Single Technology Appraisal

Relugolix–estradiol–norethisterone acetate for treating symptoms of endometriosis [ID3982]

Committee Papers

© National Institute for Health and Care Excellence 2024. All rights reserved. See Notice of Rights. The content in this publication is owned by multiple parties and may not be re-used without the permission of the relevant copyright owner.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Relugolix–estradiol–norethisterone acetate for treating symptoms of endometriosis [ID3982]

Contents:

The following documents are made available to stakeholders:

Access the **final scope** and **final stakeholder list** on the NICE website.

1. **Company submission from Gedeon Richter:**
   a. Full submission
   b. Summary of Information for Patients (SIP)

2. **Clarification questions and company responses**
   a. Additional response

3. **Patient group, professional group, and NHS organisation submissions** from:
   a. Endometriosis UK
   b. Elizabeth Bruen expert statement
   c. Karolina Afors expert statement

4. **External Assessment Report** prepared by KSR
   a. Updated appendix

5. **External Assessment Report – factual accuracy check**

Any information supplied to NICE which has been marked as confidential, has been redacted. All personal information has also been redacted.

© National Institute for Health and Care Excellence 2024. All rights reserved. See Notice of Rights. The content in this publication is owned by multiple parties and may not be re-used without the permission of the relevant copyright owner.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Relugolix–estradiol–norethisterone acetate for treating symptoms of endometriosis [ID3982; GID-TA10873]

Document B

Company evidence submission

October 2023

File name	Version	Contains confidential information	Date
ID3982_RelugolixCT_Document B_FINAL_UPDATED_07.02.24 [REDACTED]	**Final (updated)**	**Yes**	**07.02.24**

© Gedeon Richter UK Ltd (2023). All rights reserved

Contents

© Gedeon Richter UK Ltd (2023). All rights reserved

Tables and figures

Tables

© Gedeon Richter UK Ltd (2023). All rights reserved

© Gedeon Richter UK Ltd (2023). All rights reserved

© Gedeon Richter UK Ltd (2023). All rights reserved

Figures

© Gedeon Richter UK Ltd (2023). All rights reserved

© Gedeon Richter UK Ltd (2023). All rights reserved

Abbreviations

Term	Definition
AE	Adverse event
BMD	Bone mineral density
BMI	Body mass index
BSC	Best supportive care
CI	Confidence interval
CrI	Credible interval
CT	Combination therapy
CV	Cardiovascular
DIC	Deviance information criterion
DXA	Dual x-ray absorptiometry
ECG	Electrocardiogram
EHP-30	Endometriosis health profile-30
EMB	Endometrial biopsy
EOT	End of treatment
ESHRE	European Society for Human Reproduction and Embryology
FHS	Framlingham Heart Study
GnRH	Gonadotropin releasing hormone
GP	General practitioner
HDL	High density lipoprotein
HRQoL	Health-related quality of life
HRU	Healthcare resource use
ICER	Incremental cost-effectiveness ratio
ITC	Indirect treatment comparison
KOL	Key opinion leader
LA	Leuprolide acetate
LARC	Long-acting reversible contraception
LDL	Low density lipoprotein

© Gedeon Richter UK Ltd (2023). All rights reserved

LS	Least squares
LUNA	Laparoscopic uterosacral nerve ablation
MCM	Markov Chain model
mITT	Modified intent to treat
NETA	Norehisterone acetate
NMA	Network meta-analysis
NMPP	Non-menstrual pelvic pain
NSAID	Non-steroidal anti-inflammatory
NRS	Numerical rating scale
OLE	Open-label extension
OLS	Ordinary least squares
OPP	Overall pelvic pain
OWSA	One-way deterministic sensitivity analysis
PGA	Patient global assessment
PGIC	Patient global impression of change
PSA	Probabilistic sensitivity analysis
Q4W	Every 4 weeks
QALY	Quality-adjusted life year
QD	Once-daily
RCT	Randomised clinical trial
RR	Relative risk
SAE	Serious adverse event
SD	Standard deviation
SE	Standard error
SLR	Systematic literature review
TEAE	Treatment-emergent adverse event
TPP	Total pelvic pain
VAS	Visual analogue scale

© Gedeon Richter UK Ltd (2023). All rights reserved

B.1 Decision problem, description of the technology and clinical care pathway

B.1.1 Decision problem

The submission covers the technology's full marketing authorisation for this indication, i.e. symptomatic treatment of endometriosis in women with a history of previous medical or surgical treatment for their endometriosis.

Table 1: The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	Adults with symptoms of endometriosis	Same as scope The ITC and economic analysis presented in this submission focus on the subgroup of patients who remain symptomatic following treatment with conventional hormonal therapy, including combined hormonal contraception and oral and intra-uterine progestogens	

© Gedeon Richter UK Ltd (2023). All rights reserved

Intervention	Relugolix in combination with oestradiol and norethisterone acetate (also known as norethisterone acetate) *[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	
Comparator(s)	Established clinical management without relugolix in combination with oestradiol and norethisterone, including: • analgesics or non-steroidal anti-inflammatory drug (NSAID) alone or in combination with each other • neuromodulators • hormonal treatment such as combined hormonal contraception (off-label for some combined hormonal contraceptives), oral progestogens, gonadotropin-releasing hormone (GnRH) agonists.	The submission will focus on GnRH agonists as the relevant comparator for Relugolix CT	Relugolix CT will be the only oral GnRH antagonist available for the long-term management of symptoms associated with endometriosis. As such there are no direct, licensed comparators. GnRH agonists are the closest comparator in this position in the clinical pathway of care, however, please note that they are not licensed for use past 6 months.
Outcomes	The outcome measures to be considered include: • overall pain	The outcome measures in the clinical effectiveness section include: • dysmenorrhoea	Admission to hospital and fertility were not collected in the Relugolix CT clinical trials

© Gedeon Richter UK Ltd (2023). All rights reserved

	• opioid use • analgesic use • recurrence of endometriosis • admission to hospital • subsequent surgical treatment • fertility • adverse effects of treatment • complications of treatment • health-related quality of life	• non-menstrual pelvic pain • dyspareunia • EHP-30 pain domain • opioid use • analgesic use • EQ-5D-5L • adverse effects The outcome measures in the ITC include: • overall pelvic pain (OPP) • total pelvic pain (TPP) The outcome measures in the cost-effectiveness model include: • dysmenorrhoea • non-menstrual pelvic pain • recurrence of pain • analgesic use • subsequent surgical treatment • subsequent medical treatment • complications related to surgery • health-related quality of life	

© Gedeon Richter UK Ltd (2023). All rights reserved

Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective. The availability of any commercial arrangements for the intervention, comparator and subsequent treatment technologies will be taken into account.	Same as scope	
Special considerations including issues related to equity or equality		There is evidence to suggest that women from some minority ethnic groups may be underdiagnosed and/or present later for help with endometriosis and thus have more severe symptoms. The Endometriosis All-Party Parliamentary Group Report (October 2020) also highlights that Black, Asian and minority ethnic communities can receive a lower quality of care. These health	

© Gedeon Richter UK Ltd (2023). All rights reserved

		inequalities have been thought to be due to socioeconomic factors since Black, Asian, and minority ethnic women are more likely to live in areas of high deprivation, have lower incomes, experience language barriers and have poorer access to women's healthcare services	

© Gedeon Richter UK Ltd (2023). All rights reserved

B.1.2 Description of the technology being evaluated

Table 2 provides an overview of Relugolix CT. The draft Summary of Product Characteristics (SmPC) is included in Appendix C1.1. At the time of submission, there was no public assessment report available.

Table 2: The technology being evaluated

UK approved name and brand name	Relugolix in combination with oestradiol and norethisterone acetate [Brand name: Ryeqo®]
Mechanism of action	Relugolix is a non-peptide GnRH receptor antagonist that binds to and inhibits GnRH receptors in the anterior pituitary gland. In humans, inhibition of GnRH receptor results in a dose dependent decrease in the release of luteinising hormone (LH) and follicle-stimulating hormone (FSH) from the anterior pituitary gland. As a result, circulating concentrations of LH and FSH are reduced. The reduction in FSH concentrations prevents follicular growth and development, thereby reducing the production of oestrogen. Prevention of an LH surge inhibits ovulation and development of the corpus luteum, which precludes the production of progesterone.
Marketing authorisation/CE mark status	Relugolix CT is currently being appraised via the EMA reliance route. A submission was made to the EMA in September 2022. CHMP positive opinion was received in September 2023, with regulatory approval expected in November 2023.
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	The anticipated indication for Relugolix CT is "Symptomatic treatment of endometriosis in women with a history of previous medical or surgical treatment for their endometriosis" Please be aware that Relugolix CT was originally submitted for regulatory approval with the following proposed indication: "Moderate to severe pain associated with endometriosis in women with a history of previous medical or surgical treatment for their endometriosis.". However, the EMA requested a broader licence.
Method of administration and dosage	One tablet of Relugolix CT must be taken once daily, at about the same time with or without food. Relugolix CT should be taken with some liquid as needed. Each tablet of Relugolix CT contains relugolix (40mg), oestradiol (1 mg) and norethisterone acetate (0.5 mg).

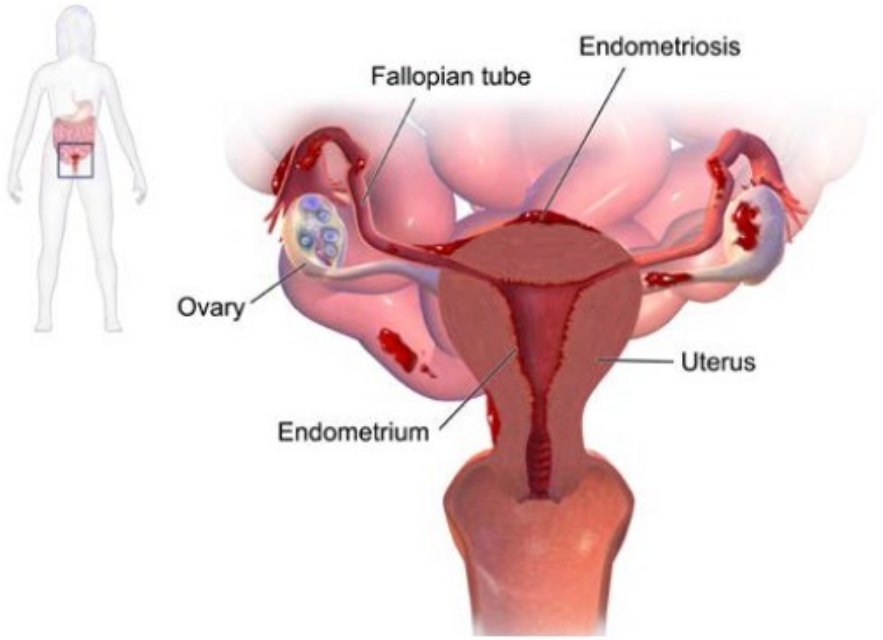
© Gedeon Richter UK Ltd (2023). All rights reserved

Additional tests or investigations	n/a
List price and average cost of a course of treatment	£72 per pack (containing 28 tablets) to be taken once daily
Patient access scheme (if applicable)	n/a

B.1.3 Health condition and position of the technology in the treatment pathway

Endometriosis is a chronic inflammatory disease associated with infertility and pelvic pain that is characterised by growth of endometrial-like tissue outside the uterus (Figure 1). It mainly affects women of reproductive age, of which an estimated 10% are affected, making endometriosis one of the most common gynaecological conditions requiring treatment. In the UK it is estimated that 1.5 million women are affected by endometriosis, similar to the number affected by diabetes mellitus. The disease is oestrogen-dependent, with endometrial-like tissue lesions requiring oestradiol for growth (1-9).

Figure 1: Endometriosis is characterised by the presence of endometrial tissue outside the endometrium (10).



© Gedeon Richter UK Ltd (2023). All rights reserved

The exact cause of endometriosis remains unclear, though several theories have been put forward to explain the disease (7, 11). The following factors are thought to contribute to the development of endometriosis:

- Retrograde menstruation – flow of endometrial cells backwards into the uterine cavity (12).
- Lymphatic or circulatory dissemination – travel of endometrial cells throughout the body via the bloodstream of lymphatic system (12, 13).
- Metaplasia – the process of differentiation into endometrial-type cells by cells in the pelvic or abdominal areas (12).
- Environmental factors – in theory, some environmental toxins including dioxin may contribute to development of endometriosis (12).
- Immune dysfunction – immunity to other conditions is often reduced in endometriosis patients, but the nature of this association is unclear (12).

Risk factors for endometriosis include the following (13, 14):

- Early menarche
- Late menopause
- Delayed childbearing
- Nulliparity
- Family history
- Vaginal outflow obstruction
- White ethnicity
- Low body mass index (BMI)
- Autoimmune disease
- Late first sexual encounter
- Smoking
- High alcohol consumption

© Gedeon Richter UK Ltd (2023). All rights reserved

The symptoms of endometriosis can be severe and wide-ranging, though some women with endometriosis will be asymptomatic. The extent of the endometriosis does not necessarily align with the amount of pain the patient experiences. This is mostly dependent on the location of the endometrial deposit (12).

Classical symptoms of endometriosis are pain during menstruation, pain during intercourse, pain while defecating, painful ovulation, pelvic pain, pain while urinating, pain radiating to the back, irregular and profuse menstruation, blood in the stool, infertility, chronic fatigue, diarrhoea and constipation, pain in the sacral region of the spine and an increasingly painful premenstrual period (12, 15, 16). Other associated symptoms include depression, nausea, fainting during periods, and frequent infections (7, 12, 15).

It remains unclear how endometriosis causes pain. Mechanisms suggested include stimulation of neural pathways, inflammation, local bleeding, hormonal stimulation of the endometrial deposits, and any combination of these factors (11).

There are several different types of endometrioses, each characterised by the location and appearance of the endometrial tissue (10, 15):

- Superficial endometriosis: This is the most common type of endometriosis and occurs when the endometrial tissue is found on the surface of the ovaries, fallopian tubes, or pelvic peritoneum.

- Ovarian endometriomas: Also known as chocolate cysts, these are large cysts filled with old blood that forms on the ovaries. These are less common and are mostly found in women with concomitant deep endometriosis or superficial endometriosis.

- Deep infiltrating endometriosis: This type of endometriosis occurs when the endometrial tissue penetrates the muscles and tissues of the pelvic cavity, including the rectum, bladder, and intestine.

- Miscellaneous types: Endometriosis can also occur in other parts of the body, such as the lungs, brain, and skin.

© Gedeon Richter UK Ltd (2023). All rights reserved

Diagnosis of endometriosis can be challenging, with the average time from first seeking treatment to diagnosis of 7.5 years (12). Diagnosis can only definitively be made via laparoscopy, although less invasive methods including ultrasound scans, MRI, pap smear, and vaginal and endocervical swabs may be used to assist diagnosis. The often-cyclical nature of symptoms, and the overlapping symptom profile with other diseases such as pelvic adhesions, gastrointestinal disorders, and adenomyosis, are both confounding factors in diagnosis of endometriosis (7, 8).

Burden associated with endometriosis

Endometriosis is the second most common gynaecological disease in the United Kingdom affecting an estimated 1.5–2 million women. It is associated with a significant economic, societal, and quality of life burden which may be underestimated due to lack of research (12, 17).

Whilst there is a lack of research on the economic burden of endometriosis in the UK, the charity Endometriosis UK estimates it to cost £8.2 billion per year in lost productivity, treatment, and healthcare costs (12, 17, 18). The 2020 APPG report on endometriosis found that 38% of those with endometriosis were worried about losing their jobs and 35% had reduced incomes due to the condition (18).

The World Endometrial Research Foundation (WERF) EndoCost study aimed to calculate the cost of endometriosis in 2008, using data from referral centres in 10 countries including the UK. The estimated cost of endometriosis was €9579 per woman, with €6298 in productivity costs, €3113 in healthcare costs, and €168 in non-healthcare costs. Twenty-nine percent of healthcare costs were due to surgery, with 19% due to monitoring tests, 18% to hospitalisation, and 16% to physician visits. This economic burden due to healthcare costs is comparable to chronic diseases such as diabetes mellitus, while the indirect costs are twice as great (17). Elsewhere in the literature, Nnoaham et al., estimated in 2013 that on average, women with endometriosis across 9 countries (including England, Ireland, and the US) lost 10.8 work hours per week, with a cost in England of ~$90 and ~$200 per week due to absenteeism and presenteeism, respectively (19). Nnoaham et al., may overestimate the burden of endometriosis by only including patients from referral centres who are presumably more severe and have higher treatment rates. However, the well

© Gedeon Richter UK Ltd (2023). All rights reserved

documented delay in diagnosis and subsequent large undiagnosed population implies that the overall economic and quality of life burden may be higher than estimated (19, 20).

Impact on quality of life

Various studies have found a significant association between endometriosis and reduced health-related quality of life (HRQoL) (17, 19, 21-26). This reduction in HRQoL appears to be strongly driven by pain, and weakly driven by infertility (26, 27).

In 2003, a long- term follow up study examining HRQoL in patients with endometriosis found that the condition led to significantly reduced EQ-5D scores vs population norm for both the physical component (43.5 vs 52.8) and mental component (46.7 vs 51.9) (28).

The WERF EndoCost study also found that the average woman with endometriosis generates 0.809 quality adjusted life years (QoL), experiencing a 19% reduction in quality of life. Key factors affecting quality of life were issues with usual activities (29%), pain and discomfort (56%), anxiety and depression (36%), issues with self-care (3%), and problems with mobility (16%). This quality-of-life burden is significant, being comparable to chronic diseases like rheumatoid arthritis (17).

Elsewhere in the literature, endometriosis is associated with depression in as many as 86% of patients, and anxiety in 87.5% (21). Sexual satisfaction is also significantly impacted in endometriosis patients in all domains of the Golombok-Rust Inventory of Sexual Satisfaction (GRISS) other than sexual communication and anorgasmia (39.27 endometriosis vs 29.79 healthy control) (24).

A further multinational study found that the impact of the disease on HRQoL is significantly worse for women with endometriosis in all SF-36v2 dimensions except physical functioning when compared to post-surgical sterilisation controls, and symptomatic controls (19).

A survey of 10,000 people living in the UK who were diagnosed with endometriosis was conducted as part of an all-party parliamentary group report on the burden of

© Gedeon Richter UK Ltd (2023). All rights reserved

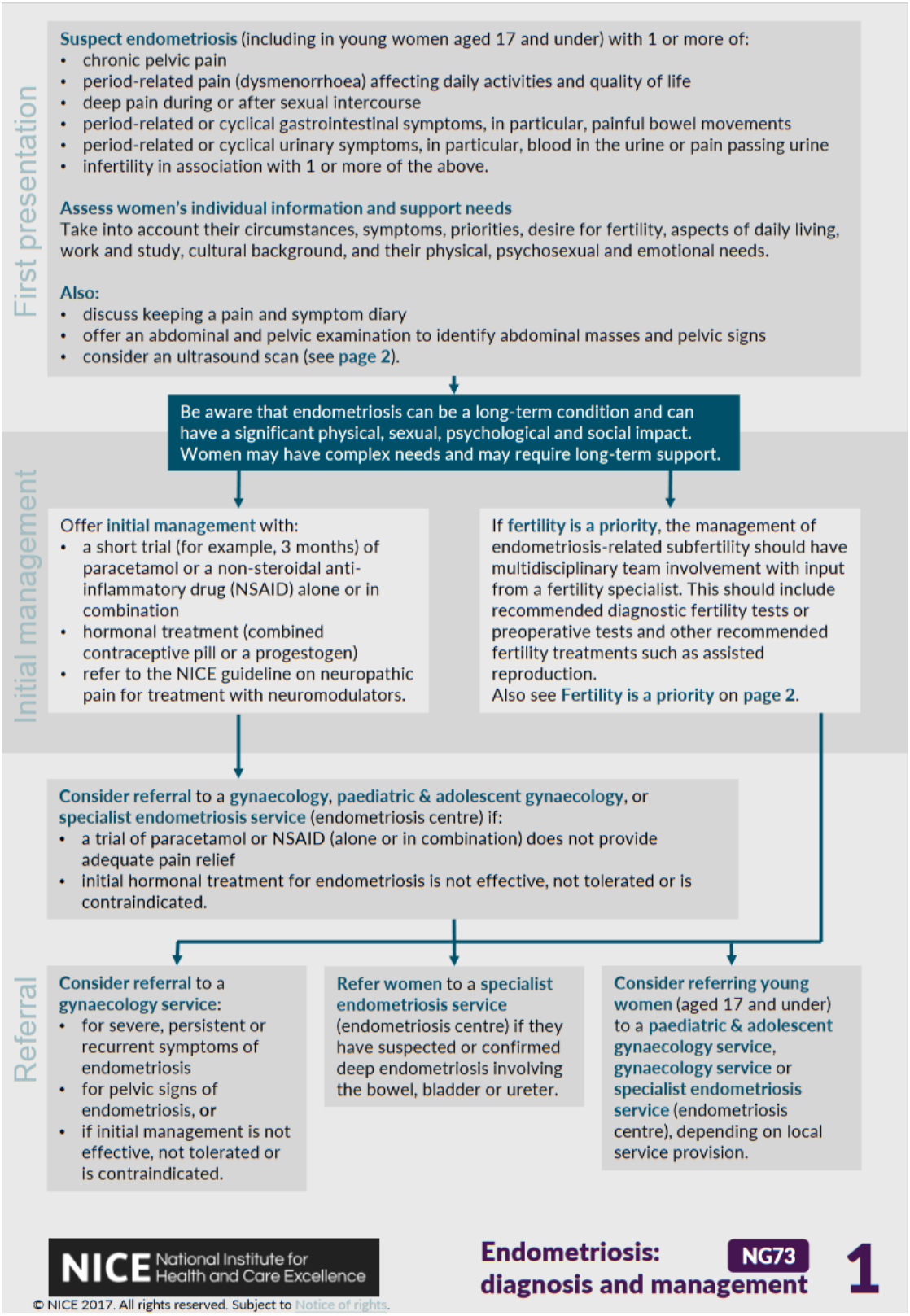
endometriosis in 2020. In the survey, 95% and 81% of respondents said that endometriosis had had a negative, or very negative, impact on their wellbeing and mental health, respectively. Furthermore, 89% felt isolated due to their condition, and 90% would have liked access to psychological support (18).

Treating endometriosis

There is currently no cure for endometriosis and the symptoms can manifest early in life, often leading to a course of disease covering multiple decades. Treatment focuses on controlling the symptoms and includes both surgical and pharmacological approaches. In 2017, NICE published a guideline for the diagnosis and management of endometriosis; Figures 2 and 3 show the recommended algorithm (7). The European Society of Human Reproduction and Embryology (ESHRE) have also produced guidelines on endometriosis, with the latest version published in 2022 (29).

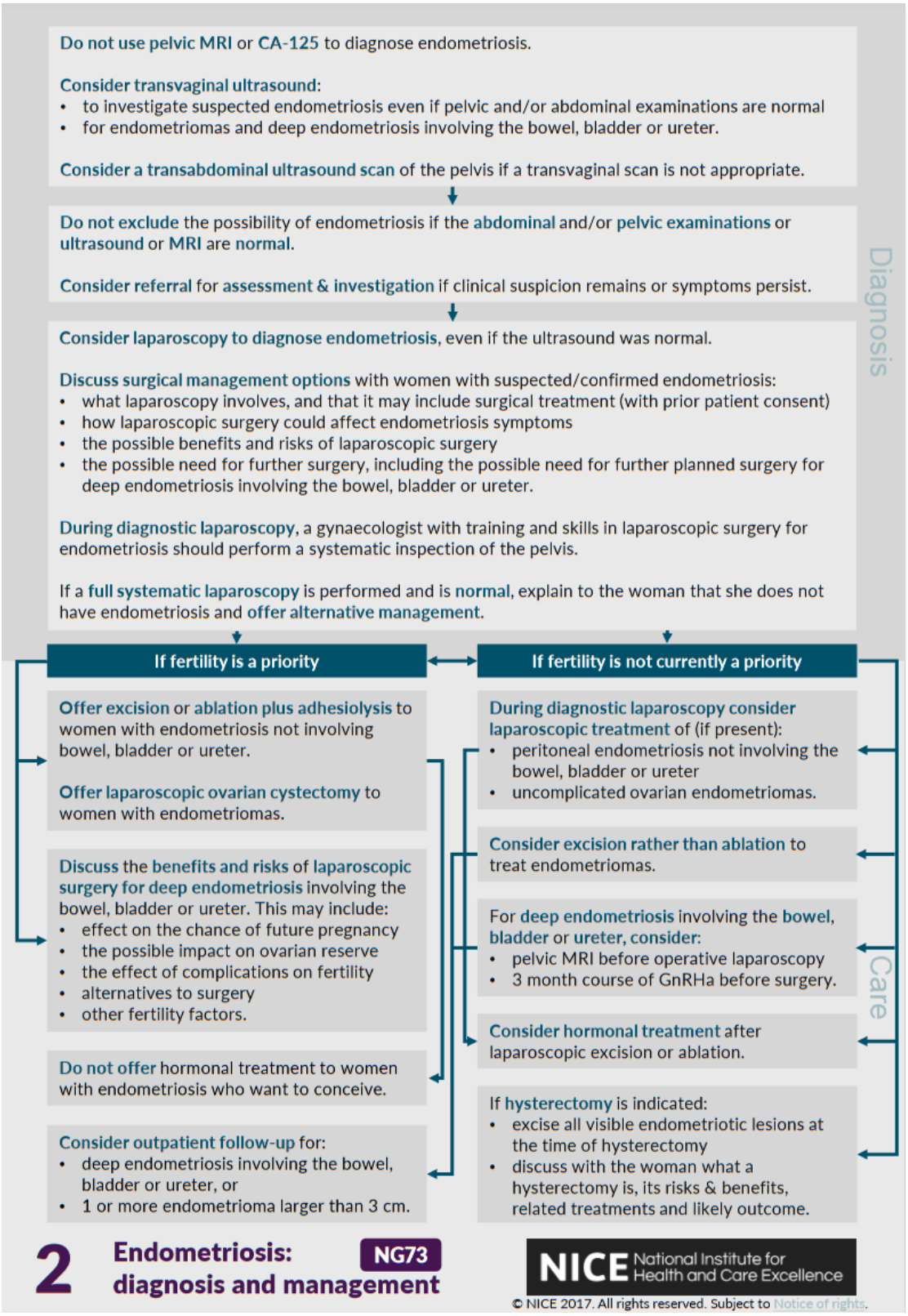
© Gedeon Richter UK Ltd (2023). All rights reserved

Figure 2:NICE algorithm for diagnosing and managing endometriosis 1 (7)



© Gedeon Richter UK Ltd (2023). All rights reserved

Figure 3: NICE algorithm for diagnosing and managing endometriosis 2 (7)



© Gedeon Richter UK Ltd (2023). All rights reserved

Analgesics

Both NICE and the ESHRE guidance (7, 29) recommend analgesics for pain management in endometriosis despite the limited available evidence to support their use (1, 7). Neuromodulators and tricyclic antidepressants have also been investigated for the treatment of endometriosis associated pain but were shown not to be superior to placebo in a recent study and are associated with severe side effects (1, 30).

