and economic analyses undertaken by the ERG

As noted in Section 5.1.3, the ERG considers the company's probabilistic analysis (see Figure 6 for the scatter plot of uncertainty around incremental costs and QALYs on the cost-effectiveness plane) to substantially under-estimate uncertainty surrounding the ICER. The ERG has therefore re-run the PSA, incorporating uncertainty surrounding the MBL (obtained from the LIBERTY studies and ITC), uncertainty surrounding the utility function parameters, and incorporation of broader uncertainty surrounding the elicitation of KOL inputs on GnRHa treatment discontinuation. The results of the ERG's preferred probabilistic analysis applied to the company's base case are illustrated in Figure 9 below. Table 36 then provides the results of all the ERG's exploratory analyses applied to the company base case ICER following clarification queries.

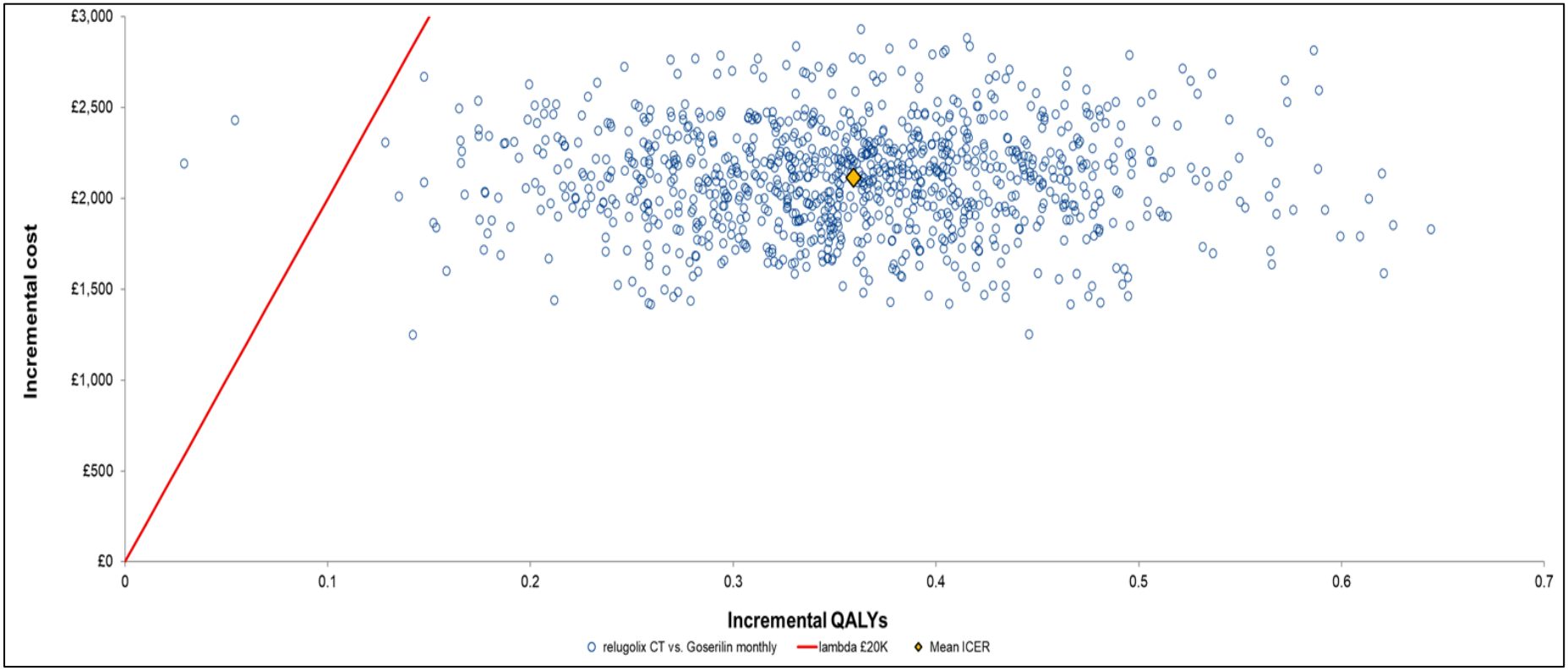


Figure 9 ERG preferred probabilistic analysis of the company's base case model configuration (relugolix CT versus goserelin monthly)

Table 36 ERG scenario analyses applied to the company base case analysis

Analysis number	Treatment	Total costs (£)	Total LYG	Total QALYs	Incremental Costs (£)	Incremental LYG	Incremental QALYs	ICER (£)
Co BC	**Company preferred base case ICER**							
	Goserelin monthly	7,742	21.525	16.530				
	Relugolix CT	9,854	21.525	16.894	2,112	0.000	0.364	5,796
1	**Remove waiting time state prior to surgery (assumes transition directly to surgery following treatment discontinuation)**							
	Goserelin monthly	8,210	21.525	17.013				
	Relugolix CT	10,111	21.525	17.116	1,901	0.000	0.103	18,470
2	**Reduce waiting time to five months**							
	Goserelin monthly	8,037	21.525	16.831				
	Relugolix CT	10,013	21.525	17.031	1,975	0.000	0.200	9,859
3	**Assume one round of surgery only**							
	Goserelin monthly	7,339	21.525	16.712				
	Relugolix CT	9,686	21.525	16.970	2,347	0.000	0.258	9,102
4	**Apply unmodified withdrawal rates for relugolix CT as per the LIBERTY studies**							
	Goserelin monthly	7,742	21.525	16.530				
	Relugolix CT	8,185	21.525	16.633	444	0.000	0.103	4,311
5	**Proportion on GnRHa treatment set to minimum of KOL input**							
	Goserelin monthly	6,775	21.525	16.414				
	Relugolix CT	9,854	21.525	16.894	3,078	0.000	0.480	6,416

Analysis number	Treatment	Total costs (£)	Total LYG	Total QALYs	Incremental Costs (£)	Incremental LYG	Incremental QALYs	ICER (£)
6	**Proportion on GnRHa treatment set to maximum of KOL input**							
	Goserelin monthly	8,781	21.525	16.650				
	Relugolix CT	9,854	21.525	16.894	1,073	0.000	0.244	4,399
7	**GnRHa treatment discontinued at 6-months in line with marketing authorisation**							
	Goserelin monthly	6,491	21.525	16.401				
	Relugolix CT	9,854	21.525	16.894	3,362	0.000	0.493	6,816
8	**Source of surgery risk from Strong et al.**							
	Goserelin monthly	7,808	21.525	16.424				
	Relugolix CT	9,870	21.525	16.851	2,061	0.000	0.426	4,836
9	**Use a utility function based on a repeated measures model to predict the impact of MBL on utilities in the 'on' and 'off' treatment states**							
	Goserelin monthly	7,742	21.525	16.441				
	Relugolix CT	9,854	21.525	16.867	2,112	0.000	0.426	4,953*
10	**Exclude disutility associated with anxiety from the waiting time for surgery state**							
	Goserelin monthly	7,742	21.525	16.536				
	Relugolix CT	9,854	21.525	16.897	2,112	0.000	0.361	5,848
11	**Use female specific UK general population utility norms**							
	Goserelin monthly	7,742	21.525	16.576				
	Relugolix CT	9,854	21.525	16.939	2,112	0.000	0.363	5,818

Analysis number	Treatment	Total costs (£)	Total LYG	Total QALYs	Incremental Costs (£)	Incremental LYG	Incremental QALYs	ICER (£)
12	**Utilities for Relugolix CT, GnRHa and BSC applied as decrements from general population norms.**							
	Goserelin monthly	7,742	21.525	17.506				
	Relugolix CT	9,854	21.525	17.022	1,982	0.000	0.484	4,098
13	**Apply one-off scan and gynaecologist consultation across all treatment states (relugolix CT, GnRHa and BSC)**							
	Goserelin monthly	5,886	21.525	16.530				
	Relugolix CT	6,935	21.525	16.894	1,048	0.000	0.364	2,877
14	**Discount rate 0%**							
	Goserelin monthly	8,752	42.086	32.268				
	Relugolix CT	11,298	42.086	32.672	2,546	0.001	0.404	6,297
15	**Discount rate 6%**							
	Goserelin monthly	7,141	15.113	11.570				
	Relugolix CT	9,001	15.113	11.910	1,861	0.000	0.340	5,469
16	**Time horizon: up to menopause (Age 51)**							
	Goserelin monthly	7,742	7.600	5.525				
	Relugolix CT	9,854	7.600	5.889	2,112	0.000	0.364	5,805
17	**Subgroup: long term use in a group who will not transition to surgery**							
	Goserelin monthly	5,927	21.525	17.061				
	Relugolix CT	8,997	21.525	16.866	3,070	0.000	0.194	15,798
18	**Subgroup: short term use in preparation for surgery**							