Hormone treatments

Evidence suggests that endometriosis is a steroid-dependent condition. Thus, hormone therapy can be offered both pre- and post-surgery (the latter where the disease is persistent) and often prior to confirmation by laparoscopy where endometriosis is suspected. Most commonly these drugs alter the hormonal environment by acting on steroid receptors and enzymes in the lesions, or by suppression of ovarian activity. Drugs in this category include combined oral contraceptives, progestogens, anti-progestogens, levonorgestrel intrauterine system, aromatase inhibitors, danazol, gonadotropin releasing hormone (GnRH) agonists, and GnRH antagonists (1). Of the GnRH agonists and antagonists, only injectable options are currently available (31-33).

These hormonal therapies all appear to have a similarly significant impact on the reduction of endometriosis related pain vs placebo (1, 7). However, despite their efficacy these therapies suffer from tolerability issues, and in practice their use must be highly individualised (1). Safety issues for GnRH agonists include vaginal dryness, headache, weight gain, loss of libido, acne, and hot flushes. Furthermore, loss of bone mineral density was shown in patients on GnRH agonist treatment (1, 34, 35), which has led to the advent of add-back therapy. In such combination therapies progestin monotherapies such as oestrogen-progestin combinations, selective oestrogen receptor modulators, bisphosphonates, tibolone, norethisterone acetate (NETA), and testosterone are used to mitigate loss of bone mineral density with some success (1, 36). Similar side effect profiles are observed in GnRH antagonist monotherapy, with long term use beyond 6 months and repeat use being subsequently restricted. Thus, similar addback therapies have been investigated (31-33, 37).

© Gedeon Richter UK Ltd (2023). All rights reserved

Poor side effect profiles necessitate trial and error when narrowing down the optimal therapy for each patient. Additionally, many of these therapies have a contraceptive element and so are contraindicated where fertility is a concern (1).

Conservative surgical treatment

Historically, surgical interventions have been integral to the management of endometriosis associated pain. Laparoscopic elimination of endometriosis via excision (most common), ablation/vaporisation, and diathermy, is still central to treatment (1). Various reviews have examined the evidence for the reduction of endometriosis associated pain. A recent review of 14 RCTs using excision, coagulation, and CO_2 laser vaporisation found it was uncertain whether laparoscopic surgery was effective for pain relief (38). However, elsewhere in the literature surgery has been found superior to diagnostic laparoscopy for laparoscopic uterosacral nerve ablation (LUNA), adhesiolysis, and CO_2 laser vaporisation (39), excision of endometrioses (28, 40). Systematic reviews of HRQoL and sexual quality of life both found significant improvement due to laparoscopic surgery (41, 42). Laparoscopic surgery for endometriosis is considered safe with only low numbers of severe complications reported in the literature (38, 43, 44).

Unfortunately, recurrence rates are high for laparoscopic surgery. In one UK-based report, ~20% of those receiving surgical treatment for endometriosis received further surgery, while elsewhere in the literature recurrence rates vary, ranging from 21.5% at two years to 40%-50% at 5 years post-surgery (45-47).

Radical surgical management

Where management of endometriosis-associated symptoms has not been successful, more invasive methods may be indicated. This involves either partial or complete removal of the uterus and is common in the UK, with an incidence of 3.55 surgeries per 1000 women (48).

While hysterectomy can provide symptom relief in endometriosis patients, it is an invasive procedure which can result in a range of short- and long-term complications ranging from blood clots and infections, to urinary incontinence and early menopause (49, 50). In addition, there is a complete loss of fertility, a severe limitation in pre-menopausal women who wish to preserve fertility.

© Gedeon Richter UK Ltd (2023). All rights reserved

Issues with current treatments and unmet need

A severe unmet need remains for long-term tolerable treatment options, as side effects from current therapies, including hot flushes, loss of bone density and weight gain, drive a significant 5-16% discontinuation rate (37, 51). Furthermore, current later line therapies consist primarily of GnRH injections and surgery. Thus, a less invasive oral alternative, with good safety profile for long term therapy is required (7).

Additionally, efficacy appears to vary with existing medical therapies. A 2017 systematic review of medical treatments in endometriosis found that 11-19% of women report no relief from pain, 5-59% continued reporting pain through to the end of the treatment period, and 17-34% report recurrent pain after treatment cessation (51, 52).

Most medical interventions are contraceptive in nature, causing further distress as women of childbearing age are forced to choose between chronic pain and delaying or forgoing starting a family. Overall, medical interventions are still restricted in their longevity, efficacy, and usefulness (1, 51-53).

Laparoscopy interventions are associated with high pain recurrence rates of 30-60% within 6-12 months, and a 7 year reoperation-free survival rate of less than 50%. Reoperation rates are significantly higher in younger women, with reoperation rates in women aged 19-29 years being 2.56 times higher than in those aged 30-39 years, and 6.66 times higher than in those 40 years and older. Overall, 20% of women appear to show no improvement following surgical intervention (45-47, 51, 53-55). Hysterectomy does not provide a guaranteed cure either, with a 7 year reoperation-free survival rate of 84.6% in women with endometriosis (54).

These limitations highlight the issues with both medical and surgical interventions, illustrating the need for additional therapies with long term tolerability and efficacy in endometriosis.

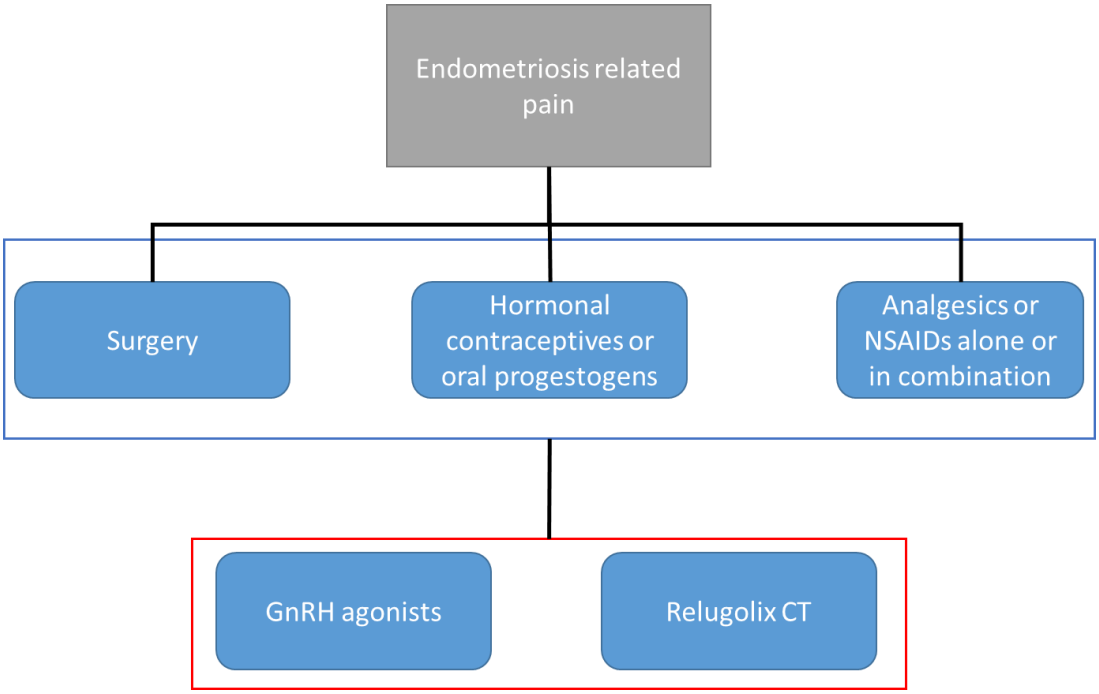
Proposed place of Relugolix CT

A severe unmet need for new therapies persists for patients with endometriosis-related symptoms as current treatments are either invasive, or not licensed for

© Gedeon Richter UK Ltd (2023). All rights reserved

longer-term treatment. The proposed place of Relugolix CT will be at second line, as an oral alternative to current GnRH agonist injections for symptomatic treatment of endometriosis in women with a history of previous medical or surgical treatment for their endometriosis. To reflect this positioning, the ITC and economic model described in this submission focus on a comparison of Relugolix CT with GnRH agonists in patients who remain symptomatic following treatment with conventional hormonal therapy, including combined hormonal contraception and oral and intra-uterine progestogens.

Figure 4: Proposed placement of Relugolix CT in the endometriosis treatment pathway



B.1.4 Equality considerations

There is evidence to suggest that women from some minority ethnic groups may be underdiagnosed (56, 57) and/or present later for help with endometriosis and thus have more severe symptoms.

The Endometriosis All-Party Parliamentary Group Report (October 2020) also highlights that Black, Asian and minority ethnic communities can receive a lower quality of care (18). These health inequalities have been thought to be due to socioeconomic factors since Black, Asian, and minority ethnic women are more likely

© Gedeon Richter UK Ltd (2023). All rights reserved

to live in areas of high deprivation, have lower incomes, experience language barriers and have poorer access to women's healthcare services (18).

© Gedeon Richter UK Ltd (2023). All rights reserved

B.2 Clinical effectiveness

B.2.1 Identification and selection of relevant studies

An SLR and pragmatic literature search were conducted to identify and select the clinical evidence relevant to Relugolix CT for the treatment of pain associated with endometriosis. Full details are given in Appendix D.

In summary, the following Relugolix CT and comparator trials were identified:

Relugolix CT studies: Two phase 3 trials and one open label extension (OLE) study relating to Relugolix CT met the inclusion criteria: SPIRIT 1, SPIRIT 2, and the SPIRIT OLE. The results from SPIRIT 1 and 2 are reported in a 2022 publication by Guidice *et al.,* in the Lancet, while findings from the SPIRIT OLE have been presented as an oral communication at the ESHRE 2022 conference with an abstract subsequently published in Human Reproduction (July 2022) (37, 58).

Comparator studies: As the identified Relugolix CT studies do not provide a direct comparison with other treatments, additional criteria were applied to identify studies for inclusion in an indirect treatment comparison (see Section B.2.9 and Appendix D for full details). Three phase 3 comparator studies were identified for inclusion in the ITC: Lang 2018, Strowitzki 2010, and D'Hooghe 2019 (59-61). A short summary of these trials is given below; further information can be found in Section B.2.9 and Appendix D.

D'Hooghe et al (2019) was a Phase II, multicentre, double-blind, randomized, parallel group, placebo-controlled study comparing the efficacy and safety of ASP1707 (Opigolix) (3 mg, 5 mg, 10 mg, 15 mg), leuprolide acetate (3.75 mg), and placebo in 540 women with endometriosis-associated pain (59).

Lang et al (2018) was a 24-week, Phase 3, randomized, double-blind, placebo-controlled multicentre study to evaluate the efficacy and safety of 2 mg dienogest once-daily in 255 women in China aged 18-45 with laproscopically-diagnosed endometriosis and endometriosis-associated pelvic pain (60).

© Gedeon Richter UK Ltd (2023). All rights reserved

Strowitzki et al (2010) was a 24-week randomized, multicentre, open-label trial comparing dienogest with leuprolide acetate in women aged 18-45 years. The study was conducted at 17 centres in Germany, Austria, Spain, Poland, Italy, and Portugal (61).

B.2.2 List of relevant clinical effectiveness evidence

The efficacy, safety, and tolerability of Relugolix CT has been demonstrated in two replicate multicentre Phase 3 trials (SPIRIT 1 & 2), and an open-label phase 3 extension (SPIRIT OLE).

The trial identifiers are as follows:

- SPIRIT 1 (MVT-601-3101): NCT03204318
- SPIRIT 2 (MVT-601-3102): NCT03204331
- SPIRIT OLE (MVT-601-3103): NCT03654274

The results of SPIRIT 1 and 2 are published in the Lancet (37) while results from the SPIRIT OLE are published as a selected communication in Human Reproduction (58). Where unavailable in the publications, data in this submission are also taken from the SPIRIT 1 and 2 clinical study reports dated February 2021 (62, 63), the SPIRIT OLE study report dated July 2022 (64) and ClinicalTrials.gov (65).

In the SPIRIT 1 and 2 trials, patients were randomised 1:1:1 to either:

- Placebo for 24 weeks
- Relugolix co-administered with oestradiol 1 mg and NETA 0.5 mg for 24 weeks (referred to as ‘Relugolix CT’)
- Relugolix monotherapy for 12 weeks followed by co-administration with oestradiol 1 mg and NETA 0.5 mg for 12 weeks (referred to in this submission document as ‘relugolix + delayed CT’)

Eligible participants were enrolled into the SPIRIT OLE study on completion of either of the two parent studies (SPIRIT 1 or 2).

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 3: Clinical effectiveness evidence SPIRIT 1 and 2

Study	SPIRIT 1 (MVT-601-3101, NCT03204318) SPIRIT 2 (MVT-601-3102, NCT03204331)
Study design	Multinational, replicate, phase 3, multicentre, randomised, double-blind, placebo-controlled trials
Population	Premenopausal women ages 18-50 with endometriosis which was surgically or directly visualised with or without histological confirmation, or histological diagnosis alone, within the past 10 years
Intervention(s)	Relugolix 40 mg in combination with oestradiol 1 mg and norethisterone acetate 0.5 mg [Relugolix CT]
Comparator(s)	Placebo
Indicate if study supports application for marketing authorisation	Yes
Indicate if study used in the economic model	Yes
Rationale if study not used in model	N/A
Reported outcomes specified in the decision problem	• **Overall pain** • **Opioid use** • **Analgesic use** • **Health-related quality of life** • **Adverse events**
All other reported outcomes	• **Dysmenorrhoea** • **Non-menstrual pelvic pain** • Dyspareunia • EHP-30 pain domain

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 4: Clinical effectiveness evidence: SPIRIT OLE

Study	SPIRIT OLE (MVT-601-3103, NCT03654274)
Study design	Multinational, phase 3, open-label, single-arm, safety and efficacy extension study
Population	Premenopausal women ages 18-51 with endometriosis which was surgically or directly visualised with or without histological confirmation, or histological diagnosis alone, within the past 10 years
Intervention(s)	Relugolix 40 mg in combination with oestradiol 1 mg and norethisterone acetate 0.5 mg [Relugolix CT]
Comparator(s)	N/A
Indicate if study supports application for marketing authorisation	Yes
Indicate if study used in the economic model	Yes
Rationale if study not used in model	N/A
Reported outcomes specified in the decision problem	- **Overall pain** - **Opioid use** - **Analgesic use** - **Health-related quality of life** - **Adverse events**
All other reported outcomes	• **Dysmenorrhoea** • **Non-menstrual pelvic pain** • Dyspareunia • EHP-30 pain domain

An overview of the comparator trials included in the ITC is available in Appendix D.

© Gedeon Richter UK Ltd (2023). All rights reserved

B.2.3 Summary of methodology of the relevant clinical effectiveness evidence

Table 5: Trial information for SPIRIT 1 and 2 and the SPIRIT OLE

Trial number (acronym)	MVT-601-3101 (SPIRIT 1) (37)	MVT-601-3102 (SPIRIT 2) (37)	MVT-601-3103 (SPIRIT OLE) (65)
Location	124 centres globally including North America (Canada and United States) and Rest of World (Argentina, Belgium, Bulgaria, Czech Republic, Finland, Hungary, Poland, Portugal, South Africa, Spain, and Ukraine).	95 centres globally including North America (United States) and Rest of World (Australia, Brazil, Chile, Czech Republic, Georgia, Italy, New Zealand, Poland, Romania, and Sweden).	171 centres globally including North America (United States) and Rest of World (Argentina, Australia, Belgium, Brazil, Bulgaria, Chile, Czech Republic, Finland, Georgia, Hungary, Italy, New Zealand, Poland, Portugal, Romania, South Africa, Spain, and Ukraine).
Trial design	International phase 3 randomised, double-blind, placebo-controlled efficacy and safety studies		International phase 3 open-label, single-arm, long-term efficacy and safety study that enrolled eligible patients who completed their participation in one of the phase 3 randomised, double-blind, placebo-controlled pivotal (also referred to as "parent") studies (MVT-601-3101 or MVT-601-3102)
Eligibility criteria for participants	Premenopausal women ages 18-50 with endometriosis which was surgically or directly visualised with or without histological confirmation, or histological diagnosis alone, within the past 10 years		Completed 24 weeks of study drug treatment and study participation in either MVT-601-3101 or MVT-601-3102. Is not expected to undergo gynaecological surgery or other surgical procedures for treatment of endometriosis (including ablation, shaving, or excision) during the study, including during the Follow-Up Period, and the patient does not

© Gedeon Richter UK Ltd (2023). All rights reserved

		desire such treatment during this time frame.
Trial drugs	Participants were randomly assigned in a 1:1:1 ratio to three treatment groups to receive Relugolix CT (40 mg relugolix in combination with 1 mg oestradiol and 0.5 mg norethisterone acetate for 24 weeks), relugolix + delayed CT (40 mg relugolix monotherapy & placebo for 12 weeks, followed by 40 mg relugolix in combination with 1 mg oestradiol and 0.5 mg norethisterone acetate for 12 weeks), or placebo (relugolix placebo tablet co-administered with oestradiol/ norethisterone acetate placebo capsule for 24 weeks) • SPIRIT 1: Relugolix CT (N=212), relugolix + delayed CT (N=213), placebo (N=213) • SPIRIT 2: Relugolix CT (N=208), relugolix + delayed CT (N=207), placebo (N=208)	Relugolix 40 mg tablets co-administered orally QD with over-encapsulated low-dose oestradiol (1 mg) and NETA (0.5 mg) on an empty stomach for up to 80 weeks
Primary outcomes	Co-primary endpoints: • Proportion of responders in the Relugolix CT group vs placebo for non-menstrual pelvic pain at 24 weeks • Proportion of responders in the Relugolix CT group vs placebo for dysmenorrhoea at 24 weeks	Co-primary endpoints: • Week 52: Proportion of patients meeting the dysmenorrhoea responder criteria (reduction in pain scores and no increase in analgesic use) and NMPP responder criteria (reduction in pain scores and no increase in analgesic use). • Week 104: Proportion of patients meeting the dysmenorrhoea responder criteria and NMPP responder criteria.
Other outcomes used in the economic	• EHP-30 pain domain score • Dysmenorrhoea NRS score	Assessed at Week 52 and Week 104:

© Gedeon Richter UK Ltd (2023). All rights reserved

model/specified in the scope	• Non-menstrual pelvic pain NRS score • Overall pelvic pain NRS score • Mean dyspareunia NRS score • Protocol-specified opioids use • Protocol-specified analgesics for endometriosis (for SPIRIT 1) • Protocol-specified analgesic use based on mean pill count (for SPIRIT 2).		• Change in EHP-30 Pain Domain scores, proportion of patients with a significant reduction in EHP-30 Pain Domain scores, • Change in dysmenorrhoea and NMPP NRS scores, proportion of patients reporting improvement on PGIC, proportion of patients not using opioids or analgesics, • Change in dyspareunia NRS scores, improvement on PGIC for dyspareunia, • Change in dyspareunia functional impairment, • Change in severity scores on PGA for pain, change in function impairment on PGA, • Change in non-pain EHP-30 domains, • Change in dysmenorrhoea-related functional effects, • Change in NMPP-related functional effects.
Pre-planned subgroups	The following subgroups were analysed: • Geographic region (North America vs rest of world) • Time since endometriosis diagnosis (<5years / >5years)	The following subgroups were analysed: • Geographic region (North America vs rest of world) • Time since endometriosis diagnosis (<5 years / >5 years)	The following subgroups were analysed: • Geographic region (North America vs rest of world) • Age (<35 years / >= 35 years) • Race (Black or African American / White)

© Gedeon Richter UK Ltd (2023). All rights reserved

	• Time since endometriosis diagnosis (<2 years / 2-5 years / >= 5 years) • Age (<30 years / 30-35 years / 35-40 years / >=40 years) • Race (Black or African American / White) • BMI at baseline (<25 / 25-30 / >=30) • Dysmenorrhoea NRS score at baseline (<7 / >=7) • Smoking history (Current smoker / former smoker / never smoked) • Alcohol use (None / Moderate) • AFSE stage (I / II / III / IV / Unknown) • Renal Function (>=60 - <90 mL/min / >=90mL/min)	• Time since endometriosis diagnosis (<2 years / 2-5 years / >= 5 years) • Age (<35 years / >= 35 years) • Age (<30 years / 30-35 years / 35-40 years / >=40 years) • Race (Black or African American / White) • BMI at baseline (<25 / 25-30 / >=30) • NMPP NRS score at baseline (>4 / 4-7 / 7-10) • Smoking history (Current smoker / former smoker / never smoked) • Alcohol use (None / Moderate) • AFSE stage (I / II / III / IV / Unknown) • Renal Function (>=60 - <90 mL/min / >=90mL/min)	• BMI at baseline (<25 / 25-30 / >=30)

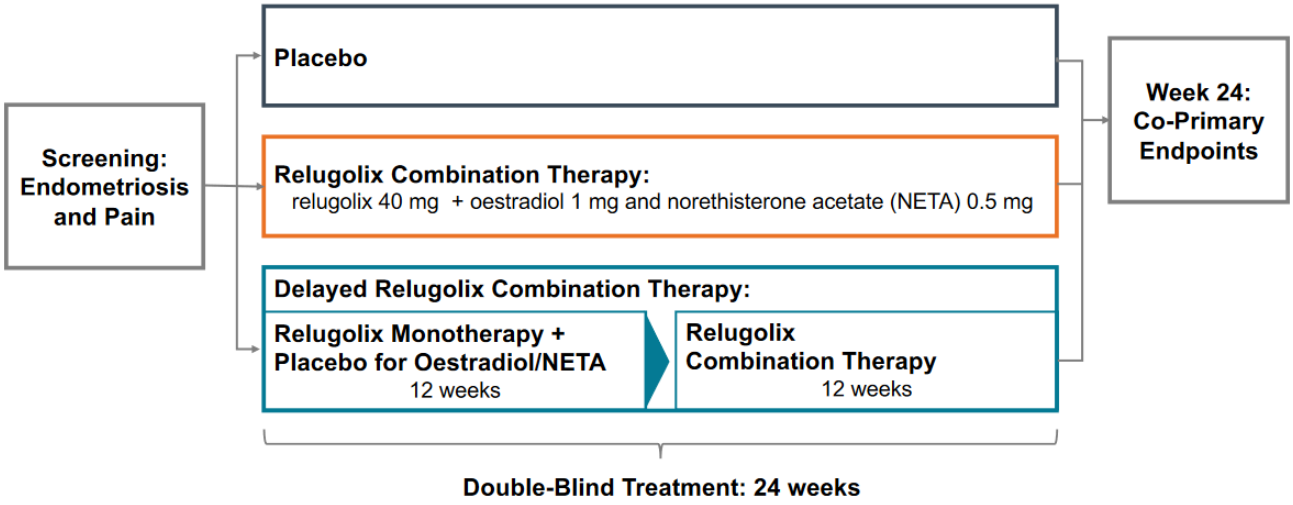
© Gedeon Richter UK Ltd (2023). All rights reserved

SPIRIT 1 and SPIRIT 2

Trial design

SPIRIT 1 and 2 were two replicate, phase 3, multicentre, randomised, double-blind, placebo-controlled efficacy and safety trials conducted between July 2017 and June 2021. Their aim was to determine the benefit and safety of Relugolix CT compared with placebo for 24 weeks on dysmenorrhoea and on non-menstrual pelvic pain (NMPP). The co-primary endpoints for both trials were proportion of responders at Week 24/EOT based on dysmenorrhoea NRS scores vs placebo, and the proportion of responders at Week 24/EOT based on non-menstrual pelvic pain NRS scores vs placebo (37).

Figure 5: Study design for SPIRIT 1 and 2 (37)



SPIRIT 1 and 2 had the same trial design with overlapping geographical regions (Figure 5). The only difference between them was the inclusion of Week 24/EOT endometrial biopsies and pharmacokinetic sampling in SPIRIT 1 only (37). Participants were randomly assigned in a 1:1:1 ratio to placebo, Relugolix CT, or relugolix + delayed CT therapy for 24 weeks. Relugolix 40 mg and relugolix placebo were supplied to the study site in blister cards co-packaged with the oestradiol/norethisterone acetate or oestradiol/norethisterone acetate placebo. The relugolix + delayed CT group received the 40 mg relugolix tablet and a placebo capsule for 12 weeks, followed by the active agent tablet and capsule for 12 weeks.

© Gedeon Richter UK Ltd (2023). All rights reserved

The relugolix + delayed CT group was included to allow for the comparison of bone mineral density and vasomotor symptoms in the combination and monotherapy groups at week 12. Study visits occurred every 4 weeks through the end of Week 24 during the Randomised Treatment Period (37).

Patient population and baseline characteristics in SPIRIT 1 and SPIRIT 2

The study population included premenopausal women aged 18 to 50 years old with endometriosis associated pain. Key inclusion and exclusion criteria are detailed in Table 6.

Table 6: Key inclusion and exclusion criteria for SPIRIT 1&2

Inclusion	Exclusion
Is a premenopausal female aged 18 to 50 years old (inclusive) on the day of signing of the informed consent form.	Has a history of chronic pelvic pain that is not caused by endometriosis.
Has agreed to use only study-specified analgesic medications during the study and is not known to be intolerant to these.	Has any chronic pain or frequently recurring pain condition, other than endometriosis that is treated with opioids or requires analgesics for ≥ 7 days per month.
Has a diagnosis of endometriosis and has had, within 10 years prior to signing the informed consent form, surgical or direct visualization and/or histopathologic confirmation of endometriosis, for example, during a laparoscopy or laparotomy.	Has had surgical procedures for treatment of endometriosis within the 3 months prior to the Screening visit.
During the Run-In Period (35 to 70 days prior to treatment period) has a dysmenorrhoea NRS score ≥ 4.0 on at least 2 days and Mean NMPP NRS score ≥ 2.5, or	Has a history of or currently has osteoporosis or other metabolic bone disease.