Analysis number	Treatment	Total costs (£)	Total LYG	Total QALYs	Incremental Costs (£)	Incremental LYG	Incremental QALYs	ICER (£)
	Goserelin monthly	8,336	21.524	17.242				
	Relugolix CT	8,406	21.524	17.243	70	0.000	0.005	261,701

Abbreviations: ICER: Incremental cost-effectiveness ratio; LYG: Life year gains, QALY: Quality-adjusted life years

* Note: the ERG was not able to reproduce the results in the company scenario analysis with utility parameters estimated from repeated measures model (Table 17 in the clarification response document). The company estimated an ICER of £4,977 while the ERG estimated an ICER of £4953. The discrepancy is likely due to rounding differences in the utility input parameters, but it was not possible to replicate the company's exact analysis as this scenario analysis was not included within the company's submitted Excel model.

6.3 ERG's preferred assumptions

The key differences between the company's and ERG's preferred base case analyses are:

- The company prefers an economic model structure based on 'treatment' states whereas the ERG prefers an economic model structure based on 'health' states, defined according to symptom control. However, the ERG couldn't construct such a model given the available data.

- The company prefers a modelling assumption where women can only be listed for surgery after treatment discontinuation, when they enter a 'waiting time' state of duration 15 months. The ERG considers it more appropriate to remove the waiting time state because, in clinical practice, most women listed for surgery would continue to receive the primary medical treatment in preparation for surgery.

- The company prefers to modify treatment discontinuation data from the LIBERTY study, based on the assumptions of clinical expert opinion that discontinuation in the trial over-estimates discontinuation in real-world clinical practice. The ERG prefers the use of relugolix CT treatment discontinuation data sourced directly from the LIBERTY study because it is more consistent with the costs required to deliver the modelled treatment benefit and also ensures consistency with the data collected in the PEARL II study for GnRH agonists.

- The company uses a mapping algorithm to transform disease-specific quality of life (UFS-QoL) to generic EQ-5D and uses a linear (OLS) utility function to model the impact of MBL on mapped EQ-5D values. The ERG would prefer more details in support of the chosen model structure and how it was derived. Based on the currently available information, the ERG considers data from the repeated measures model provided by the company in response to clarification queries (with reporting error corrected post FAC) to be more appropriate to allow estimation of appropriate standard errors for inclusion in the probabilistic analysis

- The company assumes that all patients (whether on active treatment or BSC) will receive annual examination scans, but only patients on active treatment will receive gynaecologist appointments (6-monthly). The ERG would ideally prefer a model structure that allows follow-up resource use to be linked to the patient's symptom control ('health' states) rather than their 'treatment' received (other than for Dexa- scans). In a 'treatment' state model, the ERG prefers lower resource use: a one-off gynaecologist appointment and scan to make a treatment plan whenever treatment is started or discontinued.

- The company has included the key clinical outcome from the ITC (MBL) as a fixed-point estimate in the economic model, but the ERG prefers full incorporation of uncertainty surrounding the treatment effects for relugolix CT vs. GnRH agonists and relugolix CT vs. BSC into the probabilistic analyses.

The individual impact of all the ERG's preferred scenarios has been described in Table 36 above. The cumulative impact of the ERG's preferred assumptions on the base case ICER is illustrated in Table 37 below.

Table 37 ERG's preferred model assumptions

	Total costs (£)	Total LYG	Total QALYs	Incremental Costs (£)	Incremental LYG	Incremental QALYs	ICER (£)
Company preferred base case ICER							
Goserelin monthly	7,742	21.525	16.530				
Relugolix CT	9,854	21.525	16.894	2112	0.000	0.364	5,796
+ Apply GnRHa trace correction							
Goserelin monthly	7,742	21.525	16.530				
Relugolix CT	9,854	21.525	16.894	2112	0.000	0.364	5,796
+ Apply unmodified withdrawal rates							
Goserelin monthly	7,742	21.525	16.530				
Relugolix CT	8,185	21.525	16.633	444	0.000	0.103	4,311
+ Exclude disutility from waiting time							
Goserelin monthly	7,742	21.525	16.536				
Relugolix CT	8,185	21.525	16.638	444	0.000	0.102	4,339
+ Remove waiting time before surgery							
Goserelin monthly	8,210	21.525	17.013				
Relugolix CT	8,617	21.525	17.059	407	0.000	0.046	8,784

	Total costs (£)	Total LYG	Total QALYs	Incremental Costs (£)	Incremental LYG	Incremental QALYs	ICER (£)
+ Apply utility parameters from repeated measures model							
Goserelin monthly	8,210	21.525	16.922				
Relugolix CT	8,617	21.525	16.992	407	0.000	0.070	5,846
+ Source for general population utilities: female							
Goserelin monthly	8,210	21.525	16.968				
Relugolix CT	8,617	21.525	17.037	407	0.000	0.069	5,866
+ Alternative resource use assumptions assuming one-off/routine admin/monitoring/examination costs							
Goserelin monthly	6,379	21.525	16.968				
Relugolix CT	6,573	21.525	17.037	194	0.000	0.069	2,795
ERG preferred base case analysis (deterministic)							
Goserelin monthly	6,379	21.525	16.968				
Relugolix CT	6,573	21.525	17.037	194	0.000	0.069	2,795
ERG preferred base case analysis (probabilistic)							
Goserelin monthly	6 376	--	16.957				
Relugolix CT	6 573	--	17.026	197	--	0.069	2 833

Abbreviations: ICER: Incremental cost-effectiveness ratio; LYG: Life year gains, QALY: Quality adjusted life years

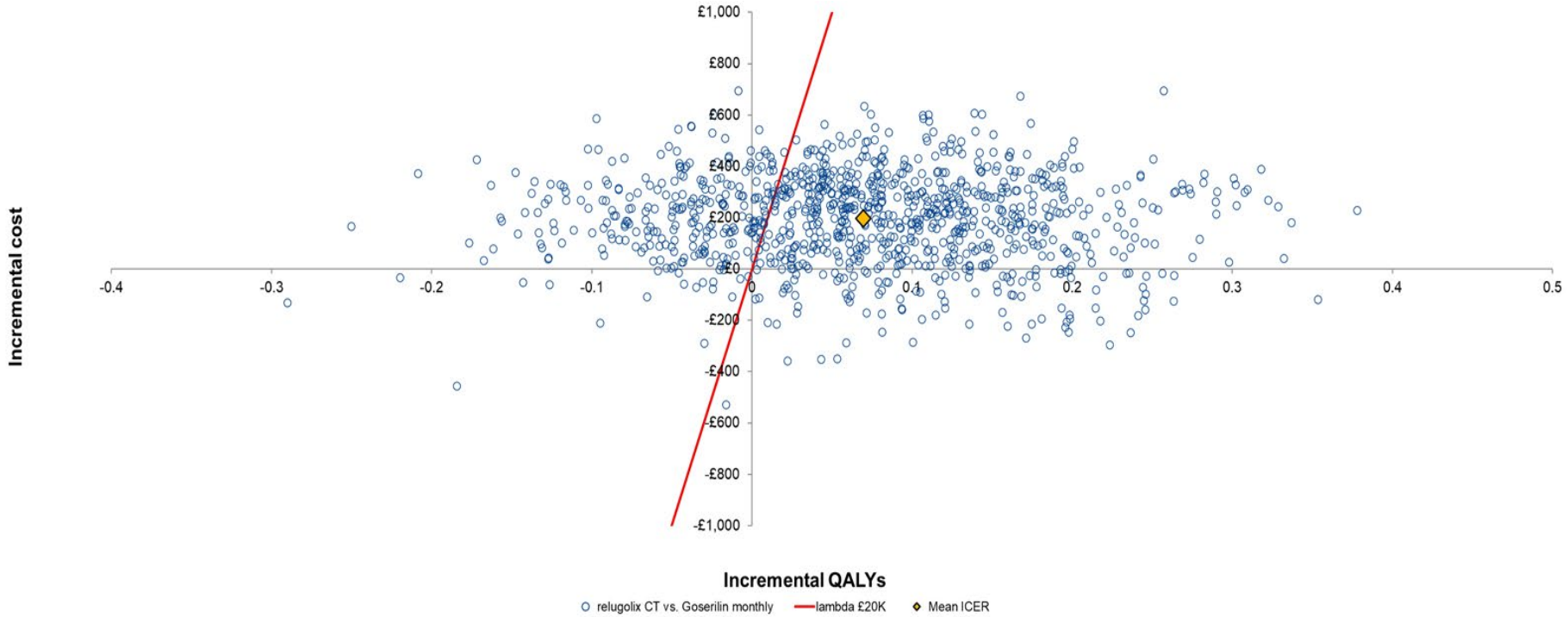


Figure 10 Scatter plot of the cost-effectiveness plane for the ERG's preferred base case probabilistic analysis (relugolix CT versus goserelin monthly)