© Gedeon Richter UK Ltd (2023). All rights reserved

Mean NMPP NRS score ≥ 1.25 and NMPP NRS score ≥ 5.0 on ≥ 4 days.	
	Has a clinically significant gynaecologic condition, other than endometriosis, identified during Screening or Run-In period transvaginal ultrasound or endometrial biopsy.

A total of 638 patients were randomised in SPIRIT 1 and 623 were randomised in SPIRIT 2.

Efficacy and safety analyses were performed using the modified Intent-to-Treat (mITT) population, unless otherwise specified. The mITT population comprised randomised patients who received any amount of study drug (relugolix, oestradiol, norethisterone acetate or placebo). Efficacy analyses were performed by treatment group as randomised. Safety data were analysed by treatment group according to the actual treatment received (not the randomised treatment) (37, 62, 63).

In SPIRIT 1, 3 patients (2 in the relugolix + delayed oestradiol /NETA group and 1 in the placebo group) were randomised but did not receive study drug. In SPIRIT 2, 7 patients were randomised but did not receive the study drug. Of these, 6 were excluded due to data integrity concerns at a study site (5 in the delayed Relugolix CT group and 1 in the Relugolix CT group), and 1 patient from the placebo group was randomised in error (62, 63).

A summary of the randomised, mITT and safety population numbers is provided in Table 7.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 7:Number of study participants in SPIRIT 1 and SPIRIT 2 (62, 63)

	SPIRIT 1	SPIRIT 2	Total
Randomised	**N=638*** Placebo (N=213) Relugolix + delayed CT (N=213) Relugolix CT (N=212)	**N=623**** Placebo (N=208) Relugolix + delayed CT (N=207) Relugolix CT (N=208)	**N=1261** Placebo (N=421) Relugolix + delayed CT (N=420) Relugolix CT (N=420)
mITT population	**N=635** Placebo (N=212) Relugolix + delayed CT (N=211) Relugolix CT (N=212)	**N=616** Placebo (N=204) Relugolix + delayed CT (N=206) Relugolix CT (N=206)	**N=1251** Placebo (N=416) Relugolix + delayed CT (N=417) Relugolix CT (N=418)
Safety population	**N=635** Placebo (N=212) Relugolix + delayed CT (N=211) Relugolix CT (N=212)	**N=616** Placebo (N=204) Relugolix + delayed CT (N=206) Relugolix CT (N=206)	**N=1251** Placebo (N=416) Relugolix + delayed CT (N=417) Relugolix CT (N=418)

*In SPIRIT 1, 3 randomised patients were randomised in error as they had not met all the eligibility requirements. These patients did not receive any study treatment and were therefore not included in either the modified Intent-to-Treat (mITT) Population or the Safety Population.
**In SPIRIT 2 7 patients were randomised but not included in the modified Intent-to-Treat (mITT) Population or Safety Populations. 6 of these were excluded due to data integrity concerns at the study site, and 1 patient was randomised in error.

© Gedeon Richter UK Ltd (2023). All rights reserved

The baseline characteristics of patients in SPIRIT 1 and SPIRIT 2 are shown in Table 8.

Table 8: Patient characteristics for SPIRIT 1 and SPIRIT 2 (mITT population) (37)

	SPIRIT 1			**SPIRIT 2**		
	Relugolix CT (n=212)	**Placebo (n=212)**	**Relugolix + delayed CT (n=211)**	**Relugolix CT (n=206)**	**Placebo (n=204)**	**Relugolix + delayed CT (n=206)**
Age, years, mean (SD)	33·9 (6·3)	34.2 (6.6)	34·3 (6·7)	33·8 (6·7)	33·6 (6·5)	33·7 (6·8)
Body mass index, mean (SD)	25·6 (6·0)	26.1 (6.4)	25·7 (6·1)	26·1 (6·5)	25·8 (6·0)	26·2 (5·9)
Race, n (%)						
White	194 (92%)	193 (91%)	194 (92%)	186 (90%)	183 (90%)	188 (91%)
Black	13 (6%)	12 (6%)	10 (5%)	14 (7%)	12 (6%)	10 (5%)
Other	5 (2%)	7 (3%)	7 (3%)	6 (3%)	9 (4%)	8 (4%)

The disease specific characteristics of patients in the mITT populations for SPIRIT 1 and SPIRIT 2 are presented in Table 9. The median time since surgical diagnosis was 3.2 years and 3.5 years for the Relugolix CT arms in SPIRIT 1 and SPIRIT 2, respectively. Overall, disease specific baseline characteristics were consistent with a population of women with endometriosis.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 9: Baseline disease specific characteristics for patients in the mITT populations of SPIRIT 1 and SPIRIT 2 (37)

Characteristics		SPIRIT 1			SPIRIT 2		
		Relugolix CT (N=212)	Placebo (N=212)	Relugolix + delayed CT (N=211)	Relugolix CT (N=206)	Placebo (N=204)	Relugolix + delayed CT (N=206)
Time since surgical diagnosis of endometriosis, years	Mean (SD)	3.8 (3.2)	3.8 (3.3)	4.4 (4.1)	4.1 (3.5)	3.8 (3.0)	4.2 (3.5)
	<5 years	151 (71%)	148 (70%)	135 (64%)	137 (67%)	143 (70%)	135 (66%)
	5–10 years	61 (29%)	64 (30%)	76 (36%)	69 (33%)	61 (30%)	71 (34%)
Bone mineral density, z-score	Lumbar spine	0.17 (1.1)	0.18 (1.1)	0.18 (1.1)	0.23 (1.1)	0.35 (1.0)	0.25 (1.1)
	Total hip	–0.01 (0.9)	0.05 (0.9)	0.05 (0.9)	0.1 (1.0)	0.12 (1.0)	0.06 (1.0)
Dysmenorrhoea NRS score	Mean (SD)	7.2 (1.7)	7.1 (1.7)	7.0 (1.8)	7.1 (1.6)	7.0 (1.6)	6.9 (1.5)
	<7	84 (40%)	90 (43%)	97 (46%)	92 (45%)	96 (47%)	97 (47%)
	≥7	128 (60%)	122 (58%)	114 (54%)	114 (55%)	108 (53%)	109 (53%)
Non-menstrual pelvic pain NRS score	Mean (SD)	5.9 (2.0)	5.8 (1.8)	5.6 (2.0)	5.8 (1.9)	5.5 (1.9)	5.5 (1.9)
	<4	43 (20%)	43 (20%)	53 (25%)	42 (20%)	45 (22%)	55 (27%)
	≥4	169 (80%)	169 (80%)	158 (75%)	164 (80%)	159 (78%)	151 (73%)
Dyspareunia NRS score	Mean (SD)	5.7 (2.3)	5.7 (2.3)	5.3 (2.4)	5.5 (2.3)	5.3 (2.3)	5.4 (2.1)
	<7	112/174 (64%)	113/165 (68%)	126/176 (72%)	127/173 (73%)	131/162 (81%)	129/167 (77%)
	≥7	62/174 (36%)	52/165 (32%)	50/176 (28%)	46/173 (27%)	31/162 (19%)	38/167 (23%)
EHP-30 pain domain	Mean (SD)	58.3 (16.7)	55.5 (16.0)	55.5 (16.8)	56.2 (17.1)	55.0 (16.2)	55.5 (15.2)
	<50	60/208 (29%)	67/208 (32%)	70/208 (34%)	62/203 (31%)	74 (36%)	62 (30%)
	≥50	148/208 (71%)	141/208 (67%)	138/208 (66%)	141/203 (69%)	130 (64%)	144 (70%)
Analgesic use	Only non-opioids	128 (60%)	137 (65%)	124 (59%)	97 (47%)	97 (48%)	94 (46%)
	Opioids	64 (30%)	52 (26%)	65 (31%)	100 (49%)	95 (47%)	101 (49%)

© Gedeon Richter UK Ltd (2023). All rights reserved

Study sites in SPIRIT 1 and SPIRIT 2

SPIRIT 1 involved 124 centres in the USA, Canada, Argentina, Belgium, Bulgaria, Czech Republic, Finland, Hungary, Poland, Portugal, South Africa, Spain, and Ukraine (62).

SPIRIT 2 involved 95 centres in the USA, Australia, Brazil, Chile, Czech Republic, Georgia, Italy, New Zealand, Poland, Romania, and Sweden (63).

Trial interventions in SPIRIT 1 and SPIRIT 2

In SPIRIT 1 and SPIRIT 2 patients were randomised in a 1:1:1 ratio to receive either:

- Relugolix 40mg co-administered with oestradiol 1 mg/NETA 0.5 mg for 24 weeks (Relugolix CT).
- Relugolix 40mg co-administered with oestradiol 0 mg/NETA 0 mg placebo for 12 weeks, followed by relugolix co-administered with oestradiol 1 mg/NETA 0.5 mg for 12 weeks.
- Relugolix placebo co-administered with oestradiol 0 mg /NETA 0mg placebo for 24 weeks.

All treatments were administered orally once daily (37).

Placebo versions of Relugolix CT were designed to match their experimental counterpart in size, shape, and colour (62, 63).

Outcomes – primary endpoint in SPIRIT 1 and SPIRIT 2 (37)

The objective of the SPIRIT 1 and SPIRIT 2 trials was to determine the benefit of Relugolix CT compared to placebo for 24 weeks on endometriosis associated non-menstrual pelvic pain and dysmenorrhoea. The primary endpoints were:

- Proportion of patients who meet the dysmenorrhoea responder criteria at the Week 24/EOT pain assessment period, achieving a mean reduction in dysmenorrhoea NRS scores of at least 2.8 points and no increase in use of analgesic medications as recorded in a daily eDiary

© Gedeon Richter UK Ltd (2023). All rights reserved

- Proportion of patients who meet the NMPP responder criteria at the Week 24/EOT pain assessment period, achieving a mean reduction in NMPP NRS scores of at least 2.1 points and no increase in use of analgesic medications as recorded in a daily eDiary.

Outcomes – Secondary endpoints in SPIRIT 1 and SPIRIT 2

Key secondary outcomes at week 24 are shown in Table 10 (37).

Table 10 Key secondary endpoints in SPIRIT 1 and SPIRIT 2 (37)

Objective	Endpoint
To determine the benefit on function measured by the Endometriosis Health Profile-30 pain domain	Change from baseline to week 24 in the Endometriosis Health Profile-30 pain domain score
To determine the benefit on dysmenorrhoea measured by the NRS	Change from baseline to week 24/end of treatment in the mean dysmenorrhoea NRS score
To determine the benefit on non-menstrual pelvic pain measured by the NRS	Change from baseline to week 24/end of treatment in the mean non-menstrual pelvic pain NRS score
To determine the benefit on overall pelvic pain measured by the NRS	Change from baseline to week 24/end of treatment in the mean overall pelvic pain NRS score
To determine the benefit on dyspareunia measured by the NRS	Change from baseline to week 24/end of treatment in the mean dyspareunia NRS score
To determine the benefit on protocol-specified opioid use (Tier 2) for endometriosis-associated pain as recorded in the electronic diary	Proportion of patients who were not using protocol-specified opioids for endometriosis-associated pain at week 24/end of treatment
To determine the benefit on protocol-specified analgesic use (Tier 1 and Tier 2) for endometriosis-associated pain as recorded in the electronic diary	Proportion of patients who were not using protocol-specified analgesics for endometriosis-associated pain at week 24/end of treatment (for SPIRIT 1) Change from baseline to week 24/end of treatment in protocol-specified analgesic use for endometriosis-associated pain based on mean pill count (for SPIRIT 2)

Safety was evaluated by monitoring adverse events, clinical laboratory data, 12-lead electrocardiograms (ECGs), vital signs, physical examinations, menstrual bleeding

© Gedeon Richter UK Ltd (2023). All rights reserved

patterns, pregnancy, overdose, BMD, and paired endometrial biopsies (37). More detail is provided in Table 11.

Table 11 Safety endpoints in the SPIRIT 1 and SPIRIT 2 trials (62, 63)

Objective	Endpoint
To determine the safety of 24 weeks of Relugolix CT or relugolix + delayed CT	Treatment-emergent adverse events, change in vital signs (including weight), clinical laboratory tests, ECGs, BMD by DXA, and EMBs
To determine the percent change from baseline to Week 12 in BMD at the lumbar spine (L1-L4) in Relugolix CT compared with relugolix + delayed CT	Percent change from baseline to Week 12 in BMD at the lumbar spine (L1-L4) as assessed by DXA
To determine the change in BMD after 24 weeks of treatment with Relugolix CT or relugolix + delayed CT	Percent change from baseline to Week 24 in BMD at the lumbar spine (L1-L4), femoral neck, and total hip as assessed by DXA
To determine the incidence of vasomotor symptoms with Relugolix CT compared with relugolix + delayed CT through Week 12	Incidence of vasomotor symptoms at Week 12

CT – Combination therapy, BMD – bone marrow density, DXA – dual-energy x-ray absorptiometry, EMB – endometrial biopsy, ECG - electrocardiogram

SPIRIT OLE

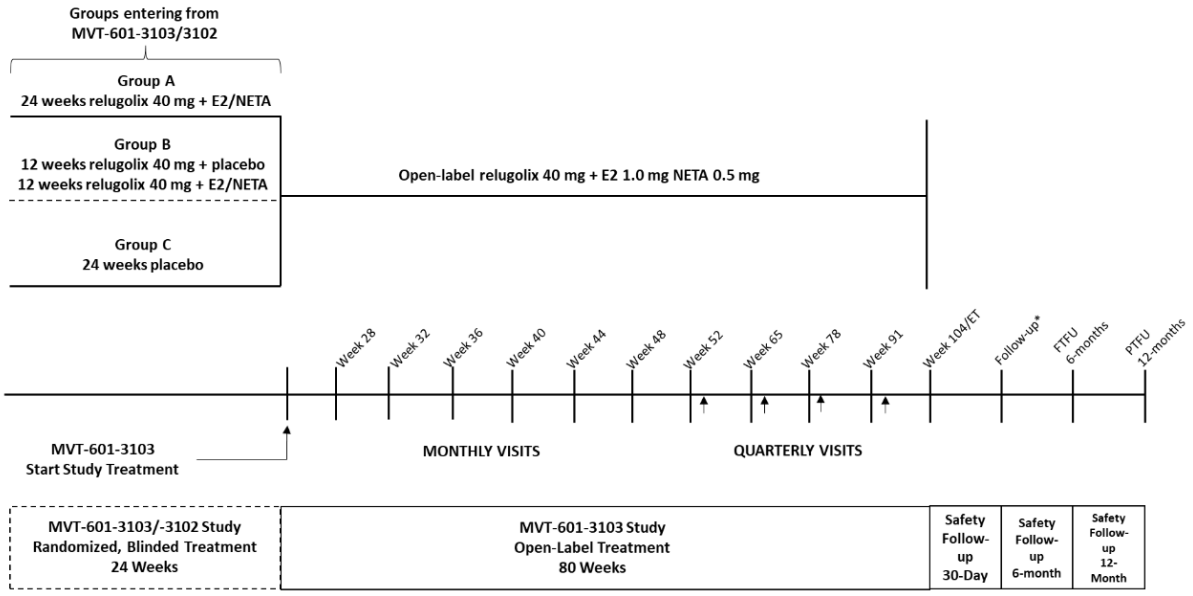
Trial design

The SPIRIT OLE study was an international phase 3, open-label, single-arm, long-term efficacy and safety extension study that enrolled eligible patients who completed the 24-week treatment period in SPIRIT 1 or SPIRIT 2. Participants received oral Relugolix CT once a day for up to 80 weeks. Therefore, over the entire trial period (i.e. SPIRIT 1 or 2 plus SPIRIT OLE), there were three possible treatment combinations: Relugolix CT for the whole trial period, placebo in SPIRIT 1 or SPRIT 2 and Relugolix CT in SPIRIT OLE or relugolix + delayed CT in SPIRIT 1 and SPIRIT 2 and Relugolix CT in SPIRIT OLE (64).

Baseline visit procedures for SPIRIT OLE were carried out at the same time as the Week 24 visits in the parent studies (hereafter referred to as the Week 24/Baseline visit). However, the baseline for analyses was the baseline in the parent studies.

© Gedeon Richter UK Ltd (2023). All rights reserved

Figure 6 SPIRIT OLE Study schematic (64)



Patient population and baseline characteristics for SPIRIT OLE

The key inclusion and exclusion criteria for SPIRIT OLE are described in Table 12.

Table 12: Key inclusion and exclusion criteria for SPIRIT OLE (64)

Inclusion	Exclusion
Completed 24 weeks of study drug treatment and study participation in either SPIRIT 1 or SPIRIT 2	Had a surgical procedure for treatment of endometriosis at any time during the parent study (SPIRIT 1 or SPIRIT 2)
Voluntarily signed and dated the informed consent form prior to initiation of any study specific procedures for SPIRIT OLE	Any chronic pain or frequently recurring pain condition, other than endometriosis, that is treated with opioids or requires analgesics for ≥7 days per month
Not expected to undergo gynecological surgery or other surgical procedures for treatment of endometriosis (including ablation, shaving, or excision) during the study, including during the Follow-Up Period, and the patient does not desire such treatment during this time frame	Z-score <-2.0 or a ≥7% decrease in BMD from the parent study Baseline at lumbar spine, total hip, or femoral neck based on the parent study Week 24 DXA assessment of BMD
Agreed to continue to use acceptable non-hormonal contraceptive methods during the Open-Label Treatment Period and for at least 30 days after the last dose of study drug. However, the patient is not required to use the specified nonhormonal contraceptive methods if she • Has a sexual partner(s) who was vasectomized at least 6 months prior to the Week 24/Baseline visit;	Any contraindication to treatment with low-dose estradiol and norethisterone acetate, including • Known, suspected, or history of breast cancer; • Known or suspected estrogen-dependent neoplasia; • Active deep vein thrombosis or pulmonary embolism, or history of these

© Gedeon Richter UK Ltd (2023). All rights reserved

• Had a bilateral tubal occlusion (including ligation and blockage methods such as Essure™), at least 6 months prior to the Week 24/Baseline visit (patients with Essure have to have prior confirmation of tubal occlusion by hysterosalpingogram) and there must be no evidence of post-Essure syndrome; • Has a non-hormonal intrauterine device (e.g. Paragard®) placed in the uterus; • Is not sexually active with men; periodic sexual relationship(s) with men requires the use of non-hormonal contraception as noted above; • Practices total abstinence from sexual intercourse, as her preferred lifestyle; periodic abstinence is not acceptable.	conditions prior to the Week 24/Baseline visit; • History of or active arterial thromboembolic disease, including stroke and myocardial infarction; • Known anaphylactic reaction or angioedema or hypersensitivity to estradiol or norethisterone acetate; • Known protein C, protein S, or antithrombin deficiency, or other known thrombophilia disorders, including Factor V Leiden; • Migraine with aura; • History of porphyria
Agreed to continue to use only study-specified analgesic medications during the study and is not known to be intolerant to these	Any of the following clinical laboratory abnormalities at the parent study Week 20 visit or, if available, any subsequent visit in one of the parent studies (SPIRIT 1 or SPIRIT 2) • Alanine aminotransferase or aspartate aminotransferase >2.0 times the upper limit of normal (ULN); or • Bilirubin (total bilirubin) >1.5 x ULN (or >2.0 x ULN if secondary to Gilbert syndrome or pattern consistent with Gilbert syndrome)
Negative urine pregnancy test at the Week 24/Baseline visit	

A total of 802 patients were enrolled in SPIRIT OLE, which represents 77% of patients who completed the parent studies. The full patient disposition for the SPIRIT OLE is provided in Appendix D1.2.

Efficacy analyses were carried out on the extension study population and safety analyses on the extension safety population. Both populations were defined as all patients who enrolled and received any amount of open-label study drug in SPIRIT OLE. Efficacy analyses were performed by treatment group as randomised in the parent trials. Safety data were analysed by pivotal phase 3 study treatment group according to the actual treatment received (not the randomised treatment in the parent study) (64).

© Gedeon Richter UK Ltd (2023). All rights reserved

The extension study population and the extension safety populations included 799 of the 802 enrolled patients; three patients from a single site were excluded owing to Good Clinical Practice (GCP) non-compliance.

The baseline characteristics of patients in SPIRIT OLE are shown in Table 13, categorised by their parent study treatment group.

Table 13: Patient characteristics for SPIRIT OLE (extension study population) (64, 65)

Characteristics		**SPIRIT OLE (Long Term Extension)**		
		Relugolix CT (N=277)	**Placebo → Relugolix CT (N=275)**	**Relugolix + delayed CT → Relugolix CT (N=247)**
Age	Mean	34.1	34.3	35.1
	SD	6.55	6.48	6.49
	<35 years	142 (51.3%)	136 (49.5%)	114 (46.2%)
	≥35 years	135 (48.7%)	139 (50.5%)	133 (53.8%)
Race	White	254 (91.7%)	248 (90.2%)	236 (95.5%)
	Black or African American	17 (6.1%)	13 (4.7%)	7 (2.8%)
	Asian	0	0	1 (0.4%)
	American Indian or Alaska Native	1 (0.4%)	0	1 (0.4%)
	Native Hawaiian or Other Pacific Islander	0	1 (0.4%)	0
	Multiple	4 (1.4%)	5 (1.8%)	2 (0.8%)
	Other	1 (0.4%)	8 (2.9%)	0
Ethnicity	Not Hispanic or Latino	249 (89.9%)	233 (87.0%)	215 (87.0%)
	Hispanic or Latino	27 (9.7%)	42 (15.3%)	31 (12.6%)
Time since surgical diagnosis	Mean	4.0	3.9	4.7
	SD	3.5	3.2	4.0
Dysmenorrhoea NRS score	Mean	7.1	7.2	7.0
	SD	1.7	1.6	1.7
NMPP NRS score	Mean	5.7	5.7	5.5
	SD	1.9	1.9	2.0
Dyspareunia NRS score	Mean	5.5	5.4	5.2
	SD	2.5	2.6	2.4
EHP-30 Pain domain score	Mean	57.3	56.2	56.2
	SD	17.1	14.9	16.4

Study sites (64)

SPIRIT OLE was conducted at 169 locations in the USA, Argentina, Australia, Belgium, Brazil, Bulgaria, Canada, Chile, Chechia, Finland, Georgia, Hungary, Italy, New Zealand, Poland, Portugal, Romania, South Africa, Spain, and Ukraine.

© Gedeon Richter UK Ltd (2023). All rights reserved

Trial interventions (64)

Relugolix CT for 80 weeks for all patients enrolled regardless of parent study group.

Outcomes (64)

The primary efficacy objectives were to evaluate the long-term efficacy of Relugolix CT once daily (OD) on endometriosis-associated pain at 52 weeks and 104 weeks among patients who previously completed a 24-week treatment period in one of the pivotal studies. The co-primary endpoints of the study were:

- Proportion of women who respond or maintain response based on assessment of dysmenorrhoea at Week 52 and Week 104. Assessed using a NRS score (11-point scale) for pain recorded daily in an e-Diary.
- Proportion of women who respond or maintain response based on assessment of NMPP at Week 52 and Week 104. Assessed using a NRS score (11-point scale) for pain recorded daily in an e-Diary.

Secondary endpoints included the change from baseline in the mean dysmenorrhoea, NMPP and pelvic pain scores. A summary of the key secondary endpoints is displayed in Table 14.

Table 14: Key secondary endpoints in SPIRIT OLE

Objectives	**Endpoints**
EHP-30 pain domain score	• Change from baseline in the EHP-30 pain domain score at Week 52 and Week 104
Dysmenorrhoea NRS score	• Change from baseline in the mean dysmenorrhoea NRS score at Week 52 and Week 104
NMPP score	• Change from baseline in the mean NMPP score at Week 52 and Week 104
Pelvic pain NRS score	• Change from baseline in the mean overall pelvic pain NRS score at Week 52 and Week 104
Dyspareunia NRS score	• Change from baseline in the mean dyspareunia NRS score at Week 52 and Week 104
Opioid use	• Proportion of patients not using protocol-specified opioids for Endometriosis-associated pain at Week 52 and Week 104
Analgesics use	• Proportion of patients not using analgesics for Endometriosis-associated pain at Week 52 and Week 104

© Gedeon Richter UK Ltd (2023). All rights reserved

Safety was evaluated by monitoring adverse events, clinical laboratory data, 12-lead ECGs, vital signs and weight, physical examinations, menstrual bleeding patterns, pregnancy, overdose, endometrial biopsies, mammograms, and BMD.

B.2.4 Statistical analysis and definition of study groups in the relevant clinical effectiveness evidence

A summary of the statistical analyses for SPIRIT 1, SPIRIT 2 and SPIRIT OLE is available in Table 15. An overview of the key aspects for each trial then follows.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 15: Summary of statistical analyses SPIRIT studies (62-64)

Trial	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
SPIRIT 1 & 2	The primary hypothesis tested for each co-primary endpoint in this study was that Relugolix CT for 24 weeks was superior to placebo. A logistic regression model was used to compare Relugolix CT with placebo for each pain measure (dysmenorrhoea or NMPP). The responder status (responder versus non-responder) was the dependent variable, treatment was the main effect, baseline pain score (dysmenorrhoea or NMPP) and stratification factors were the covariates.	The co-primary and ranked secondary efficacy analyses were performed at an overall alpha level of 0.05 (2-sided) comparing Relugolix CT with placebo. Logistic regression model	The planned sample size was planned to be approximately 600 patients per trial. (randomised 1:1:1) Actual sample size was 1251	For the primary analysis missing data handling rules were implemented for patients with missing treatment duration and pain score data at 24 weeks. A mixed-effects model approach is used to impute missing data for the primary analysis.
SPIRIT OLE	The primary efficacy endpoints for the SPIRIT OLE study were defined in a manner analogous to the co-primary endpoints for the pivotal studies SPIRIT 1 and SPIRIT 2.	The responder rate and two-sided 95% CI was presented by pivotal phase 3 study treatment group. No treatment comparisons were performed for this extension study	Expected sample size was 800 patients (67% of the total planned 1200 patients for the parent studies) Actual sample size 802. 501 patients completed to week 104	For the evaluation of the primary endpoint, missing data handling rules are implemented to derive responder status at Week 104/EOT.

© Gedeon Richter UK Ltd (2023). All rights reserved

SPIRIT 1 and SPIRIT 2

Efficacy analyses were performed using the modified Intent-to-Treat (mITT) Population, unless otherwise specified. The mITT Population was defined as all randomised patients who received any amount of study drug (relugolix/placebo or oestradiol/NETA/placebo). Efficacy analyses were performed by treatment group as randomised.

The randomisation ratio was 1:1:1 among the treatment groups: Relugolix CT, relugolix + delayed CT, and placebo. Randomisation was conducted centrally and stratified by geographic region and years since the diagnosis of endometriosis by direct surgical or laparoscopic visualisation as follows:

- Geographic region: North America versus Rest of World
- Years since endometriosis diagnosis: < 5 or ≥ 5 years

Primary efficacy analysis

The primary hypothesis tested for each co-primary endpoint in this study was that Relugolix CT for 24 weeks was superior to placebo. A logistic regression model was used to compare Relugolix CT with placebo for each pain measure (dysmenorrhoea or NMPP). The responder status (responder versus non-responder) was the dependent variable, treatment was the main effect, baseline pain score (dysmenorrhoea or NMPP) and stratification factors were the covariates (62, 63).

The threshold of a clinically meaningful response was determined for dysmenorrhoea and NMPP separately, utilizing the anchor-based cumulative distribution function/probability density function method considering the PGA for dysmenorrhoea and NMPP, respectively, as the anchors, using pooled blinded data from SPIRIT 1 and SPIRIT 2 studies (approximately 200 patients from each study). Results from a patient exit interview substudy were also available and considered as supportive information in the threshold determinations. These thresholds were pre-specified in the study protocol (62, 63).

The comparison for each co primary endpoint (dysmenorrhoea or non-menstrual pelvic pain) was done using a logistic regression model with responder status as a dependent variable, treatment as the main effect, baseline pain score

© Gedeon Richter UK Ltd (2023). All rights reserved

(dysmenorrhoea or non-menstrual pelvic pain), and the stratification factors (geographical region [North America vs all other regions]; years since surgical endometriosis diagnosis [<5 years vs ≥5 years]) as covariates (37).

Analyses of the co primary endpoints for each study were done at an overall α level of 0·05 (two sided) comparing Relugolix CT with placebo (37).

Key secondary efficacy analyses

A fixed sequence testing procedure was used to maintain the family wise type I error rate by testing the co primary and key secondary endpoints sequentially. In each study, the two co primary endpoints were tested first, and if the p value was less than 0·05 for both co primary endpoints, the seven key secondary efficacy endpoints were tested sequentially per the testing procedure for the study (37).

Statistical methods: safety

The safety population was the same as the mITT population and is defined as all randomised patients who have received any amount of study drug (37). Safety data were analysed by treatment group according to the actual treatment received (not the randomised treatment).

Safety assessments included treatment-emergent adverse events, vital signs, physical examinations, clinical laboratory tests, 12-lead ECGs, BMD, and endometrial biopsies. Safety analyses were based on all randomised patients who received any amount of randomised study drug (Safety Population). Drug exposure was summarized by descriptive statistics. Severity of all treatment-emergent adverse events was evaluated by the investigator based on the National Cancer Institute's Common Terminology Criteria for Adverse Events (NCI CTCAE) version 5.0 and were coded to preferred term, higher level term, and system organ class using MedDRA version 22.0. The number and percentage of patients with adverse events was summarised by MedDRA system organ class and preferred term, relationship to study drug, and severity (62, 63).

Sample size

In each study, a sample size of 200 patients per treatment group was planned to provide more than 90% power to detect a difference of 20% or more in each co

© Gedeon Richter UK Ltd (2023). All rights reserved

primary endpoint between the Relugolix CT and placebo groups, assuming a placebo responder rate of 30–35% (based on a range of responder rates observed in similar phase 3 endometriosis trials), and a dropout rate of 20%,8 at a two sided α level of 0·05. (randomised 1:1:1). The actual sample size across both studies was 1251.

Handling of dropouts or missing data (37)

Missing data handling rules were implemented for deriving responder status over the last 35 days of treatment (week 24), considering duration of treatment exposure and compliance with pain score entry on the daily electronic diary.

Patients who completed <5 weeks of treatment were considered non-responders for both dysmenorrhoea and non-menstrual pelvic pain.

SPIRIT OLE

Statistical methods (64)

Efficacy and safety data were analysed using descriptive statistics by the originally randomised treatment groups. There were no between-group comparisons.

The point estimate and 2-sided 95% CI for the primary efficacy endpoints were calculated for each treatment group.

The pivotal study baseline visit was used as the reference point for all change from baseline-related endpoints. The pain scores during the baseline pain assessment period of the pivotal study established the patient's baseline for both the pivotal and extension studies. No formal treatment comparisons were performed for this extension study.

Sample size (64)

As SPIRIT OLE was an extension study, the sample size was determined by the numbers of patients who completed either parent study and who were eligible and willing to participate in the extension study. It was estimated that approximately 800 patients (67% of the total planned 1200 patients who had completed either parent study SPIRIT 1 and SPIRIT 2) would be enrolled.

© Gedeon Richter UK Ltd (2023). All rights reserved

Participant flow

Details of participant flow through SPIRIT 1, SPIRIT 2, and SPIRIT OLE and the comparator trials included in the ITC are provided in Appendix D1.2.

B.2.5 Critical appraisal of the relevant clinical effectiveness evidence

SPIRIT 1, SPIRIT 2, and SPIRIT 3 were assessed for quality using the York Centre for Reviews and Dissemination guidance for undertaking reviews in healthcare. The summary of the findings is presented in Table 16 with more detailed results in Appendix D1.3.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 16: Summary of the quality assessment results

Trial number (acronym)	MVT-601-3101 (SPIRIT 1)	MVT-601-3102 (SPIRIT 2)	MVT-601-3103 (SPIRIT OLE)
Was randomisation carried out appropriately?	Yes	Yes	N/A
Was the concealment of treatment allocation adequate?	Yes	Yes	N/A
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes	Yes	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes	Yes	N/A
Were there any unexpected imbalances in drop-outs between groups?	No	No	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes	Yes	No
Was there good quality assurance for this study?	Yes	Yes	Yes
Adapted from Systematic reviews: CRD's guidance for undertaking reviews in health care (University of York Centre for Reviews and Dissemination)			

The SPIRIT 1 and SPIRIT 2 trials were good quality, robust RCTs that included randomisation, appropriate blinding of groups without any imbalances in the dropouts between groups, nor evidence to suggest any measurement of more outcomes than reported. As an open-label extension trial, SPIRIT OLE also maintained good quality standards. Randomisation and blinding were not applicable to this open-label study, however, randomisation was performed in the parent study

© Gedeon Richter UK Ltd (2023). All rights reserved

trials from which participants were enrolled, thus minimising any bias due to treatment allocation. Differences in dropouts between groups were fully documented and reported.

As detailed in Section B.2.1, the eligibility criteria of the SPIRIT studies ensured that the study population was balanced and a good representation of women with symptomatic endometriosis and significant disease burden who are likely to be treated in clinical practice.

B.2.6 Clinical effectiveness results of the relevant studies

Efficacy results from SPIRIT 1, SPIRIT 2, and SPIRIT OLE are described in this section. Note that the focus is on data from participants in the Relugolix CT and placebo groups are as these are the only study arms that are relevant to the submission population and are included in the economic model. For completeness, data from the relugolix + delayed CT group are shown in graphs, but these data are not discussed in the text.

© Gedeon Richter UK Ltd (2023). All rights reserved

SPIRIT 1 and SPIRIT 2

A summary of the results for the key efficacy endpoints in SPIRIT 1 and SPIRIT 2 is shown in Table 17.

Table 17: Results for key efficacy endpoints in SPIRIT 1 and SPIRIT 2 (37)

Endpoint **CO-PRIMARY EFFICACY ENDPOINTS**		**SPIRIT 1**	**SPIRIT 2**
Proportion of Patients Classified as Dysmenorrhoea Responders at Week 24/EOT	Relugolix CT vs Placebo	75% vs 27%	75% vs 30%
	Difference	47.6%	44.9%
	95% CI	(39.3%, 56.0%)	(36.2%, 53.5%)
	p-value	<0.0001	<0.0001
Proportion of Patients Classified as Non-Menstrual Pelvic Pain Responders at Week 24/EOT	Relugolix CT vs Placebo	59% vs 40%	66% vs 43%
	Difference	18.9%	23.4%
	95% CI	(9.5%, 28.2%)	(14.0%, 32.8%)
	p-value	<0.0001	<0.0001
SECONDARY EFFICACY ENDPOINTS			
1. Change from baseline to Week 24 in the EHP-30 Pain Domain score	Relugolix CT vs Placebo	-33.8 vs -18.7	-32.2 vs -19.9
	Difference	-15.1	-12.3
	95% CI	(-19.7, -10.5)	(-16.7, -7.9)
	p-value	<0.0001	<0.0001
2. Change from baseline to Week 24/EOT in the mean dysmenorrhoea NRS score	Relugolix CT vs Placebo	-5.1 vs -1.8	-5.1 vs -2.0
	Difference	-3.3	-3.2
	95% CI	(-3.8, -2.8)	(-3.7, -2.7)
	p-value	<0.0001	<0.0001
3. Change from baseline to Week 24/EOT in the mean NMPP NRS score	Relugolix CT vs Placebo	-2.9 vs -2.0	-2.7 vs -2.0
	Difference	-0.9	-0.7
	95% CI	(-1.4, -0.4)	(-1.2, -0.3)
	p-value	0.0002	<0.0001
4. Change from baseline to Week 24/EOT in the mean overall pelvic pain NRS score	Relugolix CT vs Placebo	-3.1 vs -1.9	-2.9 vs -2.0
	Difference	-1.1	-0.9
	95% CI	(-1.6, -0.7)	(-1.4, -0.5)
	p-value	<0.0001	<0.0001
5. Proportion of patients who are not using protocol-specified opioids for endometriosis-associated pain at Week 24/EOT	Relugolix CT vs Placebo	86% vs 76%	82% vs 66%
	Difference	9.4%	15.9%
	95% CI	(2.0%, 16.8%)	(7.5%, 24.2%)
	p-value	0.0005	<0.0001
6. Change from baseline to Week 24/EOT in the mean dyspareunia NRS score	Relugolix CT vs Placebo	-2.4 vs -1.7	-2.4 vs -1.9
	Difference	-0.7	-0.5
	95% CI	(-1.3, -0.1)	(-1.0, 0.0)
	p-value	0.0149	0.0371
7. Proportion of patients who are not using analgesics for endometriosis-associated pain at Week 24/EOT	Relugolix CT vs Placebo	56% vs 31%	54% vs 24%
	Difference	25.5%	30.8%
	95% CI	(16.4%, 34.6%)	(21.9%, 39.8%)
	p-value	<0.0001	<0.0001

© Gedeon Richter UK Ltd (2023). All rights reserved

Primary efficacy endpoints

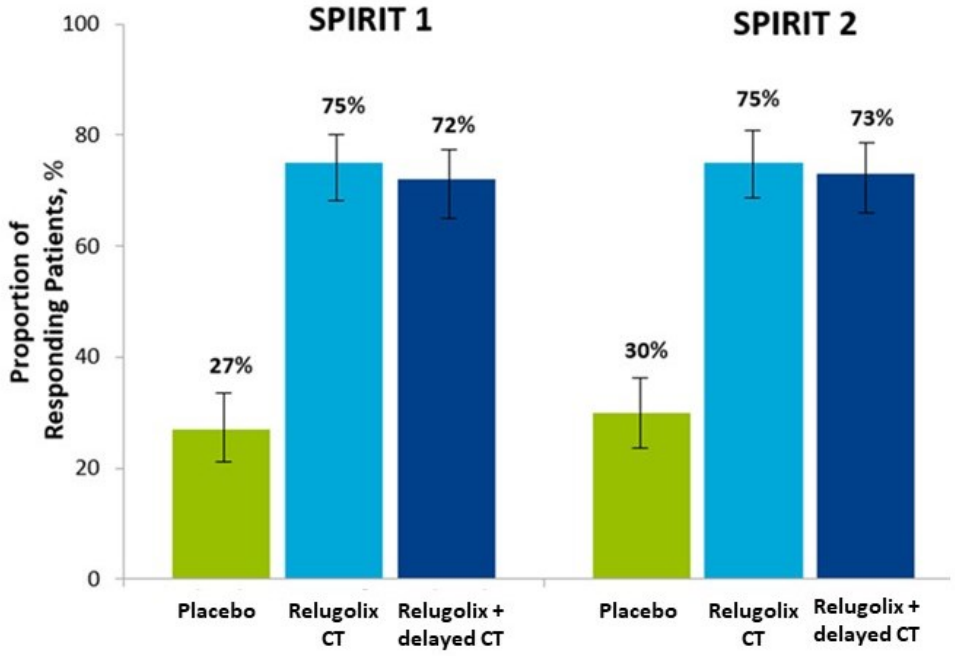
A patient was defined as a responder for the dysmenorrhoea primary endpoint if they had a mean reduction in dysmenorrhoea NRS score from baseline of at least 2.8 points without increased use of protocol-specified analgesics at Week 24. A patient was defined as a responder for the NMPP primary endpoint if they had a mean reduction in dysmenorrhoea NRS score from baseline of at least 2.1 points without increased use of protocol-specified analgesics at Week 24.

Both studies met the co-primary endpoints and first key secondary endpoint by demonstrating that Relugolix CT was statistically significantly superior to placebo.

Co-primary efficacy endpoint: Proportion of Dysmenorrhoea Responders at Week 24/EOT (mITT Population)

In both SPIRIT 1 & SPIRIT 2, 75% of patients receiving Relugolix CT achieved a decline in the dysmenorrhoea NRS score by ≥2.8 points without an increase in analgesic use, compared with 27% and 30% in the placebo groups, respectively (Figure 7). The observed difference between the two groups was 47.6% (95% CI: 39.3%, 56.0%) in SPIRIT 1 and 44.9% (95% CI: 36.2%, 53.5%) in SPIRIT 2 in favour of Relugolix CT; these differences were statistically significant (p <0.0001) (37).

Figure 7: Women achieving a mean reduction in NRS score of ≥2.8 for dysmenorrhoea and no increase in analgesic use at Week 24

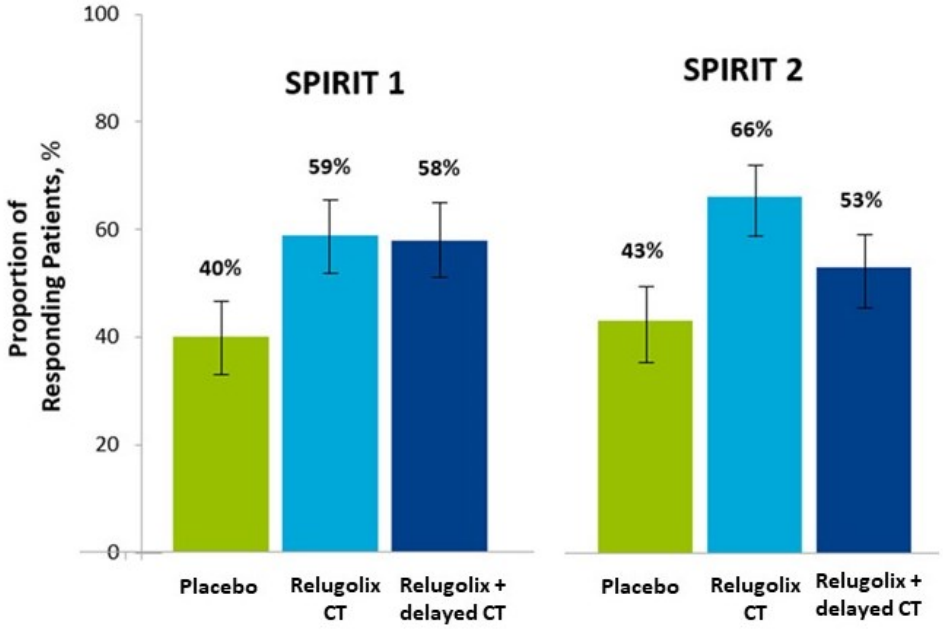


© Gedeon Richter UK Ltd (2023). All rights reserved

Co-primary efficacy endpoint: Proportion of Patients Classified as Non-Menstrual Pelvic Pain Responders at Week 24/EOT (mITT Population)

As shown in Figure 8, 59% of patients in the SPIRIT 1 Relugolix CT group were NMPP responders, compared with 40% in the placebo group. The treatment difference was 18.9% (95% CI: 9.5%, 28.2%); p <0.0001. In SPIRIT 2, 66% of patients in the Relugolix CT group were NMPP responders compared with 43% in the placebo group. The treatment difference was 23.4% (95% CI: 14.0%, 32.8%); (p <0.0001) (37).

Figure 8: Women achieving a mean reduction in NRS score of ≥2.1 points for non-menstrual pelvic pain and no Increase in analgesic use at Week 24



Secondary efficacy endpoints

Key Secondary Endpoint: Change from Baseline in the Endometriosis Health Profile-30 Pain Domain Score at Week 24, (mITT Population)

The first key secondary endpoint evaluated the functional effects of endometriosis-associated pain as assessed by the change from baseline to Week 24 in the EHP-30 Pain Domain score. Patients reported the frequency (never, rarely, sometimes, often, and always) with which they had difficulty with activities such as standing, sitting,

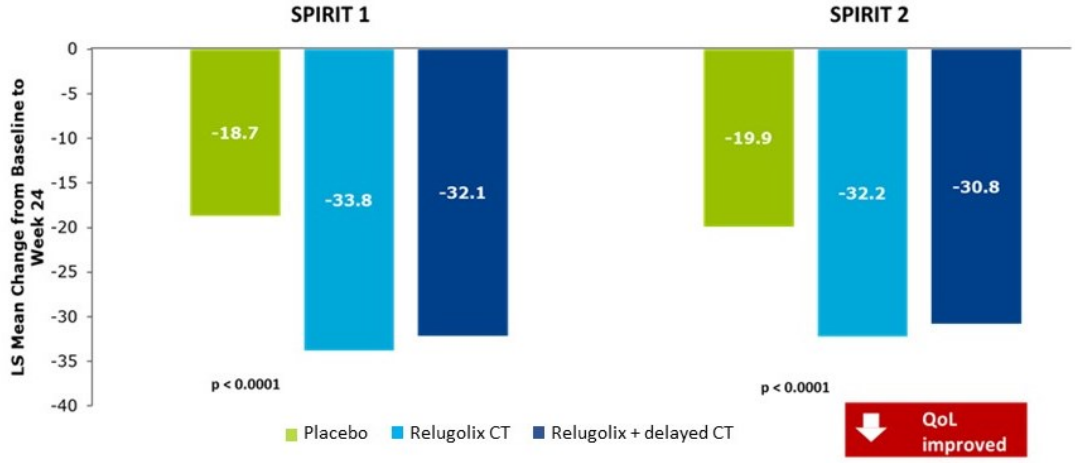
© Gedeon Richter UK Ltd (2023). All rights reserved

walking, sleeping, and performing jobs around the house because of pain. The Pain Domain scores could range from 0 to 100, with higher scores denoting greater functional impact of pain.

The baseline EHP-30 Pain Domain mean (SD) score was 58.3 (16.7) in the Relugolix CT group and 55.5 (16.0) in the placebo group in SPIRIT 1. In SPIRIT 2, the score was 56.2 (17.1) in the Relugolix CT group and 55.0 (16.2) in the placebo group. As shown in Figure 9, there was a statistically significant improvement in the EHP-30 Pain Domain score for the Relugolix CT group compared with the placebo group at Week 24 in both studies. The least squares (LS) mean (SE) change from baseline was −33.8 (1.8) versus −18.7 (1.8) (p <0.0001) in SPIRIT 1 and −32.2 (1.7) versus −19.9 (1.7) (p <0.0001) in SPIRIT 2 (37).

76.3% of patients receiving Relugolix CT in SPIRIT 1 and 72.9% of those receiving Relugolix CT in SPIRIT 2 had a meaningful improvement (i.e., reduction of at least 20 points) in the EHP-30 Pain Domain score at Week 24, compared with 48.5% and 52.5% in the respective placebo groups. The observed difference between the two groups was 27.8% (95% CI: 17.90%, 37.73%) in SPIRIT 1 and 20.5% (95% CI: 10.29%, 30.66%) in SPIRIT 2 in favour of the Relugolix CT group; these differences were statistically significant (p <0.0001) in SPIRIT 1 and p=0.0002 in SPIRIT 2.

Figure 9: LS mean change in EHP 30 pain domain from baseline to Week 24



The EHP-30 Pain Domain analyses support and extend the findings of the co-primary endpoints by demonstrating that reducing endometriosis-associated pain with Relugolix CT also meaningfully reduced the impact of pain on function.

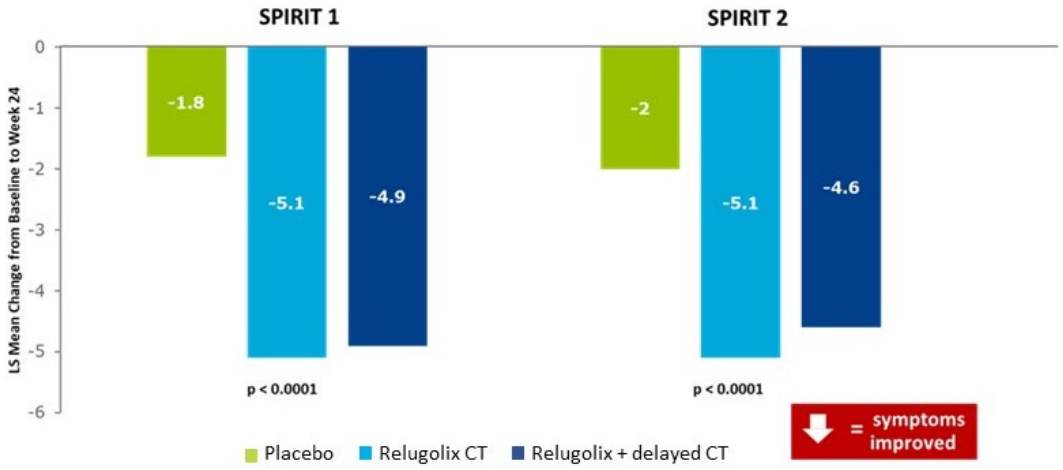
© Gedeon Richter UK Ltd (2023). All rights reserved

Morbidity from chronic pain is caused by both the aversive nature of pain as well as through its effect on limiting activities that are painful. The EHP-30 Pain Domain score, a measure of the frequency with which women reported difficulty with activities such as standing, sitting, walking, sleeping, and performing jobs around the house because of pain, was statistically and clinically significantly improved with Relugolix CT treatment versus placebo.

Secondary efficacy endpoint: Change from baseline to Week 24/EOT in the mean dysmenorrhoea NRS score

The key secondary endpoint evaluating dysmenorrhoea was the change from baseline to Week 24/EOT in the dysmenorrhoea NRS score. Patients were to report their pelvic pain on an 11-point NRS (0 = no pain to 10 = pain as bad as you can imagine) daily in an eDiary. In the Relugolix CT group, the LS change from baseline to Week 24 in the dysmenorrhoea NRS score was greater than that in the placebo group (-5.1 versus -1.8 in SPIRIT 1 and -5.1 versus -2.0 in SPIRIT 2) as shown in Figure 10, and the differences between the two groups were statistically significant (p <0.0001) (37).

Figure 10: LS mean change in dysmenorrhoea NRS score from baseline to Week 24



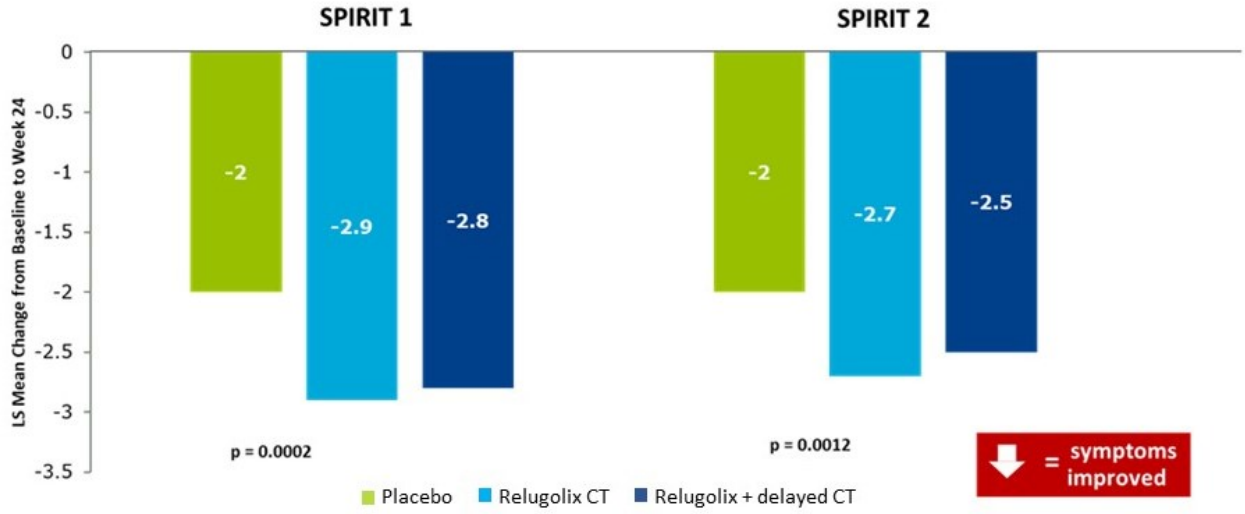
© Gedeon Richter UK Ltd (2023). All rights reserved

Secondary efficacy endpoint: Change from baseline to Week 24/EOT in the mean NMPP NRS score

The key secondary endpoint evaluating NMPP was the change from baseline to Week 24/EOT in the NMPP NRS score. Patients reported their pelvic pain on an 11-point NRS (0 = no pain to 10 = pain as bad as you can imagine) daily in an eDiary.

In the Relugolix CT group, the LS mean change from baseline to Week 24 in the NMPP NRS score was greater than that in the placebo group (-2.9 versus -2.0 in SPIRIT 1 and -2.7 versus -2.0 in SPIRIT 2), as shown in Figure 11 (p <0.0002 in SPIRIT 1; p = 0.0012 in SPIRIT 2) (37).