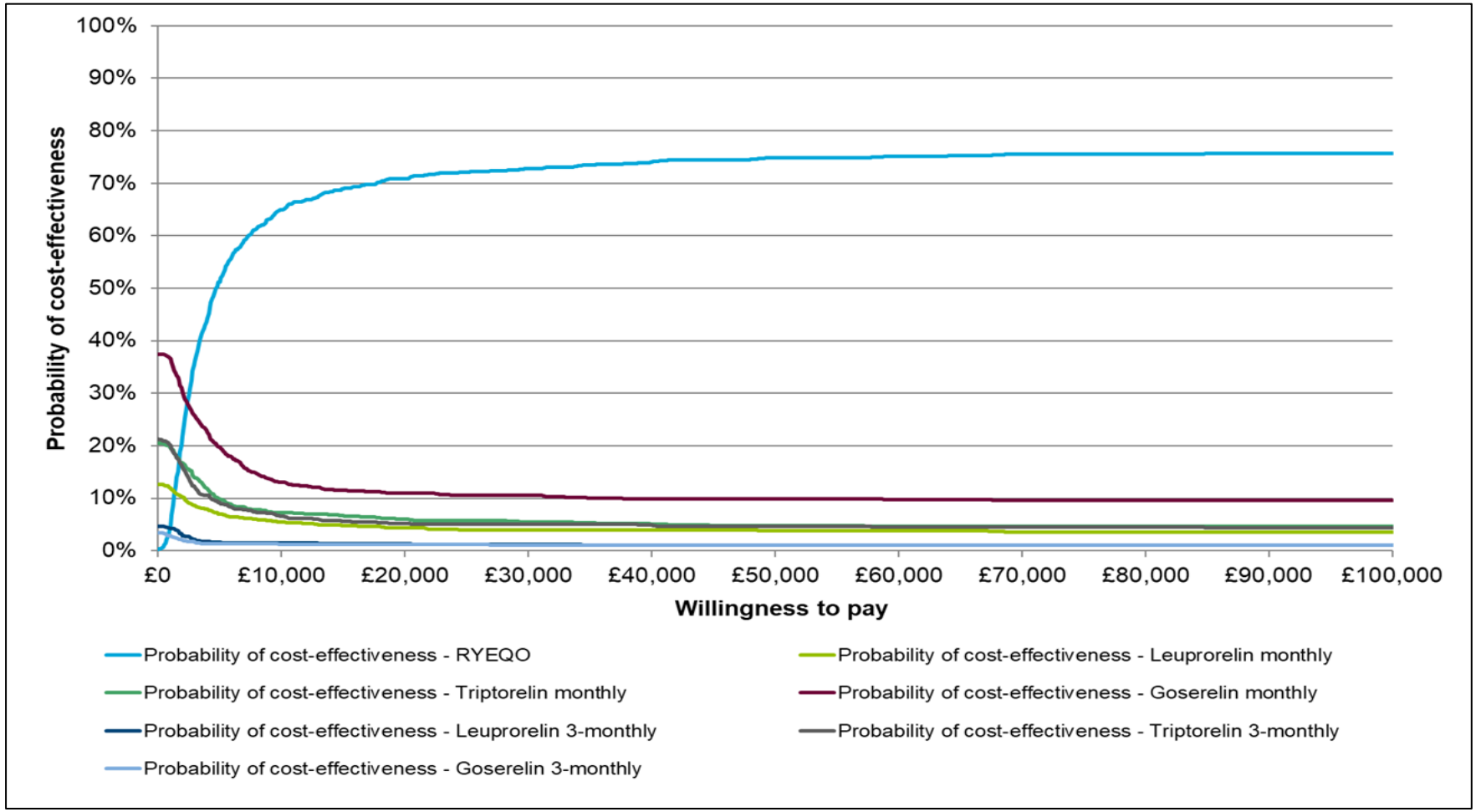


Figure 11 Cost-effectiveness acceptability curves for the ERG's preferred base case probabilistic analysis