Figure 11: LS mean change in NMPP NRS score from baseline to Week 24



Secondary efficacy endpoint: Change from baseline to Week 24/EOT in the mean overall pelvic pain NRS score

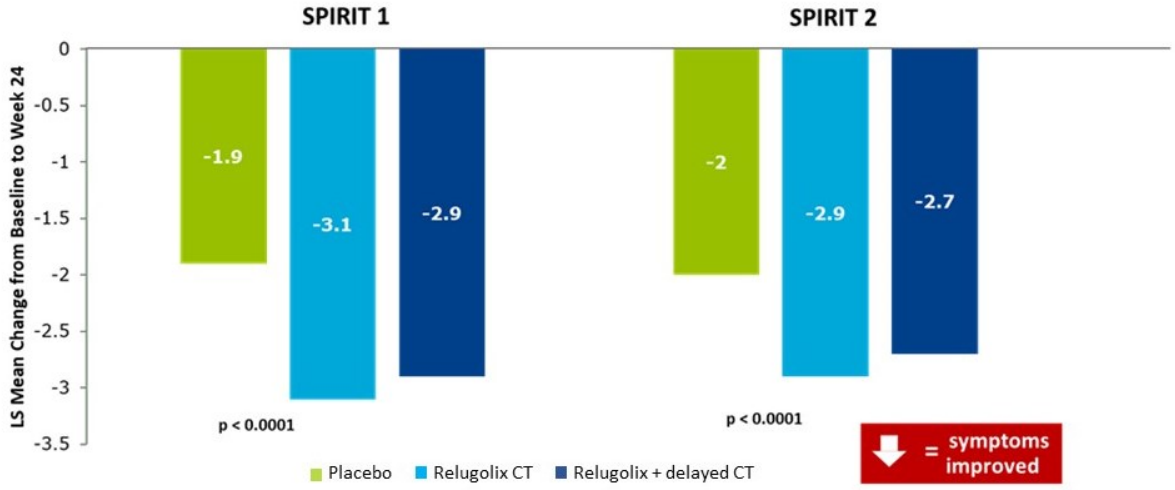
The key secondary endpoint evaluating overall pelvic pain was the change from baseline to Week 24/EOT in the pelvic pain NRS scores irrespective of menstruation status. Patients reported their pelvic pain on an 11-point NRS (0 = no pain to 10 = pain as bad as you can imagine) daily in an eDiary.

There was a significantly greater improvement in the overall pelvic pain NRS score in the Relugolix CT groups compared with the placebo groups at Week 24/EOT.

© Gedeon Richter UK Ltd (2023). All rights reserved

In the Relugolix CT group, the LS change from baseline to Week 24 in the overall pelvic pain NRS score was greater than that in the placebo group (-3.1 versus -1.9 in SPIRIT 1 and -2.9 versus -2.0 in SPIRIT 2) as shown in Figure 12; p <0.0001 (37).

Figure 12: LS mean change in overall pelvic pain NRS score from baseline to Week 24



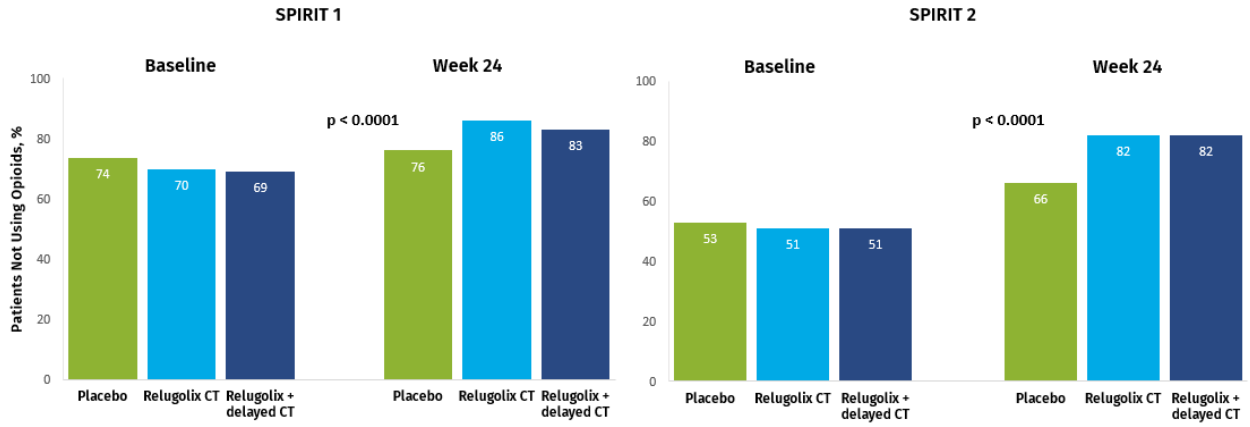
Secondary efficacy endpoint: Proportion of patients who are not using protocol-specified opioids for endometriosis-associated pain at Week 24/EOT

The key secondary endpoint evaluating opioid use was based on the proportion of patients who were not using protocol-specified opioids for endometriosis-associated pain at Week 24/EOT.

Figure 13 shows the proportion of patients not using opioids from baseline through Week 24.

© Gedeon Richter UK Ltd (2023). All rights reserved

Figure 13: Proportion of patients not using opioids increased significantly from baseline to Week 24



In SPIRIT 1, 182 patients (86%) in the Relugolix CT group, and 162 (76%) in the placebo group were not using protocol-specified opioids at Week 24/EOT. Similarly, in SPIRIT 2, 169 patients (82%) in the Relugolix CT group, and 135 (66%) in the placebo group were not using protocol-specified opioids at Week 24/EOT. The between-group difference was 9.4% (95% CI: 2.0%, 16.8%) in SPIRIT 1 and 15.9% (95% CI: 7.5%, 24.2%) in SPIRIT 2 in favour of the Relugolix CT group; these differences were statistically significant (SPIRIT 1: p = 0.0005, SPIRIT 2: p <0.0001) (37).

Secondary efficacy endpoint: Change from baseline to Week 24/EOT in the mean dyspareunia NRS score

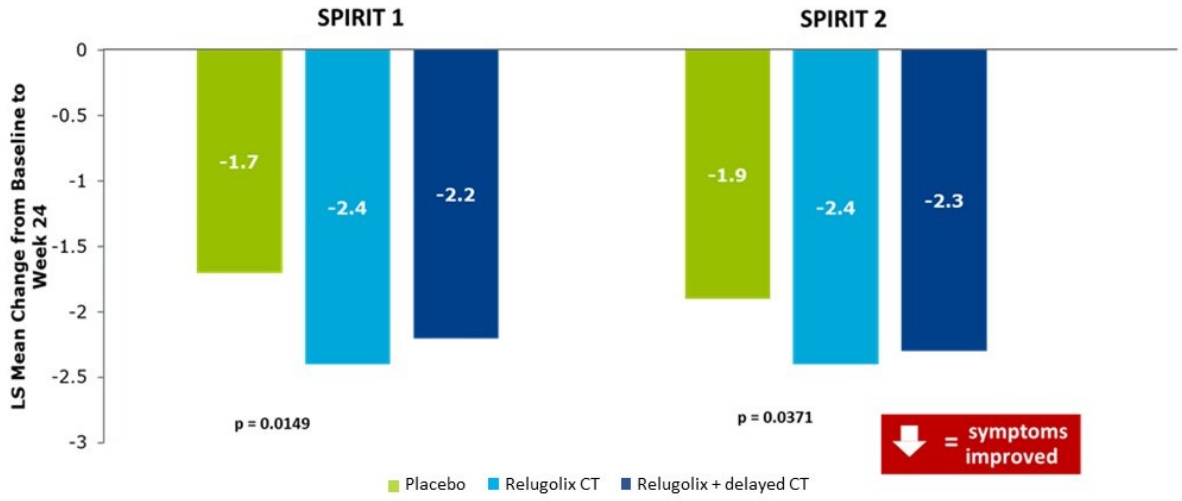
The key secondary endpoint evaluating dyspareunia was the change from baseline to Week 24/EOT in the dyspareunia NRS scores among patients reporting at least one vaginal sexual intercourse with a non-zero pain score at baseline. Patients were to report whether they had vaginal sexual intercourse and, if so, their level of pelvic pain during vaginal sexual intercourse on an 11-point NRS (0 = no pain to 10 = pain as bad as you can imagine) daily in an eDiary.

The change in mean dyspareunia NRS score from baseline to Week 24/EOT is shown in Figure 14. There was a significantly greater improvement in the dyspareunia NRS scores in the Relugolix CT groups compared with the placebo

© Gedeon Richter UK Ltd (2023). All rights reserved

groups (-2.4 versus -1.7 in SPIRIT 1, p =0.0149, and -2.4 versus -1.9 in SPIRIT 2, p=0.0371) (37).

Figure 14: LS mean change in dyspareunia NRS score from baseline to Week 24



The dyspareunia NRS score at Week 24/EOT represented a 40.1% improvement for patients in the Relugolix CT group compared with a 23.1% improvement for patients in the placebo group in SPIRIT 1. In SPIRIT 2, the dyspareunia NRS score at Week 24/EOT represented a 46.0% improvement for patients in the Relugolix CT group compared with a 34.1% improvement for patients in the placebo group in SPIRIT 2.

Secondary efficacy endpoint: Change from baseline to Week 24/EOT in protocol-specified analgesic use for endometriosis-associated pain based on mean pill count

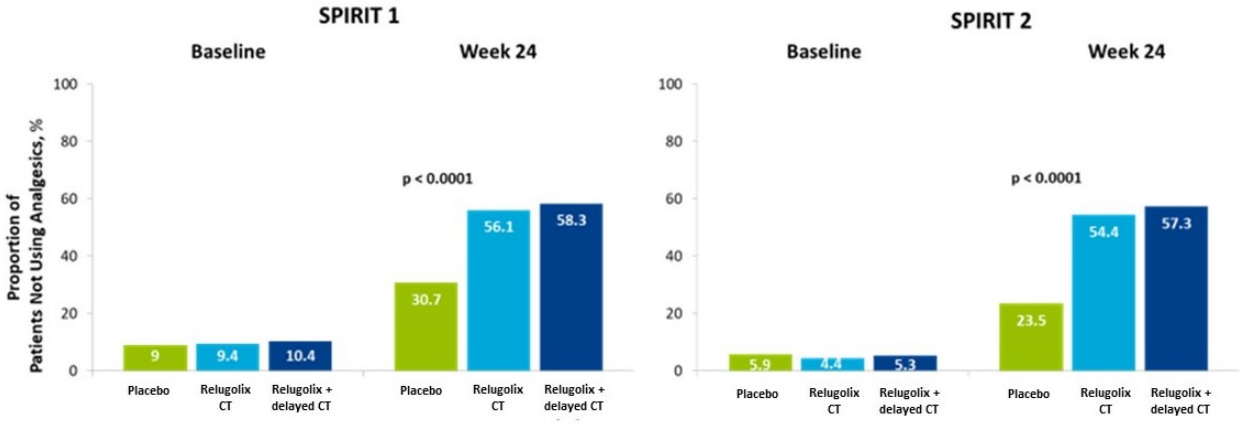
The secondary endpoint evaluating overall analgesic use was the proportion of patients who were not using protocol-specified analgesics for endometriosis-associated pain at Week 24/EOT (prespecified key secondary endpoint in SPIRIT 1, post hoc analysis in SPIRIT 2).

The proportion of patients who were not using analgesics increased significantly between baseline and Week 24/EOT (Figure 15). In SPIRIT 1, 119 patients (56.%) patients in the Relugolix CT group and 65 (31%) in the placebo group, were not using protocol-specified analgesics at Week 24/EOT. In SPIRIT 2, the corresponding numbers were 112 (54.%) patients in the Relugolix CT group and 48 (24%) in the placebo group. The between-group difference was 25.5% (95% CI: 16.4%, 34.6%) in

© Gedeon Richter UK Ltd (2023). All rights reserved

SPIRIT 1 and 30.8% (95% CI: 21.9%, 39.8%) in SPIRIT 2 in favour of the Relugolix CT group (both p <0.0001) (37).

Figure 15: No. of patients not using analgesics increased significantly from baseline through to Week 24



EQ-5D-5L

EQ-5D data from SPIRIT 1 and 2 were used to generate utility values for the economic analysis. Table 18 shows the improvements from baseline in EQ-5D-5L to Week 24 (62, 63).

Table 18: EQ-5D-5L change from baseline to Week 24

	SPIRIT 1		SPIRIT 2	
Domain Baseline (n)	**Relugolix CT (n = 208)**	**Placebo (n = 207)**	**Relugolix CT (n = 203)**	**Placebo (n = 203)**
Mobility				
No problems walking	70 (33.7%)	84 (40.6%)	94 (46.3%)	94 (46.3%)
Slight problems walking	63 (30.3%)	70 (33.8%)	61 (30.0%)	62 (30.5%)
Moderate problems walking	61 (29.3%)	44 (21.3%)	42 (20.7%)	43 (21.2%)
Severe problems walking	13 (6.3%)	9 (4.3%)	5 (2.5%)	4 (2.0%)
Unable to walk	1 (0.5%)	0	1 (0.5%)	0
Change at Week 24 (n)	173	165	173	162
No change	58 (33.5%)	85 (51.5%)	86 (49.7%)	88 (54.3%)
1 to 2 category improvement	97 (56.1%)	71 (43.0%)	78 (45.1%)	58 (35.8%)
3 to 4 category improvement	6 (3.5%)	3 (1.8%)	3 (1.7%)	2 (1.2%)
Self-care				
No problems washing or dressing myself	133 (63.9%)	155 (74.9%)	143 (70.4%)	150 (73.9%)
Slight problems washing or dressing myself	41 (19.7%)	33 (15.9%)	41 (20.2%)	38 (18.7%)

© Gedeon Richter UK Ltd (2023). All rights reserved

Moderate problems washing or dressing myself	32 (15.4%)	19 (9.2%)	16 (7.9%)	13 (6.4%)
Severe problems washing or dressing myself	2 (1.0%)	0	3 (1.5%)	2 (1.0%)
Unable to wash or dress myself	0	0	0	0
Change at Week 24 (n)	173	165	173	162
No change	106 (61.3%)	120 (72.7%)	122 (70.5%)	125 (77.2%)
1 to 2 category improvement	63 (36.4%)	37 (22.4%)	47 (27.2%)	31 (19.1%)
3 to 4 category improvement	1 (0.6%)	0	1 (0.6%)	2 (1.2%)
Usual activities				
No problems doing my usual activities	42 (20.2%)	48 (23.2%)	49 (24.1%)	71 (35.0%)
Slight problems doing my usual activities	69 (33.2%)	75 (36.2%)	75 (36.9%)	67 (33.0%)
Moderate problems doing my usual activities	71 (34.1%)	68 (32.9%)	63 (31.0%)	47 (23.2%)
Severe problems doing my usual activities	26 (12.5%)	15 (7.2%)	12 (5.9%)	17 (8.4%)
Unable to do my usual activities	0	1 (0.5%)	4 (2.0%)	1 (0.5%)
Change at Week 24 (n)	173	165	173	162
No change	53 (30.6%)	54 (32.7%)	58 (33.5%)	66 (40.7%)
1 to 2 category improvement	104 (60.1%)	90 (54.5%)	91 (52.6%)	74 (45.7%)
3 to 4 category improvement	10 (5.8%)	5 (3.0%)	11 (6.4%)	5 (3.1%)
Pain/discomfort				
No pain or discomfort	6 (2.9%)	9 (4.3%)	11 (5.4%)	18 (8.9%)
Slight pain or discomfort	50 (24.0%)	52 (25.1%)	49 (24.1%)	58 (28.6%)
Moderate pain or discomfort	99 (47.6%)	107 (51.7%)	102 (50.2%)	94 (46.3%)
Severe pain or discomfort	47 (22.6%)	37 (17.9%)	34 (16.7%)	30 (14.8%)
Extreme pain or discomfort	6 (2.9%)	2 (1.0%)	7 (3.4%)	3 (1.5%)
Change at Week 24 (n)	173	165	173	162
No change	33 (19.1%)	45 (27.3%)	27 (15.6%)	50 (30.9%)
1 to 2 category improvement	117 (67.6%)	87 (52.7%)	121 (69.9%)	87 (53.7%)
3 to 4 category improvement	16 (9.2%)	9 (5.5%)	13 (7.5%)	4 (2.5%)
Anxiety/depression				
Not anxious or depressed	55 (26.4%)	73 (35.3%)	55 (27.1%)	70 (34.5%)
Slightly anxious or depressed	64 (30.8%)	59 (28.5%)	65 (32.0%)	60 (29.6%)
Moderately anxious or depressed	61 (29.3%)	58 (28.0%)	63 (31.0%)	46 (22.7%)
Severely anxious or depressed	25 (12.0%)	13 (6.3%)	17 (8.4%)	24 (11.8%)
Extremely anxious or depressed	3 (1.4%)	4 (1.9%)	3 (1.5%)	3 (1.5%)
Change at Week 24 (n)	173	165	173	162
No change	63 (36.4%)	64 (38.8%)	64 (37.0%)	66 (40.7%)
1 to 2 category improvement	85 (49.1%)	59 (35.8%)	80 (46.2%)	66 (40.7%)
3 to 4 category improvement	8 (4.6%)	6 (3.6%)	6 (3.5%)	3 (1.9%)

© Gedeon Richter UK Ltd (2023). All rights reserved

In SPIRIT 1, the mean (SD) overall health status on the VAS at baseline was 55.3 (18.74) in the Relugolix CT group and 55.6 (18.57) in the placebo group. The mean (SD) improvement in overall health status at Week 24 was greater in the Relugolix CT group than in the placebo group: 22.8 (21.31) versus 14.0 (23.52). A similar pattern was seen in SPIRIT 2: the mean (SD) overall health status on the VAS at baseline was 57.0 (20.07) in the Relugolix CT group and 60.7 (21.50) in the placebo group. The mean (SD) improvement in overall health status at Week 24 20.2 (23.68) in the Relugolix CT group and 12.7 (24.75) in the placebo group.

In both studies, most patients reported baseline pain/discomfort, anxiety/depression, problems with walking, and problems doing usual activities. The percentage of patients reporting improvements with walking, self-care, usual activities, pain or discomfort, and anxiety or depression were all numerically higher in the Relugolix CT group compared with the placebo group at Week 24. These findings are consistent with the observed improvements in the co-primary pain endpoints.

SPIRIT OLE

Primary efficacy endpoints

The primary efficacy endpoints for this OLE study were defined in a manner analogous to the co-primary endpoints for the pivotal studies SPIRIT 1 and SPIRIT 2. Both endpoints were assessed at Week 52 and at Week 104/EOT.

Co-primary efficacy endpoint: Proportion of women who respond or maintain response based on assessment of dysmenorrhoea at Week 52 and Week 104 (Extension Study Population)

A patient was defined as a responder for the dysmenorrhoea primary endpoints if the NRS score for dysmenorrhoea declined from baseline to the endpoint timepoint (Week 52 or Week 104/EOT) by at least 2.8 points without increased use of protocol-specified analgesics for pelvic pain at the endpoint timepoint (Week 52 or Week 104/EOT) relative to baseline.

The analyses of the dysmenorrhoea primary efficacy endpoints, the proportion of patients meeting the dysmenorrhoea responder definition at Week 52 and Week

© Gedeon Richter UK Ltd (2023). All rights reserved

104/ET and the components of the endpoints are presented in Table 19 below (64, 65).

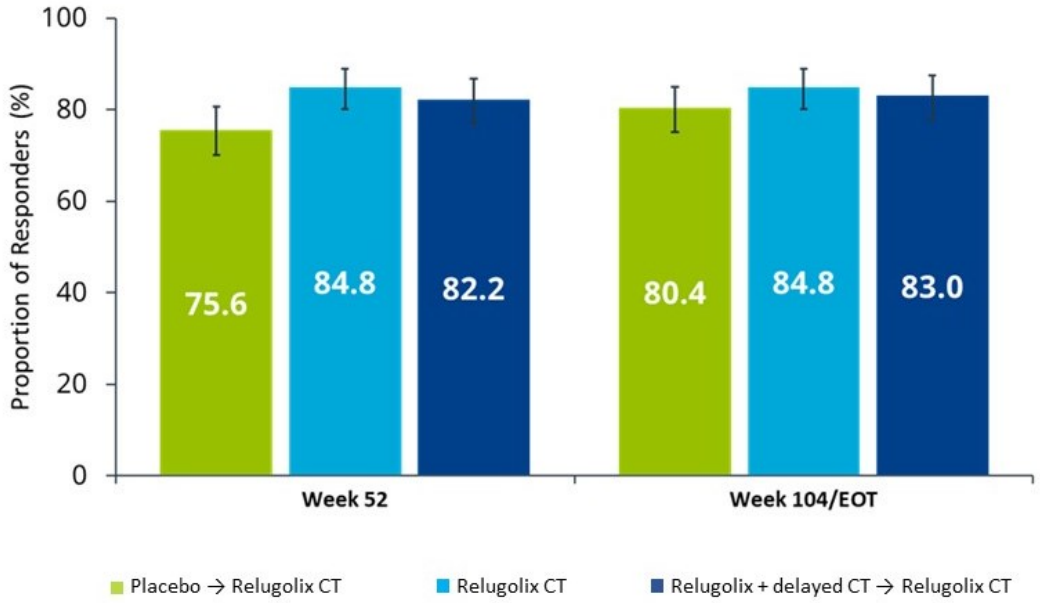
Table 19: SPIRIT OLE Primary efficacy analysis: Proportion of Patients Classified as Dysmenorrhoea Responders at Week 52 and Week 104/EOT (Extension Study Population)

	Relugolix CT (N=277)	**Placebo → Relugolix CT (N=275)**
Number (%) of responders at Week 52 (95% CI)	235 (84.8%) (80.06%, 88.85%)	208 (75.6%) (70.12%, 80.59%)
Number (%) of patients with a reduction of at least 2.8 points from baseline in mean dysmenorrhoea NRS score at Week 52 (95% CI)	[redacted]	[redacted]
Number (%) of patients with no increase in analgesic use from baseline at Week 52 (95% CI)	[redacted]	[redacted]
Number (%) of responders at Week 104/EOT (95% CI)	235 (84.8%) (80.06%, 88.85%)	221 (80.4%) (75.17%, 84.89%)
Number (%) of patients with a reduction of at least 2.8 points from baseline in mean dysmenorrhoea NRS score at Week 104/EOT (95% CI)	[redacted]	[redacted]
Number (%) of patients with no increase in analgesic use from baseline at Week 104/EOT (95% CI)	[redacted]	[redacted]

Figure 16 shows the proportion of patients who met the dysmenorrhoea responder definition over time (64, 65).

© Gedeon Richter UK Ltd (2023). All rights reserved

Figure 16: SPIRIT OLE: women achieving a mean reduction in NRS Score of ≥2.8 for dysmenorrhoea and no increase in analgesic use



In the Relugolix CT group, 235 patients (84.8%) met the dysmenorrhoea responder definition at Week 52. At Week 104/EOT, the responder rate remained unchanged. As shown in Table 19, the proportion of patients meeting the criteria for the individual components of this composite primary endpoint was high, indicating that neither component (i.e., reduction in NRS by ≥2.8 or lack of increase in analgesics from baseline) drove the results for the primary endpoint.

In the placebo group (the group that had received active treatment with Relugolix CT for up to 80 weeks in the extension study), 208 patients (75.6%) met the dysmenorrhoea responder definition at Week 52. At Week 104/EOT, the responder rates increased slightly: 221 patients (80.4%) met the dysmenorrhoea responder definition. A high percentage of patients met both components of the endpoint at both timepoints.

© Gedeon Richter UK Ltd (2023). All rights reserved

Co-primary efficacy endpoint: Proportion of women who respond or maintain response based on assessment of NMPP at Week 52 and Week 104 (Extension Study Population)

A patient was defined as a responder for the NMPP primary endpoints if the NRS score for NMPP declined from baseline to the endpoint timepoint (Week 52 or Week 104/EOT) by at least 2.1 points without increased use of protocol-specified analgesics for pelvic pain at the endpoint timepoint (Week 52 or Week 104/EOT) relative to baseline.

The analyses of the primary efficacy endpoints, proportion of patients meeting the NMPP responder definition at Week 52 and Week 104/ET, are presented in Table 20 (64, 65).

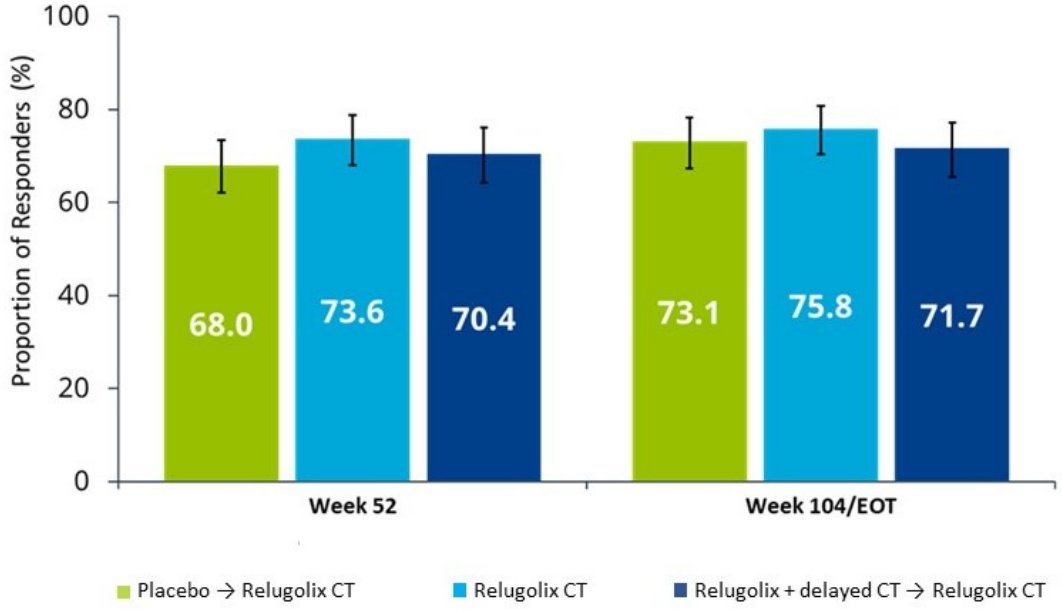
Table 20: Primary efficacy analysis: proportion of patients classified as NMPP responders at Week 52 and Week 104/EOT (extension study population)

	Relugolix CT	**Placebo → Relugolix CT**
Number (%) of responders at Week 52 (95% CI)	204 (73.6%) (68.04%, 78.74%)	187 (68.0%) (62.13%, 73.47%)
Number (%) of patients with a reduction of at least 2.1 points from baseline in mean NMPP NRS score at Week 52 (95% CI)	[redacted]	[redacted]
Number (%) of patients with no increase in analgesic use from baseline at Week 52 (95% CI)	[redacted]	[redacted]
Number (%) of responders at Week 104/EOT (95% CI)	210 (75.8%) (70.33%, 80.74%)	201 (73.1%) (67.44%, 78.24%)
Number (%) of patients with a reduction of at least 2.1 points from baseline in mean NMPP NRS score at Week 104/EOT (95% CI)	[redacted]	[redacted]
Number (%) of patients with no increase in analgesic use from baseline at Week 104/EOT (95% CI)	[redacted]	[redacted]

The proportions of patients who met the NMPP responder definition over time are presented in Figure 17 (64, 65).