Table 38 Scenario and exploratory analyses applied to the ERG preferred base case

	Total Costs (£)	Total LYG	Total QALYs	Incremental Costs (£)	Incremental LYG	Incremental QALYs	ICER (£)
ERG preferred base case analysis							
Goserelin monthly	6,379	21.525	16.968				
Relugolix CT	6,573	21.525	17.037	194	0.000	0.069	2,795
One round of surgery (i.e. no one needs repeat surgery - cure rate = 100%)							
Goserelin monthly	5,787	21.525	17.082				
Relugolix CT	6,054	21.525	17.136	267	0.000	0.054	4,983
Apply the minimum KOL max cap on the proportion on GnRHa treatment							
Goserelin monthly	5,928	21.525	16.966				
Relugolix CT	6,573	21.525	17.037	645	0.000	0.072	9,014
Apply the maximum KOL max cap on the proportion on GnRHa treatment							
Goserelin monthly	6,891	21.525	17.370				
Relugolix CT	6,573	21.525	17.037	-318	0.000	0.059	Dominant
Use GnRHa within its licence (6 months)							
Goserelin monthly	5,768	21.524	16.970				
Relugolix CT	6,573	21.525	17.037	805	0.000	0.068	11,901
Source for surgery rates (Strong et al.)							
Goserelin monthly	6,559	21.525	16.912				
Relugolix CT	6,727	21.525	16.989	167	0.000	0.077	2,163

	Total Costs (£)	Total LYG	Total QALYs	Incremental Costs (£)	Incremental LYG	Incremental QALYs	ICER (£)
Utilities for Relugolix CT, GnRHa, and BSC applied as decrements from general population norms							
Goserelin monthly	6 379	21.525	17.405				
Relugolix CT	6 573	21.525	17.498	194	0.000	0.093	2,082
0% discount rate							
Goserelin monthly	7,086	42.085	32.699				
Relugolix CT	7,335	42.086	32.775	249	0.000	0.075	3,302
6% discount rate							
Goserelin monthly	5,957	15.113	11.977				
Relugolix CT	6,118	15.113	12.042	161	0.000	0.066	2,454
Time horizon: until menopause (aged 51)							
Goserelin monthly	6,379	5.920	6.330				
Relugolix CT	6,573	5.990	6.341	194	0.000	0.069	2,797
Subgroup analysis: long-term treatment = 0% transition to surgery							
Goserelin monthly	2,856	21.525	16.868				
Relugolix CT	3,452	21.525	16.768	596	0.000	0.100	5,967
Subgroup analysis: short-term treatment = 100% transitions to surgery							
Goserelin monthly	8,519	21.524	17.241				
Relugolix CT	8,536	21.524	17.245	17	0.000	0.004	4,563

6.4 Conclusions of the cost-effectiveness section

The company preferred base case analysis and associated scenario analyses generate ICERs well below £20,000 per QALY gained. The ERG's suggested alternative base case also generates a similar ICER, but with substantially lower incremental costs and incremental QALY gains compared to the company base case. As noted in the critique throughout this report, the ERG's main conclusion is that it is very difficult to draw a clear conclusion on the most appropriate base case set of assumptions as data are often sparse and assumptions unclear. Plausible combinations of different scenario analyses would lead to wide variation in the ICER, and results are highly uncertain. The revised ERG probabilistic analyses illustrate substantial uncertainty that is not apparent when examining deterministic analyses alone. The ERG view is that it is essential that decision-makers are aware of this uncertainty and consider it in their judgments.

7 References

1. Walker CL, Stewart EA. Uterine fibroids: the elephant in the room. Science. 2005;**308**(5728):1589-92.

2. Stewart EA. Uterine fibroids. Lancet. 2001;**357**(9252):293-8.

3. NHS England. Fibroids 2018. Available from: https://www.nhs.uk/conditions/fibroids/ (Accessed 18 November 2021).

4. BMJ Best Practice. Uterine fibroids 2021. Available from: https://bestpractice.bmj.com/topics/en-gb/567 (Accessed 18 November 2021).

5. Guy's and St Thomas' NHS Foundation Trust. Could you have fibroids? 2016. Available from: https://www.guysandstthomas.nhs.uk/resources/patient-information/gynaecology/Fibroids.pdf (Accessed 18 November 2021).