© Gedeon Richter UK Ltd (2023). All rights reserved

Figure 17: SPIRIT OLE: women achieving a mean reduction in NRS Score of ≥2.1 points for non-menstrual pelvic pain and no increase in analgesic use



In the Relugolix CT group, 204 patients (73.6%) met the NMPP responder definition at Week 52. At Week 104/EOT, the responder rate remained essentially unchanged: 210 patients (75.8%) met the NMPP responder definition. As shown in Table 20, the proportion of patients meeting the criteria for the individual components of this composite primary endpoint were high, indicating that neither component (i.e., reduction in NRS by ≥2.1 or lack of increase in analgesics from baseline) drove the results for the primary endpoint.

In the placebo group (those that had received active treatment with Relugolix CT for up to 80-weeks in the extension study), 187 patients (68.0%) met the NMPP responder definition at Week 52. At Week 104, the responder rates remained similar: 201 patients (73.1%) met the NMPP responder definition at Week 104/EOT. A high percentage of patients met both endpoints, the decline of the NRS score for NMPP and the non-increase of analgesics, at both timepoints (Week 52 and Week 104).

Secondary efficacy endpoints: Overview

Table 21 summarizes the results of all secondary efficacy endpoints in SPIRIT OLE (64, 65). Results for each endpoint are further described in following sections.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 21: SPIRIT OLE secondary efficacy endpoints

		Relugolix CT	Placebo → Relugolix CT
Week 52 Key secondary endpoints			
Change from baseline in the EHP-30 pain domain score at Week 52	LS mean	-37.7	-35.1
	SE	1.34	1.32
Change from baseline in the mean dysmenorrhoea NRS score at Week 52	LS mean	-5.9	-5.3
	SE	0.15	0.15
Change from baseline in the mean NMPP score at Week 52	LS mean	-3.6	-3.4
	SE	0.15	0.15
Change from baseline in the mean overall pelvic pain NRS score at Week 52	LS mean	-3.9	-3.6
	SE	0.15	0.15
Change from baseline in the mean dyspareunia NRS score at Week 52	LS mean	-3.3	-3.0
	SE	0.18	0.18
Proportion of patients not using protocol-specified opioids for endometriosis-associated pain at Week 52	n	201	208
	%	86.3%	88.9%
Proportion of patients not using analgesics for endometriosis-associated pain at Week 52	n	151	165
	%	64.8%	70.5%
Week 104 Key secondary endpoints			
Change from baseline in the EHP-30 pain domain score at Week 104	LS mean	-41.3	-37.7
	SE	1.33	1.29
Change from baseline in the mean Dysmenorrhoea NRS score at Week 104	LS mean	-5.9	-5.6
	SE	0.17	0.17
Change from baseline in the mean NMPP score at Week 104	LS mean	-4.0	-3.8
	SE	0.16	0.16
Change from baseline in the mean overall pelvic pain NRS score at Week 104	LS mean	-4.2	-4.0
	SE	0.16	0.16
Change from baseline in the mean dyspareunia NRS score at Week 104	LS mean	-3.5	-3.4
	SE	0.21	0.21
Proportion of patients not using protocol-specified opioids for endometriosis-associated pain at Week 104	n	252	249
	%	91.0%	90.5%
Proportion of patients not using analgesics for endometriosis-associated pain at Week 104	n	208	209
	%	75.1%	76.0%

© Gedeon Richter UK Ltd (2023). All rights reserved

Secondary efficacy endpoint: Change from baseline in the EHP-30 pain domain score at Week 52 and Week 104 (Extension Study Population)

The EHP-30 Pain Domain evaluates the functional effects of endometriosis-associated pain. Patients reported the frequency (never, rarely, sometimes, often, and always) with which they had difficulty with activities such as standing, sitting, walking, sleeping, and performing jobs around the house because of pain. The Pain Domain normalised scores could range from 0 to 100, with higher scores denoting greater functional impact of pain.

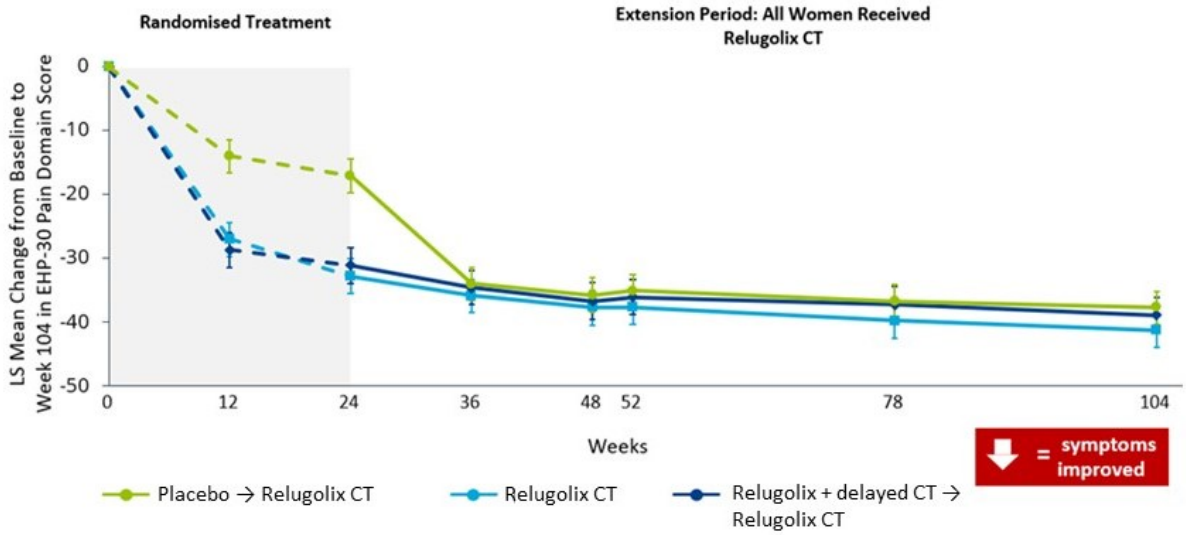
Mean changes from baseline to Week 52 and to Week 104 on the EHP-30 Pain Domain scores are provided in Figure 18.

The baseline EHP-30 pain domain score in the Relugolix CT and placebo groups was similar (57.3 and 56.2, respectively) (64).

At Week 52, the LS mean (standard error, [SE]) change from baseline in the Relugolix CT group was -37.7 (1.34) (95% CI: -40.3, -35.0), representing a 66.4% decrease (improvement) from baseline. At Week 104, the EHP-30 change from baseline remained consistent (LS mean -41.3 [1.33] [95% CI: -43.9, -38.7]), representing a 72.2% decrease (improvement) from baseline (64, 65). A meaningful response in the EHP-30 pain domain was pre-specified as a ≥20-point improvement.

© Gedeon Richter UK Ltd (2023). All rights reserved

Figure 18: LS mean change in EHP 30 Pain domain from baseline to week 104



In the Relugolix CT group, the responder rate for function, as assessed by the EHP-30 pain domain, increased from [redacted] at Week 12 to 83.6% (95% CI: 78.22, 88.14) at Week 52 and 88.6% (95% CI: 82.80, 93.01) at Week 104 (64).

In the placebo group, the percentage with a meaningful functional response was lower than in the relugolix groups through Week 24, but increased following initiation of treatment with Relugolix CT at Week 24 and was similar to that in the relugolix groups between Week 36 and Week 104

In summary, treatment with Relugolix CT was associated with improved functioning on all daily activities assessed and a high proportion of patients achieved and maintained clinically meaningful functional improvements (83.6% at Week 52 and 88.6% at Week 104). The time course of improvement in function, is consistent with the time course of improvement in dysmenorrhoea and NMPP, consistent with the hypothesis that reducing pain in women with endometriosis would improve their functioning.

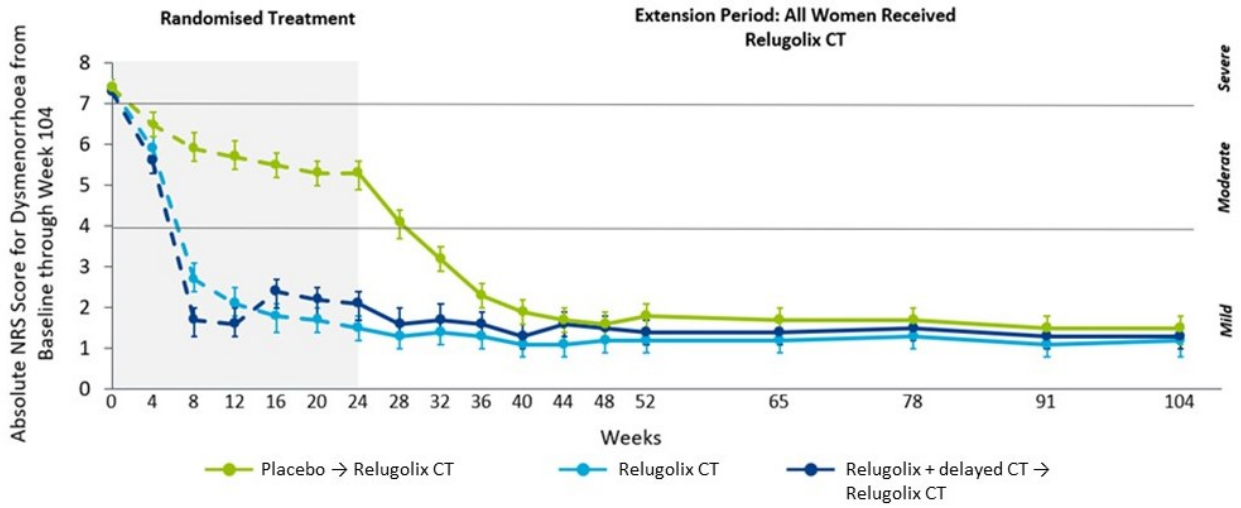
Secondary efficacy endpoint: Change from baseline in the mean dysmenorrhoea NRS score at Week 52 and Week 104 (Extension Study Population)

The secondary endpoints evaluating dysmenorrhoea based on the NRS were the change and percent change from baseline to Week 52 and from baseline to Week 104. Patients were to report their pelvic pain on an 11-point NRS (0 = no pain to 10 = pain as bad as you can imagine) daily in an eDiary.

© Gedeon Richter UK Ltd (2023). All rights reserved

The change from baseline in average dysmenorrhoea NRS score by visit is shown in Figure 19 (64).

Figure 19: LS Mean Change in Dysmenorrhoea NRS Score from Baseline to Week 104



The baseline LS mean (SE) dysmenorrhoea NRS scores were 7.4 (0.11) for the Relugolix CT group and 7.4 (0.11) for the placebo group. In the Relugolix CT group, the LS mean dysmenorrhoea score at Week 52 decreased to 1.2 (0.15), a change of -5.9 (95% CI: -6.2, -5.6), representing an 83.9% decrease from baseline and a reduction in pain scores from severe to mild. At Week 104, the LS mean dysmenorrhoea score was sustained (1.2 [0.17]), representing an 84.0% decrease from baseline.

Separation in the mean change dysmenorrhoea NRS scores between the Relugolix CT and placebo groups was evident (visually) starting with the first post-baseline time point (Week 4), with a sharp decline in between Weeks 4 and 8 and a near maximum reduction in the Relugolix CT by Week 16 that was sustained through Week 104.

In the placebo group, the LS mean dysmenorrhoea NRS score remained higher than in the relugolix groups through Week 24. Following initiation of treatment with Relugolix CT, there was a sharp decline in the score between Weeks 24 and 36. With continued treatment, the LS mean dysmenorrhoea score in this group became similar to the scores in the Relugolix CT group.

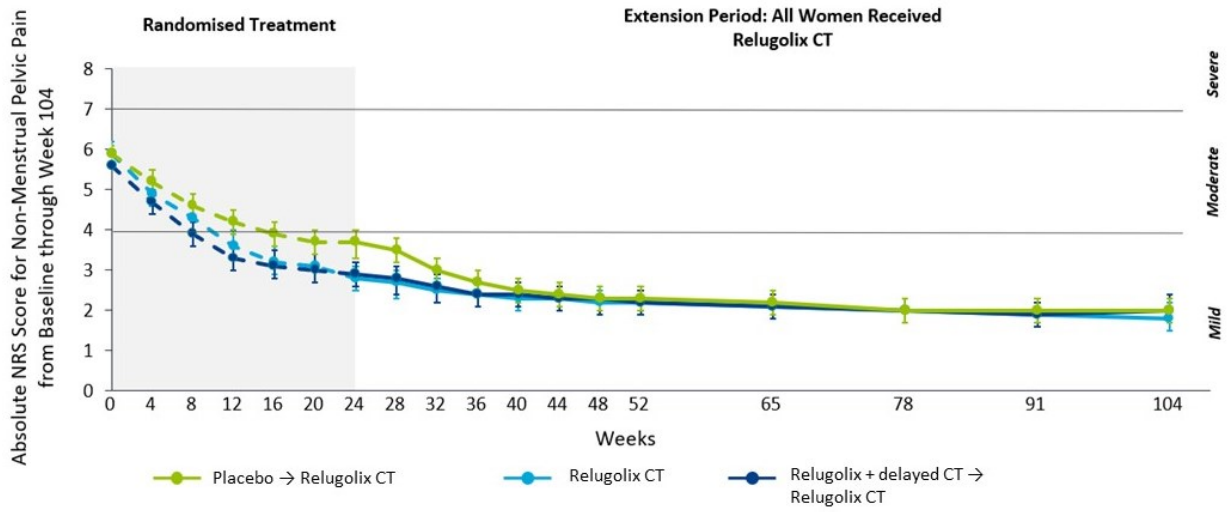
© Gedeon Richter UK Ltd (2023). All rights reserved

Dysmenorrhoea improved as early as Week 8 (two menstrual cycles) in the treatment course with Relugolix CT, reached near maximal improvement at approximately Week 28, and was sustained through Week 104. At Week 52 and Week 104, there was an 83.9% and 84.0% reduction from baseline in dysmenorrhoea, respectively, with an absolute score that was indicative of minimal pain.

Secondary efficacy endpoint: Change from baseline in the mean NMPP score at Week 52 and Week 104 (Extension Study Population)

The secondary endpoints evaluating NMPP based on the NRS were the change and percent change from baseline to Week 52 and from baseline to Week 104. Patients were to report their pelvic pain on an 11-point NRS (0 = no pain to 10 = pain as bad as you can imagine) daily in an eDiary. The change from baseline in average NMPP NRS score by visit is shown in Figure 20 (64).

Figure 20: LS Mean Change in NMPP Score from Baseline to Week 104



The baseline LS mean (SE) NMPP NRS scores were 5.9 (0.13) for the Relugolix CT group and 5.9 (0.13) for the placebo group.

In the Relugolix CT group, the LS mean NMPP NRS score decreased at Week 52 to 2.2 (0.15), a change of -3.6 (95% CI: -3.9, -3.3), representing a 63.5% decrease from baseline and a reduction in pain from moderate to mild. At Week 104, the LS mean

© Gedeon Richter UK Ltd (2023). All rights reserved

NMPP NRS score was sustained (1.8 [0.17]), representing a 68.9% decrease from baseline (64, 65).

The scores in the Relugolix CT group steadily declined over time, through Week 52 and were then sustained through Week 104. Separation in the curves for the change from baseline NRS scores between the Relugolix CT and placebo groups were evident (visually) starting at Week 12.

In the placebo group, the LS mean NMPP NRS score remained higher than in the relugolix groups through Week 24. Following initiation of treatment with Relugolix CT, starting at Week 28, the LS mean NMPP score in this group decreased noticeably and subsequently became similar to the scores in the Relugolix CT group.

NMPP improved early in the treatment course (three menstrual cycles) with Relugolix CT and the improvement in pain continued to steadily decline through Week 52; this decline was subsequently sustained through Week 104.

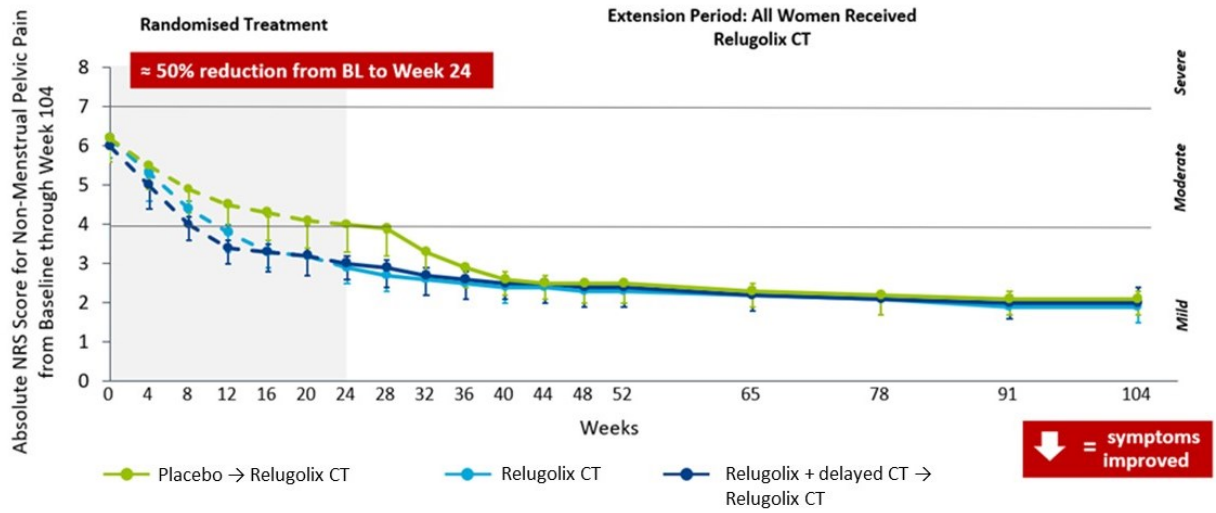
Secondary efficacy endpoint: Change from baseline in the mean overall pelvic pain NRS score at Week 52 and Week 104 (Extension Study Population)

The secondary endpoints evaluating overall pelvic pain based on the pelvic pain NRS irrespective of menstruation status were the change and percent change from baseline to Week 52 and from baseline to Week 104. Patients were to report their pelvic pain on an 11-point NRS (0 = no pain to 10 = pain as bad as you can imagine) daily in an eDiary.

The change from baseline in average overall pelvic pain NRS score at Week 104 is shown in Figure 21 (64).

© Gedeon Richter UK Ltd (2023). All rights reserved

Figure 21: LS Mean Change in overall pelvic pain NRS Score from Baseline to Week 104



The baseline LS mean (SE) overall pelvic pain NRS scores were 6.0 (0.12) for all treatment groups.

In the Relugolix CT group, the LS mean NRS score decreased at Week 52 to 2.3 (0.15), a change of -3.9 (95% CI: -4.1, -3.6), representing a 64.5% decrease from baseline and a reduction in pain from moderate to mild. At Week 104, the LS mean overall pelvic pain score was sustained (1.9 [0.16]), representing a 69.4% decrease from baseline (64, 65).

The scores in the Relugolix CT group steadily declined over time, through Week 52 and were then sustained through Week 104. Separation in the curves for the change from baseline NRS scores between the Relugolix CT and placebo groups were evident (visually) starting at Week 8.

Overall pelvic pain improved early in the treatment course with Relugolix CT, and the pain continued to steadily decline through Week 52; this decline was subsequently sustained through Week 104. The reduction in pain was substantial. At Week 52 and Week 104, there was a 64.5% and 69.4%, reduction from baseline in overall pelvic pain, respectively, with an absolute score that was indicative of mild pain.

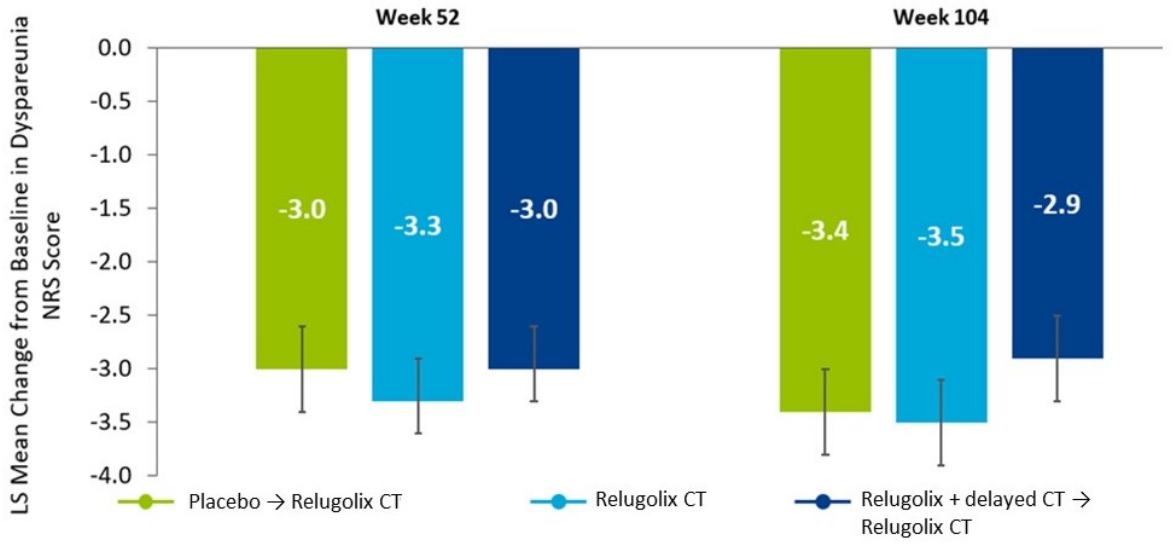
© Gedeon Richter UK Ltd (2023). All rights reserved

Secondary efficacy endpoint: Change from baseline in the mean dyspareunia NRS score at Week 52 and Week 104 (Extension Study Population)

The secondary endpoints evaluating dyspareunia was the change from baseline to Week 52 and from baseline to Week 104 in the dyspareunia NRS scores among patients reporting at least one vaginal sexual intercourse with a non-zero pain score at baseline. Patients were to report whether they had vaginal sexual intercourse and, if so, their level of pelvic pain during vaginal sexual intercourse on an 11-point NRS (0 = no pain to 10 = pain as bad as you can imagine) daily in an eDiary.

The change from baseline in average dyspareunia NRS score is shown in Figure 22 (64).

Figure 22: LS Mean Change in Dyspareunia NRS Score from Baseline to Week 104



The baseline LS mean (SE) dyspareunia NRS scores were 5.9 (0.17) for the Relugolix CT group and 5.8 (0.17) for the placebo group (64).

In the Relugolix CT group, the LS mean NRS score decreased at Week 52 to 2.5 (0.19), a change from baseline of -3.3 (95% CI: -3.6, -2.9), representing a 62.7% decrease and a reduction in pain from moderate to mild. Similar reductions were observed for Week 104 (-3.5, 58.6% decrease) (64, 65).

© Gedeon Richter UK Ltd (2023). All rights reserved

In the placebo group, the LS mean change from baseline in the dyspareunia NRS scores were numerically less than those in the relugolix groups through Week 28. At Week 52, the LS mean change in dyspareunia in this group was in the same range as in the Relugolix CT group (-3.0 [95% CI: -3.4, -2.6]) and sustained through Week 104(64, 65).

Dyspareunia improved with Relugolix CT treatment and the improvement in pain continued to steadily decline through Week 52 and was subsequently sustained through Week 104. The reduction in pain was substantial. At Week 52 and Week 104, there was an approximately 62% reduction from baseline in dyspareunia with an absolute score that was indicative of mild pain.

Secondary efficacy endpoint: Proportion of patients not using protocol-specified opioids for endometriosis-associated pain at Week 52 and Week 104 (Extension Study Population)

The secondary endpoints evaluating opioid use was based on the proportion of patients who were not using protocol-specified opioids for endometriosis-associated pain at Week 52 and at Week 104/EOT. Protocol-specified opioid use is presented in Table 22 (64, 65).

Table 22: Protocol-Specified Opioid Use at Week 52 and Week 104/EOT (Extension Study Population)

	Relugolix CT	**Placebo → Relugolix CT**
Run-in Period, n	277	275
Number (%) of patients using opioid[1]	109 (39.4%)	95 (34.5%)
Pain Assessment Period (Week 52), n	233	234
Number (%) of patients using opioid[1]	32 (13.7%)	26 (11.1%)
Number (%) of patients not using opioid (95% CI)[2]	201 (86.3%) (81.2%, 90.4%)	208 (88.9%) (84.1%, 92.6%)
Pain Assessment Period (Week 104/EOT), n	277	275
Number (%) of patients using opioid[1]	25 (9.0%)	26 (9.5%)
Number (%) of patients not using opioid (95% CI)[2]	252 (91.0%) (87.0%, 94.1%)	249 (90.5%) (86.5%, 93.7%)

[1]Tier 2 analgesic (ie, tramadol (37.5 mg) / paracetamol (325 mg), tramadol (50 mg), codeine (30 mg), codeine (30 mg) / paracetamol (300 mg), codeine (30 mg) / paracetamol (500 mg), codeine (15 mg) / paracetamol (500 mg), and hydrocodone (5 mg) / acetaminophen (325 mg)). [2]Based on exact binomial 95% CI (Clopper-Pearson)

© Gedeon Richter UK Ltd (2023). All rights reserved

At baseline, 39.4% (109 patients) and 34.5% (95 patients) in the Relugolix CT group and placebo group, respectively were receiving opioids (Table 22).

In the Relugolix CT group:

- The proportion of patients receiving opioids declined, to [redacted] at Week 24 and this decline was sustained at Week 52 (13.7% [32 patients]).
- At Week 104/EOT, the percentage receiving opioids was 9.0% (25 patients), representing a relative reduction from baseline by 65.2% at Week 52 (from 39.4% to 13.7%) and 77.1% (from 39.4% to 9.0%) at Week 104/EOT in the proportion of patients receiving opioids.
- The percentage of patients meeting the opioid-free endpoint at Week 52 and Week 104/EOT was 86.3% (95% CI: 81.2, 90.4) and 91.0 (95% CI: 87.0, 94.1), respectively (Table 22).