6. Yang Q, Ciebiera M, Victoria Bariani M, Ali M, Elkafas H, Boyer TG, et al. Comprehensive review of uterine fibroids: developmental origin, pathogenesis, and treatment. Endocr Rev. 2021.

7. Stewart EA, Cookson CL, Gandolfo RA, Schulze-Rath R. Epidemiology of uterine fibroids: a systematic review. BJOG. 2017;**124**(10):1501-12.

8. NHS England. Fibroids: Diagnosis. 2018. Available from: https://www.nhs.uk/conditions/fibroids/diagnosis/ (Accessed 18 November 2021).

9. Drayer SM, Catherino WH. Prevalence, morbidity, and current medical management of uterine leiomyomas. Int J Gynaecol Obstet. 2015;**131**(2):117-22.

10. NHS Digital. Hospital Admitted Patient Care Activity, 2020-21: Diagnosis. 2021. Available from: https://digital.nhs.uk/data-and-information/publications/statistical/hospital-admitted-patient-care-activity/2020-21#resources (Accessed 18 November 2021).

11. National Institute for Health and Care Excellence. NICE Pathways: Heavy menstrual bleeding. 2021. Available from: https://pathways.nice.org.uk/pathways/heavy-menstrual-bleeding#path=view%3A/pathways/heavy-menstrual-bleeding/managing-heavy-menstrual-bleeding.xml&content=view-index (Accessed 18 November 2021).

12. Centre for Reviews and Dissemination. Systematic reviews: CRD's guidance for undertaking systematic reviews in health care. University of York: 2009. Available from: URL: http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm.

13. Myovant Sciences GmbH. LIBERTY 1 (MVT-601-3001) Clinical Study Report [Data on file]. Basel: Myovant Sciences GmbH; 2020.

14. Myovant Sciences GmbH. LIBERTY 2 (MVT-601-3002) Clinical Study Report [Data on file]. Basel: Myovant Sciences GmbH; 2020.

15. Donnez J, Tatarchuk TF, Bouchard P, Puscasiu L, Zakharenko NF, Ivanova T, et al. Ulipristal acetate versus placebo for fibroid treatment before surgery. N Engl J Med. 2012;**366**(5):409-20.

16. Donnez J, Tomaszewski J, Vázquez F, Bouchard P, Lemieszczuk B, Baró F, et al. Ulipristal acetate versus leuprolide acetate for uterine fibroids. N Engl J Med. 2012;**366**(5):421-32.

17. Badiani B, Chiumente M, Messori A. Ulipristal acetate for pre-operative management of uterine fibroids: Modeling outcomes and costs. Eur J Obstet Gynecol Reprod Biol. 2018;**222**:84-8.

18. Nagy B, Tímár G, Jozwiak-Hagymásy J, Kovacs G, Merész G, Vámossy I, et al. PIH24 Economic evaluation of ulipristal acetate tablets for the treatment of patients with moderate and severe symptoms of uUterine fibroids. Value Health. 2012;**15**(7):A540.

19. Geale K, Saridogan E, Lehmann M, Arriagada P, Hultberg M, Henriksson M. Repeated intermittent ulipristal acetate in the treatment of uterine fibroids: a cost-effectiveness analysis. Clinicoecon Outcomes Res. 2017;**9**:669-76.

20. Choi YR, Lee S, Choi I, Suh D. Cost-utility analysis of ulipristal acetate 5mg for treating patients with uterine fibroids in South Korea. Value Health. 2016;**19**(3):A177.

21. Lorenzovici L, Székely A, Ágh T, Vámossy I, Kelemen LJ, Finta H, et al. Economic evaluation of ulipristal acetate for the treatment of patients with moderate and severe symptoms of uterine fibroids in Romania. Value Health. 2014;**17**(7):A510.

22. Maratea D. Repeated-intermittent use of ulipristal acetate for the management of uterine fibroids: an Italian pharmacoeconomic evaluation. Minerva Ginecol. 2016;**68**(1):15-20.

23. Nagy B, Timár G, Józwiak-Hagymásy J, Kovács G, Merész G, Vámossy I, et al. The cost-effectiveness of ulipristal acetate tablets in treating patients with moderate to severe symptoms of uterine fibroids. Eur J Obstet Gynecol Reprod Biol. 2014;**175**:75-81.

24. Paladio-Hernández J, Rosas R, Pozo L, Del Cid O, Robles_Valencia JA. Economic evaluation of ulipristal acetate for the treatment of patients with moderate and severe symptoms of uterine fibroids before surgery in Mexico. Value Health. 2015;**18**(7):A834.