In the placebo group:

- The proportion of patients receiving opioids was [redacted] at Week 24
- This percentage declined, after transition to Relugolix CT to 11.1% at Week 52 and to 9.5% (26 patients) at Week 104/EOT, representing a relative reduction from baseline by 72.4% (from 34.5% to 9.5%) in the proportion of patients receiving opioids at Week 104/EOT
- The percentage of patients meeting the opioid-free endpoint at Week 52 and Week 104/EOT were 88.9% and 90.5%, respectively (Table 22).

These data support and extend the findings from the primary endpoints and overall pelvic pain, and dyspareunia secondary endpoint analyses by showing that over 52 weeks and 104 weeks of treatment with Relugolix CT, reductions in dysmenorrhoea, NMPP, overall pelvic pain, and dyspareunia were achieved, enabling a relative reduction in the percentage of patients using opioids by 65% at Week 52 and 77% at Week 104/EOT. In absolute terms, 86.3% of patients at Week 52 and 91.0% of patients at Week 104/EOT were opioid-free.

© Gedeon Richter UK Ltd (2023). All rights reserved

Secondary efficacy endpoint: Proportion of patients not using analgesics for Endometriosis-associated pain at Week 52 and Week 104 (Extension Study Population)

The proportion of patients not using analgesics at Week 52 and Week 104/EOT is provided Table 23 (64, 65).

Table 23: Protocol-specified analgesic use at Week 52 and Week 104/EOT (extension study population)

	Relugolix CT	**Placebo → Relugolix CT**
Run-in Period, n	277	275
Number (%) of patients using analgesics	257 (92.8%)	255 (92.7%)
Tier1 Analgesics	[redacted]	[redacted]
Tier2 Analgesics	[redacted]	[redacted]
Number (%) of patients not using analgesics[1]	20 (7.2%)	20 (7.3%)
Pain Assessment Period (Week 52), n	233	234
Number (%) of patients using analgesics	82 (35.2%)	69 (29.5%)
Tier1 Analgesics	[redacted]	[redacted]
Tier2 Analgesics	[redacted]	[redacted]
Number (%) of patients not using analgesics[1]	151 (64.8%)	165 (70.5%)
(95% CI)[2]	(58.3%, 70.9%)	(64.2%, 76.3%)
Pain Assessment Period (Week 104/EOT), n	277	275
Number (%) of patients using analgesics	69 (24.9%)	66 (24.0%)
Tier1 Analgesics	[redacted]	[redacted]
Tier2 Analgesics	[redacted]	[redacted]
Number (%) of patients not using analgesics[1]	208 (75.1%)	209 (76.0%)
(95% CI)[2]	(69.6%, 80.1%)	(70.5%, 80.9%)

[1]Patients who were not using either Tier 1 analgesic (Ibuprofen (200 mg dose strength)) or Tier 2 analgesics (tramadol (37.5 mg) / paracetamol (325 mg), tramadol (50 mg), codeine (30 mg), codeine (30 mg) / paracetamol (300 mg), codeine (30 mg) / paracetamol (500 mg), codeine (15 mg) / paracetamol (500 mg), and hydrocodone (5 mg) / acetaminophen (325 mg)). [2]Based on exact binomial 95% CI (Clopper-Pearson)

At baseline, 92.8% (257 patients) and 92.7% (255 patients) in the Relugolix CT and placebo groups, respectively, were receiving Tier 1 and/or Tier 2 analgesics.

© Gedeon Richter UK Ltd (2023). All rights reserved

In the Relugolix CT group:

- The proportion of patients receiving any analgesic (i.e., Tier 1 and/or Tier 2) declined to [redacted] at Week 24, and further declined at Week 52 (35.2% [82 patients]).

- At Week 104/EOT, the percentage receiving analgesics was 24.9% (69 patients), representing a relative reduction from baseline by 62.1% (from 92.8% to 35.2%) at Week 52 and 73.1% (from 92.8% to 24.9%) at Week 104 in the proportion of patients receiving any analgesics.

- The percentage of patients meeting the analgesic-free endpoint at Week 52 and Week 104/EOT was 64.8% (95% CI: 58.3, 70.9) and 75.1% (95% CI: 69.6, 80.1), respectively (Table 23).

In the placebo group:

- The proportion of patients receiving any analgesic was [redacted] at Week 24. This percentage declined, after transition to Relugolix CT, to 29.5% at Week 52 and to 24.0% (66 patients) at Week 104/EOT, representing a relative reduction from baseline by 74.1% (from 92.7% to 24.0%) in the proportion of patients receiving analgesics.

- The percentage of patients meeting the analgesic-free endpoint at Week 52 and Week 104/EOT was 70.5% and 76.0%, respectively (Table 23).

These data support and extend the findings from the primary endpoints, overall pelvic pain and dyspareunia secondary endpoint analyses by showing that over 52 weeks and 104 weeks of treatment with Relugolix CT, reductions in dysmenorrhoea, NMPP, overall pelvic pain, and dyspareunia were achieved, enabling a relative reduction in the percentage of patients using any analgesics by over 62% at Week 52 and 73% at Week 104/EOT. In absolute terms, 65% of patients at Week 52 and 75% of patients at Week 104/EOT were analgesic-free.

© Gedeon Richter UK Ltd (2023). All rights reserved

B.2.7 Subgroup analysis

No subgroups were analysed in the economic analyses, however, please find below the subgroups analysed within the SPIRIT 1, SPIRIT 2 and SPIRIT OLE clinical trials.

SPIRIT 1 and SPIRIT 2 (62, 63)

Subgroup analyses were conducted for the co-primary efficacy endpoints by geographic region, time since surgical diagnosis of endometriosis, AFS endometriosis stage, age, race, BMI, smoking status, dysmenorrhoea NRS score at baseline, NMPP NRS score at baseline, and renal function based on the Cockcroft-Gault formula for calculated creatinine clearance.

Consistent with the findings for the overall population, treatment differences with regard to the co-primary endpoints were consistent across nearly all subgroups as demonstrated by the odds ratio point estimate consistently favouring Relugolix CT over placebo on the dysmenorrhoea and NMPP co-primary endpoints.

Together, these data provide support for the efficacy of Relugolix CT across age groups, race, BMI, level of pain at baseline, disease duration, renal function, smoking status, and geography.

SPIRIT OLE (64)

The dysmenorrhoea responder and NMPP responder primary endpoints at Week 104/EOT were analysed by predefined subgroups of the study population including geographic region, age, race, and baseline BMI. For patients reporting multiple races, those who reported “Black/African-American” as one of the races were included in the “Black or African-American” category. Subgroup analyses for patients in the Relugolix CT group are presented in Table 24 and Table 25 for the dysmenorrhoea responder and NMPP responder endpoints, respectively.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 24: SPIRIT OLE Proportion of Patients Classified as Dysmenorrhoea Responders at Week 104/EOT, Subgroup Analyses, Relugolix CT Group

Subgroups	Category	Number of Evaluable Patients	Number (%) of Responders	95% CI
Overall		277	235 (84.8%)	(80.06%, 88.85%)
Geographic region	North America Rest of the World	48 229	35 (72.9%) 200 (87.3%)	(58.15%, 84.72%) (82.32%, 91.35%)
Age (years)	< 35 years >= 35 years	142 135	114 (80.3%) 121 (89.6%)	(72.78%, 86.48%) (83.21%, 94.21%)
Race	Black/African American White	18 258	14 (77.8%) 220 (85.3%)	(52.36%, 93.59%) (80.35%, 89.36%)
BMI (kg/m^2) at baseline	< 25 25 - <30 >=30	161 65 51	142 (88.2%) 56 (86.2%) 37 (72.5%)	(82.19%, 92.74%) (75.34%, 93.47%) (58.26%, 84.11%)

Table 25:SPIRIT OLE Proportion of Patients Classified as Non-menstrual Pelvic Pain Responders at Week 104/EOT, Subgroup Analyses, Relugolix CT Group

Subgroups	Category	Number of Evaluable Patients	Number (%) of Responders	95% CI
Overall		277	210 (75.8%)	(70.33%, 80.74%)
Geographic region	North America Rest of the World	48 229	33 (68.8%) 177 (77.3%)	(53.75%, 81.34%) (71.31%, 82.55%)
Age (years)	< 35 years >= 35 years	142 135	108 (76.1%) 102 (75.6%)	(68.18%, 82.81%) (67.42%, 82.54%)
Race	Black/African American White	18 258	14 (77.8%) 196 (76.0%)	(52.36%, 93.59%) (70.28%, 81.05%)
BMI (kg/m^2) at baseline	< 25 25 - <30 >=30	161 65 51	125 (77.6%) 49 (75.4%) 36 (70.6%)	(70.41%, 83.82%) (63.13%, 85.23%) (56.17%, 82.51%)

Some subgroup analyses (e.g. 5 categories of BMI) with many categories yielded subgroups < 30 patients within a treatment group and greater variability. Nevertheless, in the Relugolix CT group, all subgroups, for both primary endpoints (dysmenorrhoea and NMPP) showed consistent point estimates and confidence intervals, overlapping with those of the overall population.

Together, these data provide support for the overall efficacy of Relugolix CT. The responder rate is comparable in each of the subgroups to that of the overall population.

© Gedeon Richter UK Ltd (2023). All rights reserved

B.2.8 Meta-analysis

Not applicable

B.2.9 Indirect and mixed treatment comparisons

Since direct head-to-head randomised control trial (RCT) data is not available, an indirect treatment comparison (ITC) was conducted to compare the efficacy of Relugolix CT with comparator therapies for the treatment of endometriosis-associated pain.

The initial SLR identified a total of 58 studies. To be included in the ITC, studies had to fulfil the following criteria:

- Directly connect a comparator of interest to the intervention Relugolix CT, or
- Indirectly connect Relugolix CT with a comparator of interest (e.g., through placebo).

Studies that did not fulfil these two criteria (i.e., that were disconnected from the Relugolix CT network) were excluded.

A summary of the ITC methods is given below. Full details of the methods are given in Appendix D.

No separate network was synthesized for dysmenorrhea owing to inconsistencies in the this was measured across the trials. However, dysmenorrhea was captured in the ITC as an element of the TPP endpoint. In addition, analgesic and opioid use were not included in the ITC owing to the amount of heterogeneity between studies in terms of permitted use and reporting of use.

Overall pelvic pain

Three studies identified in the SLR reported results for OPP; all were eligible for inclusion in the ITC (Table 26).

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 26: Summary of the trials used to carry out the indirect or mixed treatment comparison for OPP

	Relugolix CT (relugolix 40 mg with oestradiol 1 mg and norethisterone acetate 0.5 mg	**Placebo**	**Leuprolide acetate 3.75 mg**	***ASP1707 / Opigolix 3 mg, 5 mg, 10 mg**
SPIRIT 1&2	Yes	Yes		
D'Hooghe et al., 2019		Yes	Yes	Yes

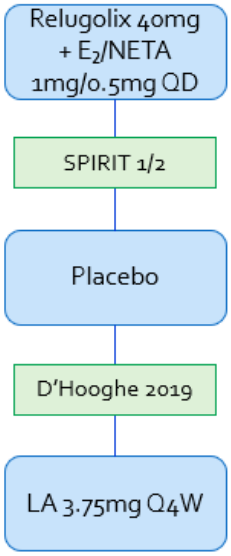
Opigolix arm was excluded from the ITC

D'Hooghe et al (2019) was a Phase II, multicentre, double-blind, randomized, parallel group, placebo-controlled study comparing the efficacy and safety of ASP1707 (Opigolix) (3 mg, 5 mg, 10 mg, 15 mg), leuprolide acetate (3.75 mg), and placebo in 540 women with endometriosis-associated pain. The 24-week assessment period was divided into two 12-week parts. In part 1, subjects received either once-daily oral ASP1707 tables, monthly subcutaneous leuprolide acetate or once-daily placebo tablets. In part 2, patients in the placebo group were re-randomized to either ASP1707 or leuprolide acetate and patients in the active treatment groups continued with their allocated treatment (59). The study was carried out in Europe and Japan.

The SPIRIT 1&2 trials connected Relugolix CT to placebo. D'Hooghe et al. 2019, connected placebo to leuprolide acetate (LA) 3.75 mg Q4W (59). The full connected network is presented in Figure 23.

© Gedeon Richter UK Ltd (2023). All rights reserved

Figure 23:Evidence network of all connected studies reporting information on OPP at week 12



The NRS, an 11-point Likert scale ranging from 0 (no pain) to 10 (worst possible pain), was used to measure pain in all three trials. All three studies reported the results on a continuous scale.

The studies reported results at different time points. The SPIRIT 1&2 trials reported results at 4, 8, 12, 16, 20 and 24 weeks, whereas the study by D'Hooghe et al. 2019 reported results at 4, 8 and 12 weeks. The network analysis of OPP was therefore constrained to data reported at the 12-week time point.

Total pelvic pain

Eleven studies identified in the SLR reported results for TPP; four were eligible for inclusion in the ITC (Table 27). A list of the seven excluded studies is provided in Appendix D.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 27: Summary of the trials used to carry out the indirect or mixed treatment comparison for TPP

	Relugolix CT (relugolix 40 mg with oestradiol 1 mg and norethisterone acetate 0.5 mg	**Placebo**	**Dienogest 2mg**	**Leuprolide acetate 3.75mg**
SPIRIT 1&2	Yes	Yes		
Lang 2018		Yes	Yes	
Strowitzki et al., 2010			Yes	Yes

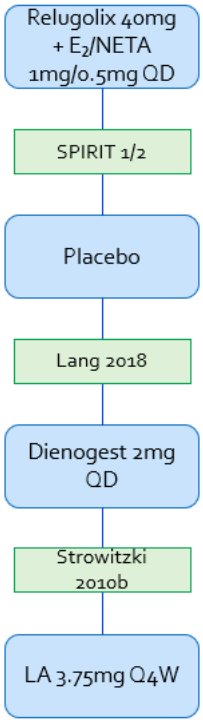
Lang et al. 2018 was a 24-week, Phase 3, randomized, double blind, placebo-controlled multicentre study to evaluate the efficacy and safety of 2 mg dienogest once-daily in 255 women in China aged 18-45 with laproscopically-diagnosed endometriosis and endometriosis-associated pelvic pain (60).

Strowitzki et al. 2010 was a 24-week randomized, multicentre, open-label trial comparing dienogest 2 mg with leuprolide acetate in women aged 18-45 years. The study was conducted at 17 centres in Germany, Austria, Spain, Poland, Italy, and Portugal (61).

The SPIRIT 1 & 2 trials connected Relugolix CT to placebo. Lang et al. 2018 connected dienogest 2 mg to placebo (60), and Strowitzki et al. 2010 (61) connected dienogest 2 mg to leuprolide acetate 3.75 mg Q4W. The full connected network is presented in Figure 24.

© Gedeon Richter UK Ltd (2023). All rights reserved

Figure 24: Evidence network of all connected studies reporting information on TPP at week 24



The trials evaluated total pelvic pain using the Biberoglu-Behrman score (B&B) or a modified version of the B&B score. Total pelvic pain severity is usually measured as the combination of three patient assessed pain symptoms (dysmenorrhoea, NMPP/PP and dyspareunia). The SPIRIT 1&2 trials captured TPP as the sum of dysmenorrhea, dyspareunia and NMPP. Each symptom was scored between 0 and 3, corresponding to absent, mild, moderate, and severe. SPIRIT 1 & 2 and Lang et al 2018 reported the results using a continuous scale (60, 62, 63). Strowitzki et al 2010 reported the proportion of patients reporting pain as none, mild, moderate, severe, and very severe (61).

The network analysis of TPP considered results at 24 weeks only.

Combined network analysis for OPP and TPP

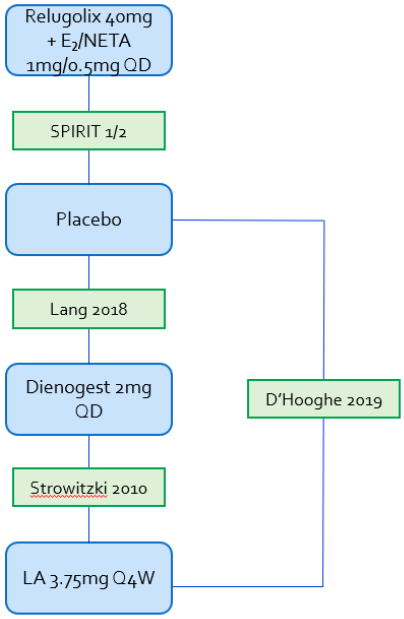
The secondary objective of the ITC was to synthesize the available evidence for the treatment efficacy of relugolix-CT compared with GnRH agonists by pooling the networks for OPP and TPP to achieve a greater number of studies included in a single network. For this analysis, it was assumed that the clinical endpoints OPP and TPP are sufficiently similar to allow for a combined analysis. The rationale for the

© Gedeon Richter UK Ltd (2023). All rights reserved

above-mentioned assumption was discussed and agreed on during a global advisory board meeting involving clinical experts within endometriosis, where the clinical experts agreed that the treatment effect would be similar on the TPP and OPP outcome scale (66).

All studies that were included in the final networks for OPP and TPP were included in a pooled network for OPP and TPP (Figure 25). Owing to limited data availability, the study endpoints were included irrespective of the time point.

Figure 25:Combined evidence network of all connected studies reporting information on OPP and TPP



Abbreviations: E_2 = Estradiol, LA = Leuprolide Acetate, NETA = Norethisterone Acetate, Q4W = Every 4 weeks, QD = Once daily.

Results of the ITC

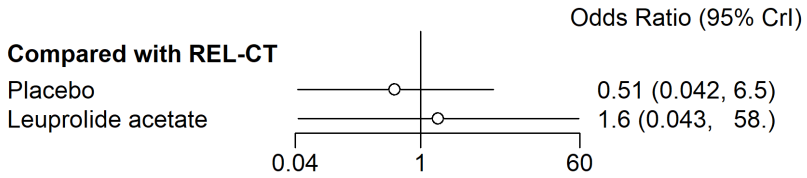
OPP

Base case

The forest plot showing the odds ratio for OPP at 12 weeks is presented in Figure 26. No significant differences were found in terms of the treatment effect on OPP between Relugolix CT, placebo, and LA. The odds ratios for the two comparators were not found to be statistically significant, with wide credible intervals that encompassed the null value of one.

© Gedeon Richter UK Ltd (2023). All rights reserved

Figure 26: Forest plot of odds ratios for OPP (weakly informative priors)

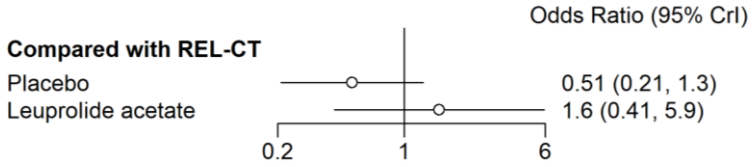


Abbreviations: CrI: credible intervals, REL-CT: Relugolix combination therapy
Note: An odds ratio on the left side of the vertical line indicates that the comparison favors REL-CT, an odds ratio on the right side favors the comparator treatment or placebo.

Sensitivity analysis

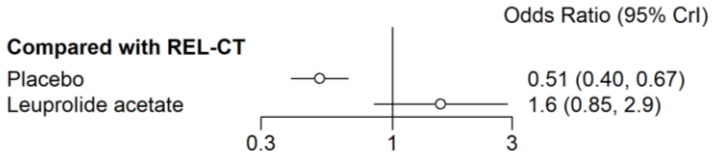
Except for the odds ratio of placebo compared to Relugolix CT estimated using a fixed effects model, no statistically significant difference could be found in the scenario analyses (Figure 27 to Figure 29). The fixed effects model resulted in an odds ratio of below one, indicating better treatment outcomes with Relugolix CT compared to placebo.

Figure 27:Forest plot of odds ratios for OPP (empirical priors)



Abbreviations: CrI: credible intervals, REL-CT: Relugolix combination therapy
Note: An odds ratio on the left side of the vertical line indicates that the comparison favors REL-CT, an odds ratio on the right side favors the comparator treatment or placebo.

Figure 28: Forest plot of odds ratios for OPP (fixed effects)



Abbreviations: CrI: credible intervals, REL-CT: Relugolix combination therapy
Note: An odds ratio on the left side of the vertical line indicates that the comparison favors REL-CT, an odds ratio on the right side favors the comparator treatment or placebo.

OPP model fit and convergence

The assessment of model fit was based on the posterior mean of the residual deviance and DIC, with lower values generally indicating better fit and less unnecessary complexity (67). There was no meaningful difference in the DIC across

© Gedeon Richter UK Ltd (2023). All rights reserved

the models, indicating similar model fit after adjusting for model complexity. All models resulted in a ratio close to one when considering the posterior mean of the residual deviance and the number of data points, indicating a sufficient model fit for each model (Table 28).

Table 28: Summary of fit statistics for models evaluated for the NMA of OPP

Model type	Priors	Posterior mean of the residual deviance	Ratio of the posterior mean of the residual deviance and data points (n=3)	DIC
Random effects	Uniform (0, 2)	1.998	0.999	3.97
Random effects	Log-normal (-3.23, 1.88)	2.002	1.001	3.98
Fixed effects	Not applicable	5.930	1.483	3.99

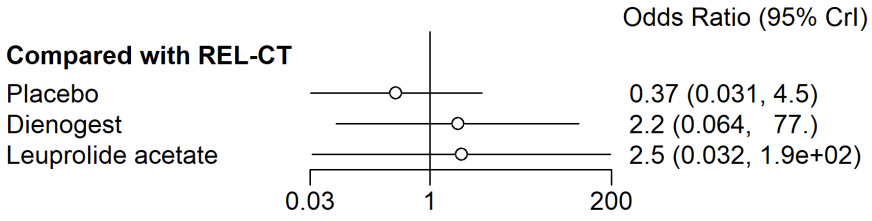
Abbreviations: DIC=Deviance Information Criterion

TPP

Base case

The forest plot in Figure 29 shows the results of the NMA comparing the effects of Relugolix CT to placebo, dienogest, and LA in terms of relative treatment effect measured by TPP at 24 weeks. The results showed very wide credible intervals for the estimates and none of the treatments showed a statistically significant difference, suggesting that there is no difference in TPP between Relugolix CT, placebo, dienogest and LA.

Figure 29: Forest plot of odds ratios for TPP (weekly informative priors)



Abbreviations: CrI: credible intervals, REL-CT: Relugolix combination therapy
Note: An odds ratio on the left side of the vertical line indicates that the comparison favors REL-CT, an odds ratio on the right side favors the comparator treatment or placebo.

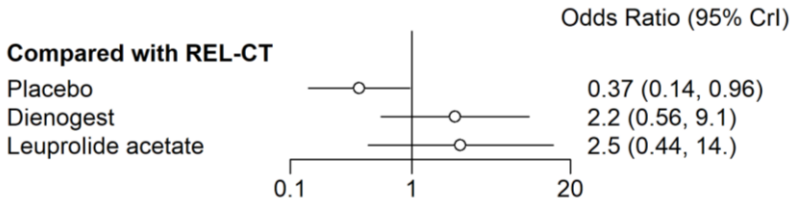
Sensitivity analysis

Unlike the base case analysis, both scenario analyses for TPP showed a statistically significant difference between placebo and Relugolix CT, favouring Relugolix CT (Figure 30 and Figure 31). The use of a fixed effects model resulted in a statistically

© Gedeon Richter UK Ltd (2023). All rights reserved

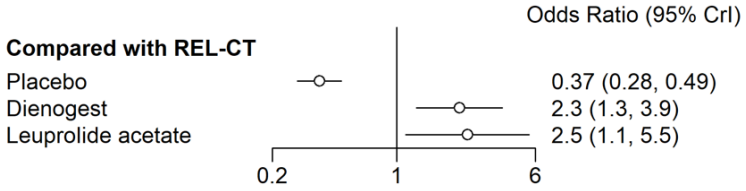
significant difference for dienogest and LA (odds ratios above 1) compared to Relugolix CT, indicating improved treatment outcomes for dienogest and LA.

Figure 30: Forest plot of odds ratios for TPP (empirical priors)



Abbreviations: CrI: credible intervals, REL-CT: Relugolix combination therapy
Note: An odds ratio on the left side of the vertical line indicates that the comparison favors REL-CT, an odds ratio on the right side favours the comparator treatment or placebo.

Figure 31: Forest plot of odds ratios for TPP (fixed effects)



Abbreviations: CrI: credible intervals, REL-CT: Relugolix combination therapy
Note: An odds ratio on the left side of the vertical line indicates that the comparison favors REL-CT, an odds ratio on the right side favours the comparator treatment or placebo.

TPP model fit and convergence

There was no meaningful difference in the DIC across the models, indicating similar model fit after adjusting for model complexity. All models resulted in a ratio close to one when considering the posterior mean of the residual deviance and the number of data points, indicating a sufficient model fit for each model (Table 29).

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 29: Summary of fit statistics for models evaluated for the NMA of TPP

Model type	Priors	Posterior mean of the residual deviance	Ratio of the posterior mean of the residual deviance and data points (n=3)	DIC
Random effects	Uniform (0, 2)	2.997	0.999	5.98
Random effects	Log-normal (-3.23, 1.88)	2.995	0.998	6.00
Fixed effects	Not applicable	2.997	0.999	5.99

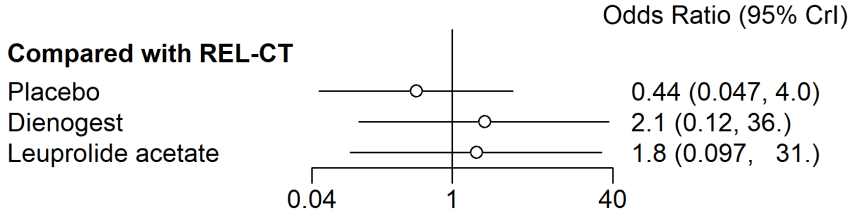
Abbreviations: DIC=Deviance Information Criterion

Combined evidence synthesis using OPP values from SPIRIT 1&2

The forest plot in Figure 32 shows the results of the combined NMA comparing the effects of Relugolix CT with placebo, dienogest and LA, using weakly informative priors (at 24 weeks). None of the odds ratios were found to be statistically significant, with wide credible intervals that encompassed the null value of one. This suggests that the true difference in effect of these interventions and placebo on the outcome is uncertain. The same was observed for the model using empirical priors (scenario analysis, Figure 33). No statistically significant difference between the treatments could be observed. Only the fixed effects model resulted in statistically significant differences between Relugolix CT and placebo (Figure 34). The odds ratios comparing dienogest and LA to Relugolix CT, resulted in point estimates above one and hence indicated an improvement of the treatment outcomes when compared to Relugolix CT.