25. Paquete A, Lehmann M, Silva DP, Pinto CG. Cost utility analysis of ulipristal acetate (5 Mg) for long-term treatment of moderate to severe symptoms of uterine fibroids in Portugal. Value Health. 2016;**19**(7):A403.

26. Tsoi B, Blackhouse G, Ferrazzi S, Reade CJ, Chen I, Goeree R. Incorporating ulipristal acetate in the care of symptomatic uterine fibroids: a Canadian cost-utility analysis of pharmacotherapy management. Clinicoecon Outcomes Res. 2015;**7**:213-25.

27. Lethaby A, Vollenhoven B, Sowter M. Pre-operative GnRH analogue therapy before hysterectomy or myomectomy for uterine fibroids. Cochrane Database Syst Rev. 2001(2):Cd000547.

28. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal [PMG9]. 2013. Available from: https://www.nice.org.uk/process/pmg9/chapter/foreword (Accessed 30 November 2021).

29. NHS England. Guide to NHS waiting times in England 2019. Available from: https://www.nhs.uk/nhs-services/hospitals/guide-to-nhs-waiting-times-in-england/ (Accessed 30 November 2021).

30. Gupta JK, Sinha A, Lumsden MA, Hickey M. Uterine artery embolization for symptomatic uterine fibroids. Cochrane Database Syst Rev. 2014(12):Cd005073.

31. Gorny KR, Woodrum DA, Brown DL, Henrichsen TL, Weaver AL, Amrami KK, et al. Magnetic resonance-guided focused ultrasound of uterine leiomyomas: review of a 12-month outcome of 130 clinical patients. J Vasc Interv Radiol. 2011;**22**(6):857-64.

32. NHS Digital. Hospital Outpatient Activity - 2012-13. 2013. Available from: https://digital.nhs.uk/data-and-information/publications/statistical/hospital-outpatient-activity/hospital-outpatient-activity-2012-13 (Accessed 30 November 2021).

33. Carls GS, Lee DW, Ozminkowski RJ, Wang S, Gibson TB, Stewart E. What are the total costs of surgical treatment for uterine fibroids? J Womens Health (Larchmt). 2008;**17**(7):1119-32.

34. Settnes A, Topsoee MF, Moeller C, Dueholm M, Kopp TI, Norrbom C, et al. Reduced complications following implementation of laparoscopic hysterectomy: a Danish population-based cohort study of minimally invasive benign gynecologic surgery between 2004 and 2018. J Minim Invasive Gynecol. 2020;**27**(6):1344-53.e3.

35. Szende A, Janssen B, Cabases J. Self-reported population health: an international perspective based on EQ-5D. Dordrecht (NL): Springer; 2014.

36. NHS Business Services Authority. NHS England and Wales electronic drug tariff. 2021. Available from: https://www.drugtariff.nhsbsa.nhs.uk/#/00812617-DC/DC00812614/Home (Accessed 30 November 2021).

37. NHS England. National Cost Collection for the NHS 2019/20. 2021. Available from: https://www.england.nhs.uk/national-cost-collection/ (Accessed 18 November 2021).

38. Curtis L, Burns A. Unit Costs of Health and Social Care 2020. Canterbury: Personal Social Services Research Unit.; 2020. Available from: https://doi.org/10.22024/UniKent%2F01.02.84818.

39. Medicines Complete. British National Formulary London: Royal Pharmaceutical Society; 2020. Available from: https://www.medicinescomplete.com/mc/bnf/current/index.htm (Accessed 18 November 2021).

40. Brummer TH, Jalkanen J, Fraser J, Heikkinen AM, Kauko M, Mäkinen J, et al. FINHYST, a prospective study of 5279 hysterectomies: complications and their risk factors. Hum Reprod. 2011;**26**(7):1741-51.

41. Manyonda IT, Bratby M, Horst JS, Banu N, Gorti M, Belli AM. Uterine artery embolization versus myomectomy: impact on quality of life - results of the FUME (Fibroids of the Uterus: Myomectomy versus Embolization) Trial. Cardiovasc Intervent Radiol. 2012;**35**(3):530-6.

42. Tappenden P, Chilcott JB. Avoiding and identifying errors and other threats to the credibility of health economic models. Pharmacoeconomics. 2014;**32**(10):967-79.