Base case

Figure 32: Forest plot of odds ratios for the combined network - OPP (weakly informative priors)

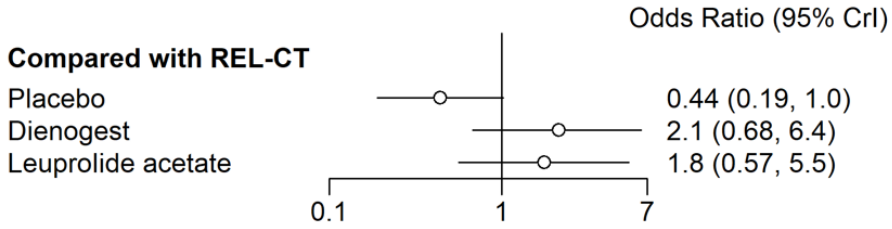


Abbreviations: CrI: credible intervals, REL-CT: Relugolix combination therapy
Note: An odds ratio on the left side of the vertical line indicates that the comparison favors REL-CT, an odds ratio on the right side favors the comparator treatment or placebo.

© Gedeon Richter UK Ltd (2023). All rights reserved

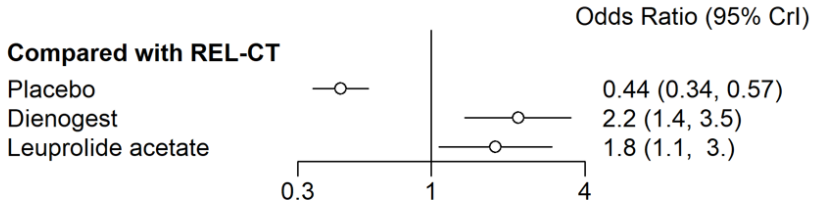
Scenario analysis

Figure 33: Forest plot of odds ratios for the combined network - OPP (empirical priors)



Abbreviations: CrI: credible intervals, REL-CT: Relugolix combination therapy
Note: An odds ratio on the left side of the vertical line indicates that the comparison favors REL-CT, an odds ratio on the right side favors the comparator treatment or placebo.

Figure 34: Forest plot of odds ratios for the combined netwrok - OPP (fixed effects)



Abbreviations: CrI: credible intervals, REL-CT: Relugolix combination therapy
Note: An odds ratio on the left side of the vertical line indicates that the comparison favors REL-CT, an odds ratio on the right side favors the comparator treatment or placebo.

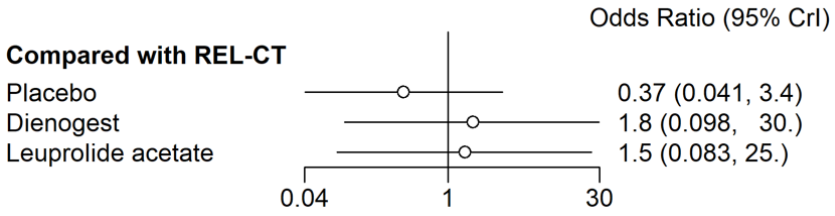
Combined ITC using TPP values from SPIRIT 1&2

Similar to the combined analysis using the TPP data from the SPIRIT 1&2 trials, no statistically significant differences could be observed between Relugolix CT, the included treatments and placebo (base case analysis, Figure 35). However, the two alternative models that were explored in a scenario analysis showed statistically significant differences between Relugolix CT and placebo, indicating improved treatment outcomes with Relugolix CT versus placebo (the random effects model using empirical prior, Figure 36 and the fixed effects model, Figure 37). There was also a significant difference in the treatment effect for dienogest versus Relugolix CT (the fixed effects model, Figure 37), indicating an improved treatment outcome with dienogest.

© Gedeon Richter UK Ltd (2023). All rights reserved

Base case

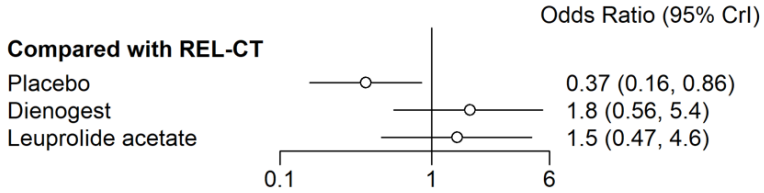
Figure 35:Forest plot of odds ratios for the combined network - TPP (weakly informative priors)



Abbreviations: CrI: credible intervals, REL-CT: Relugolix combination therapy
Note: An odds ratio on the left side of the vertical line indicates that the comparison favors REL-CT, an odds ratio on the right side favors the comparator treatment or placebo.

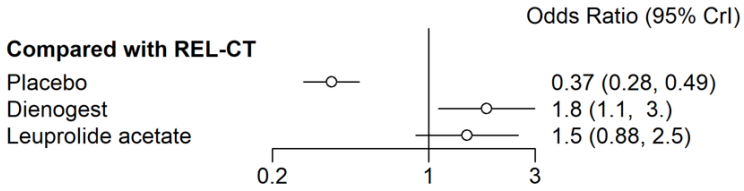
Scenario analysis

Figure 36: Forest plot of odds ratios for the combined network - TPP (empirical priors)



Abbreviations: CrI: credible intervals, REL-CT: Relugolix combination therapy
Note: An odds ratio on the left side of the vertical line indicates that the comparison favors REL-CT, an odds ratio on the right side favors the comparator treatment or placebo.

Figure 37: Forest plot of odds ratios for the combined network - TPP (fixed effects)



Abbreviations: CrI: credible intervals, REL-CT: Relugolix combination therapy
Note: An odds ratio on the left side of the vertical line indicates that the comparison favors REL-CT, an odds ratio on the right side favors the comparator treatment or placebo.

Heterogeneity in the OPP NMA studies

An underlying assumption of the network meta-analysis is similarity of studies, i.e., clinical trials are assumed to be similar in characteristics other than the intervention itself. To ensure a robust comparison, potential imbalances in terms of baseline clinical and disease characteristics from relevant clinical trials were assessed. This included age, race, and baseline severity scores for NMPP, dysmenorrhoea, OPP

© Gedeon Richter UK Ltd (2023). All rights reserved

and TPP, as well as prior and concomitant medications, with a focus on the use of analgesics (full details of the data assessed are provided in **Error! Reference source not found.**, **Error! Reference source not found.**, and **Error! Reference source not found.** in Appendix D.1.1).

Demographic characteristics

All participants in the trials were pre-menopausal females. The mean age across studies and relevant treatment arms was similar, ranging from approximately 33 to 34 years. Where reported, patients were primarily white (up to 91.9% in one arm of the SPIRIT 1 trial). The same SPIRIT 1 arm also had the lowest percentage of black patients (4.7%). The differences in the racial makeup of the included studies may be partially due to the geographic location of the studies (SPIRIT 1&2 included research centres in Africa, Australasia, Europe, North America, and South America while D'Hooghe et al. 2019 included only centres in Europe and Japan) (37, 59). Research suggests that black and Hispanic women are less likely to be diagnosed with endometriosis which may translate into a general underrepresentation of black or Hispanic women in clinical studies (57).

Severity of pain at baseline

All included studies provided information on the baseline severity of pain. The SPIRIT 1&2 trials reported baseline scores for various pain categories, including dysmenorrhea (7.1 for placebo and 7.2 for Relugolix CT, respectively), dyspareunia (5.7 for placebo and Relugolix CT, respectively), NMPP (5.8 and 5.9 for placebo and Relugolix CT, respectively), OPP (6.1 and 6.0 for placebo and Relugolix CT, respectively) using the NRS scale, and TPP (4.9 and 5.2 for placebo and Relugolix CT, respectively) using the mB&B scale. The study by D'Hooghe et al. 2019 reported baseline scores for OPP of approximately 4 on the NRS scale (4.2 for placebo and 4.1 for LA, respectively) at baseline.

The baseline severity levels of the study populations appear different, with the SPIRIT 1&2 trials generally reporting increased and more severe baseline scores for OPP compared to the D'Hooghe et al. 2019 study. The latter study provided only limited information in this regard.

© Gedeon Richter UK Ltd (2023). All rights reserved

Prior treatments

Prior treatments included both medication treatment and surgical treatment. In the SPIRIT 1&2 trials, close to all patients had received at least one prior medication before study begin (>98% of patients). Additionally, the majority of patients had received at least one surgical treatment for their endometriosis (>79% of patients) prior to the study.

As part of the exclusion criteria, the SPIRIT 1&2 trials did not include patients that received hormonal treatment and specific non-hormonal treatments for the management of endometriosis within pre-defined time periods prior to the study initiation. The use of estrogens and intrauterine devices was prohibited from 56 days prior to study initiation. Hormonal contraceptives and aromatase inhibitors were not allowed from 28 days and anti-androgens were not allowed from 12 weeks prior to study initiation. The use of GnRH analogues was not allowed from 35 days prior to study initiation.

D'Hooghe 2019 et al. reported that up to 57% of patients received prior medication treatment and 97% of patients received surgical treatment before participating in the clinical trial. Patients were not eligible for the study if they received hormonal treatment or other treatments with effects on gynaecological endocrinology within four weeks prior to the start of screening. Other treatments such as depotmedroxyprogesterone acetate or danazol as well as anticoagulants or drugs with effects on BMD were prohibited within 12 weeks prior to the start of screening, and the use of GnRH agonists had to be terminated 24 weeks prior to the start of screening. Any surgery had to be at least four months before signing informed consent. Previous hysterectomy or bilateral oophorectomy were an exclusion criterion.

Concomitant treatment

The use of analgesics (concomitant) medication differed across the studies. While participants in the SPIRIT 1&2 trials were allowed to use non-steroidal anti-inflammatory drug (NSAIDs) and opioids, participants in the D'Hooghe et al. 2019 study were only allowed to use a NSAID (ibuprofen).

© Gedeon Richter UK Ltd (2023). All rights reserved

Heterogeneity in the TPP NMA studies

As for OPP, potential imbalances in terms of baseline clinical and disease characteristics from relevant clinical trials were assessed. This included age, race, and baseline severity scores for dysmenorrhea, dyspareunia, NMPP, OPP and TPP, as well as prior and concomitant medications, with a focus on the use of analgesics (full details of the data assessed are provided in **Error! Reference source not found.**, **Error! Reference source not found.** and **Error! Reference source not found.** in Appendix D.1.1).

Demographic characteristics

All participants in the trials were pre-menopausal females. The mean age across studies was similar, ranging from approximately 31 to 34 years. Race was only reported for SPIRIT 1&2, which included primarily white women (up to 91.9% in one arm of the SPIRIT 1 trial).

Severity of pain at baseline

All included studies provided information on the baseline severity of pain. The SPIRIT 1&2 trials were particularly informative, reporting baseline scores for various pain categories, including dysmenorrhoea (7.1 for placebo and 7.2 for Relugolix CT, respectively), dyspareunia (5.7 for placebo and Relugolix CT, respectively), NMPP (5.8 and 5.9 for placebo and Relugolix CT, respectively), OPP (6.1 and 6.0 for placebo and Relugolix CT, respectively) based on the NRS scale, and TPP (4.9 and 5.2 for placebo and Relugolix CT, respectively) based on the mB&B scale. The study by Lang et al. 2018 (60) reported baseline scores for TPP based on the B&B scale, 4.3 for dienogest and 4.4 for placebo indicating moderate pain (68)). Strowitzki and colleagues (61) reported baseline OPP based on VAS (53.3 mm for dienogest and 55.4 mm for LA). Baseline TPP was reported based on the B&B scale (including proportion pain categories rather than a score). In total, 69% of patients who received dienogest and 63% of patients who received LA reported severe or very severe pain at baseline. The remaining patients reported moderate (31% for dienogest and LA) or mild (6% for LA) TPP.

The baseline pain severity differed across studies. The SPIRIT 1&2 trials reported higher OPP scores compared to Strowitzki et al. 2010 (61) indicating more severe

© Gedeon Richter UK Ltd (2023). All rights reserved

pain levels at baseline. Only small differences became apparent regarding TPP between the SPIRIT 1&2 studies and the study published by Lang et al. 2018 (60). The reported TPP at baseline from the study by Strowitzki et al. 2010 (61) may not be comparable with the other studies due to the form of reporting (proportions). The discrepancy in baseline severity and limited information between studies is a source of between-study heterogeneity.

Prior treatments

Prior treatments included both medication treatment and surgical treatment. In the SPIRIT 1&2 trials, close to 100% of patients had received at least one prior medication before the study (>98% of patients). Additionally, most patients had received at least one surgical treatment (>79% of patients) for their endometriosis prior to the study.

As part of the exclusion criteria, the SPIRIT 1&2 trials did not include patients that received hormonal treatment and specific non-hormonal treatments for the management of endometriosis within pre-defined time periods prior to the study initiation and the use was not allowed during the study.

The study published by Lang et al. 2018 only reported surgical treatment for endometriosis. Similarly to the SPIRIT 1&2 trials, the majority of patients had received surgical treatment prior to study initiation (>86%). It is worth noticing that the proportion of patients that had received a prior surgical treatment in the dienogest arm in the Lang et al. 2018 study (94.4%) was considerably higher compared to the REL-CT arm in the SPIRIT 1&2 trials (≈80%). In the Lang et al. 2018 study women were excluded in case of recent use of hormonal agents (GnRH agonists within 6 months, long-acting agents such as depot progestins within 3 months, or short-acting agents such as oral contraceptives within one month) or required surgical treatment for endometriosis at the time of study inclusion. No use of previous treatment was reported by Strowitzki and colleagues. The exclusion criteria included amongst others previous use of hormonal agents (e.g., GnRH agonists ≤6 months, progestins or danazol ≤3 months or oral contraceptives ≤1 month before screening).

© Gedeon Richter UK Ltd (2023). All rights reserved

Concomitant treatment

The use of analgesics (concomitant) medication differed across the studies. While participants in the SPIRIT 1&2 trials were allowed to use NSAIDs and opioids, participants in the Lang et al. 2018 (60) study were only allowed to use an NSAID, namely ibuprofen. Patients in the SPIRIT 1&2 trials took ≤0.8 tablets per day (average), patients in the Lang et al. 2018 study took ≈1.6 tablets per day. It is worth noting that the use of NSAIDs differed considerably between the placebo groups at 24 weeks. While patients in the placebo arm in the SPIRIT 1&2 trials took 0.3 tablets on average at that time, patients in the Lang et al. 2018 study took 1.9 per day. However, the strength of the medication was not accounted for. Strowitzki et al. 2010 (61) mentioned that the use of concomitant medication, including analgesic medication, was allowed and that the medications recorded did not differ between the groups at baseline or during the trial. No further information was provided.

Uncertainties in the indirect and mixed treatment comparisons

OPP

As described above, it appears there is a considerable degree of between-study heterogeneity. The use of a fixed effects model was therefore considered inappropriate, and the results of the latter should be interpreted with caution. Instead, a random effects model was favoured as it allows to account for the between-study heterogeneity.

The NMA of the OPP network was linked to some limitations. The network only included three studies. The impact of an individual study on the NMA estimate and potential bias is large. This may be of importance considering that the change from baseline values and odds ratios were calculated based on graphical estimation in the D'Hooghe et al. 2019 study as they were not reported. The calculations were based on assumptions or approximations, which potentially introduced inaccuracies in the estimates. Inaccurate estimates can lead to biased results and affect the interpretation of the meta-analysis findings.

The risk of bias assessment (see Appendix D) did not show any potential risk of bias for the studies included in the OPP network. Lastly, as the OPP network was too small to assess network coherence, potential inconsistencies may exist.

© Gedeon Richter UK Ltd (2023). All rights reserved

TPP

The risk of bias assessment (see Appendix D) showed potential bias for the Strowitzki et al. 2010 study due to the absence of blinding. The study could not be removed from the network as it connected Relugolix CT with LA. The risk of bias remains, and the results of the comparison should be interpreted with caution. The same applies to the between study heterogeneity that could not be accounted for due to limited data availability. Due to the heterogeneity between studies, a fixed effects model was deemed not appropriate. The TPP network was too small to assess network coherence and thus potential inconsistencies may exist.

Generally, the NMA of the TPP network should be interpreted with care due to the risk of bias as described above as well as the small number of studies. When there is limited data, it becomes challenging to explore and account for the sources of heterogeneity, which ultimately can impact the reliability and generalisability of the results.

Combined network

Using all studies in the combined network decreased the point estimate of the OR for Relugolix CT vs LA (though still significantly different in the fixed effects combined network using OPP values from SPIRIT 1&2). However, the impact of the Lang et al. 2018 decreased by adding more evidence and the estimate may be more balanced compared to the findings from the TPP network alone.

The between-study heterogeneity and risk of bias as described above are true for the combined analysis as well. Consequently, a fixed effects model was deemed inappropriate. Moreover, there is potential bias due to the inclusion of the Strowitzki et al. 2010 study. Additionally, the assumption that OPP and TPP are sufficiently similar was based on clinical expert opinion. However, due to the different methods and instruments (e.g., B&B scale versus NRS) that were used to derive these measures, there may be relevant differences between these endpoints compromising the validity of the analysis. This, in turn, may result in a biased NMA estimate. The combined networks allowed for the assessment of network coherence and no inconsistencies were identified.

© Gedeon Richter UK Ltd (2023). All rights reserved

B.2.10 Adverse reactions

SPIRIT 1 and SPIRIT 2

Safety was evaluated by monitoring adverse events, clinical laboratory data, 12-lead ECGs, vital signs, physical examinations, menstrual bleeding patterns, pregnancy, overdose, BMD, and paired endometrial biopsies. An overview of the key safety endpoints is provided in Table 30. Relugolix CT maintains estradiol concentrations in a range that enables to control the adverse events, e.g., by maintaining BMD, while inhibiting endometriomal growth.

For completeness, the safety results for the relugolix + delayed CT arms are also presented in this section.

Table 30:Key safety endpoints for SPIRIT 1 and SPIRIT 2

Objective	Endpoint
Safety of 24 weeks of once-daily Relugolix CT or relugolix + delayed CT	To determine the safety of 24 weeks of Relugolix CT or relugolix + delayed CT
Change in BMD (lumbar spine) at Week 12	To determine the percent change from baseline to Week 12 in BMD at the lumbar spine (L1-L4) in Relugolix CT compared with relugolix + delayed CT
Change in BMD at Week 24	To determine the change in BMD after 24 weeks of treatment with Relugolix CT or relugolix + delayed CT
Incidence of vasomotor symptoms	To determine the incidence of vasomotor symptoms with Relugolix CT compared with relugolix + delayed CT through Week 12

Treatment emergent adverse events

The frequency of subjects who reported treatment-emergent adverse events (TEAEs) in the placebo group was similar to the Relugolix CT treatment groups. In SPIRIT 1, 140 (66.0%) subjects from the placebo group and 151 (71.2%) subjects from the Relugolix CT group experienced at least one TEAE while in SPIRIT 2, 153 (75.0%) subjects from the placebo group and 166 (80.6%) subjects from the Relugolix CT group experienced at least one TEAE (see Table 31)

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 31: Summary of adverse events in SPIRIT 1 and SPIRIT 2 (37, 62, 63)

Characteristics	SPIRIT 1						SPIRIT 2					
	Relugolix CT		Placebo		Relugolix + delayed CT		Relugolix CT		Placebo		Relugolix + delayed CT	
Any	151	71%	140	66%	163	77%	166	81%	153	75%	168	82%
Leading to discontinuation	8	4%	4	2%	9	4 %	11	5%	8	4%	15	7%
Leading to drug interruption	0		6	2.8%	3	1.4%	1	0.5%	4	2.0%	4	1.9%
Related to study drug	86	40.6%	73	34.4%	125	59.2%	112	54.4%	83	40.7%	117	56.8%
Grade 3 or above	10	5%	12	6%	9	4%	14	7%	7	3%	12	6%
Grade 3 or above related to study drug	6	2.8%	5	2.4%	5	2.4%	6	2.9%	3	1.5%	7	3.4%
Serious	3	1%	5	2%	3	1%	9	4%	4	2%	6	3%
Serious and related to study drug	0		0		1	0.5%	5	2.4%	1	0.5%	2	1.0%
Serious leading to discontinuation	1	0.5%	1	0.5%	1	0.5%	3	1.5%	4	2.0%	3	1.5%
Fatal outcome	0		0		0		0		0		0	

Generally, AEs were reported with similar frequency in all treatment groups, and observed differences were typically small, sometimes favouring one treatment group and other times favouring another without a clear discernible pattern .

© Gedeon Richter UK Ltd (2023). All rights reserved

A more detailed summary of AEs reported for more than 5% in any group is provided in Table 32. The incidence of AEs with Relugolix CT was similar to that observed with placebo in both SPIRIT 1 and SPIRIT 2 (37).

Table 32: Adverse events reported for >5% in any group in SPIRIT 1 or SPIRIT 2

Characteristics	SPIRIT 1						SPIRIT 2					
	Relugolix CT		Placebo		Relugolix + delayed CT		Relugolix CT		Placebo		Relugolix + delayed CT	
Headache	57	27%	46	22%	67	32%	81	39%	64	31%	79	38%
Hot flush	22	10%	21	10%	71	34%	28	14%	7	3%	72	35%
Nasopharyngitis	13	6%	12	6%	10	5%	29	14%	17	8%	14	7%
Toothache	5	2%	3	1%	3	1%	18	9%	7	3%	7	3%
Nausea	13	6%	11	5%	9	4%	12	6%	6	3%	9	4%
Back pain	8	4%	5	2%	7	3%	12	6%	7	3%	12	6%
Arthralgia	4	2%	2	1%	9	4%	11	5%	7	3%	10	5%
Bone density decreased	5	2%	4	2%	8	4%	11	5%	5	2%	13	6%
Libido decreased	5	2%	1	<1%	7	3%	11	5%	4	2%	8	4%
Urinary tract infection	4	2%	6	3%	9	4%	11	5%	5	2%	10	5%
Vitamin D decreased	4	2%	15	7%	8	4%	1	1%	3	1%	0	-
Acne	2	1%	13	6%	1	<1%	7	3%	11	5%	7	3%

© Gedeon Richter UK Ltd (2023). All rights reserved

The most frequently reported adverse events (≥5% of patients in any treatment group) included headache and hot flush. Both of these events were numerically more common in one or both Relugolix CT groups compared with the placebo groups (62, 63).

Table 33 shows the percentage change from baseline to week 24 in lumbar spine BMD and total hip BMD.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 33: LS mean change in lumbar spine and total hip BMD from baseline (37)

	SPIRIT 1			SPIRIT 2		
	Relugolix CT	**Placebo**	**Relugolix + delayed CT**	**Relugolix CT**	**Placebo**	**Relugolix + delayed CT**
Lumbar spine (L1-L4)						
Week 12						
n	177	172	181	172	166	166
LS mean % change from baseline	-0.52 (0.239)	0.29 (0.242)	-1.69 (0.243)	-0.47 (0.217)	-0.14 (0.219)	-1.87 (0.224)
95% CI	(-0.99, -0.05)	(-0.18, 0.77)	(-2.16, -1.21)	(-0.90, -0.05)	(-0.90, -0.05)	(-2.31, -1.43)
Week 24						
n	164	161	174	168	156	163
LS mean % change from baseline	-0.70 (0.255)	0.21 (0.256)	-1.99 (0.256)	-0.78 (0.233)	0.02 (0.237)	-1.92 (0.239)
95% CI	(-1.20, -0.20)	(-0.30, 0.71)	(-2.49, -1.48)	(-1.23, -0.32)	(-0.45, 0.48)	(-2.39, -1.45)
Total hip						
Week 12						
n	171	172	181	173	166	162
LS mean % change from baseline	0.01 (0.209)	0.25 (0.211)	-0.65 (0.211)	-0.31 (0.185)	-0.02 (0.187)	-0.81 (0.193)
95% CI	(-0.40, 0.42)	(-0.16, 0.67)	(-1.06, -0.23)	(-0.67, 0.06)	(-0.38, 0.35)	(-1.19, -0.43)
Week 24						
n	164	161	173	169	157	163
LS mean % change from baseline	-0.11 (0.216)	0.27 (0.217)	-0.74 (0.217)	-0.56 (0.196)	-0.19 (0.199)	-0.89 (0.202)
95% CI	(-0.53, 0.31)	(-0.16, 0.70)	(-1.17, -0.32)	(-0.95, -0.18)	(-0.58, 0.20)	(-1.29, -0.50)

BMD, bone mineral density; CI, confidence interval; CT, combination therapy; LS, least squares

© Gedeon Richter UK Ltd (2023). All rights reserved

Noticeable differences exist between the treatment groups. BMD was preserved in the Relugolix CT group and there was no difference compared to the placebo group through 24 weeks of treatment. As expected, relugolix + delayed CT was associated with a decline in bone mass that stabilised with transition to Relugolix CT. Furthermore, the percent change in BMD at the lumbar spine and hip at Week 12 in the Relugolix CT group was lower than in the relugolix + delayed CT group, reflecting the benefit of combined treatment with Relugolix CT to minimise bone loss (62, 63).

Deaths and serious adverse events (SAE)

No deaths were reported during the studies.

In SPIRIT 1, SAEs were reported for 3 patients (1%), 5 patients (2%), and 3 patients (1%) in the Relugolix CT group, placebo group and relugolix + delayed CT group, respectively (Table 34). In SPIRIT 2, SAEs were reported for 9 patients (4%), 4 patients (2%), and 6 patients (3%), in the Relugolix CT group, placebo group and relugolix + delayed CT group, respectively (Table 34).

Abdominal pain, a symptom of endometriosis, was reported as a SAE after discontinuation of Relugolix CT or relugolix + delayed CT for 3 patients in SPIRIT 2, possibly reflecting symptom exacerbation following loss of efficacy with treatment discontinuation (63). There were nine reports of suicidal ideation across both studies: (placebo run-in n = 2, Relugolix CT n=3, placebo n=2 and relugolix + delayed CT n = 3). All reports were in women with a history of psychiatric disorders. All patients who had suicidal ideation discontinued from the studies (37).

In summary, the overall incidence of SAEs was low and similar across treatment groups in SPIRIT 1 & 2, suggesting that SAEs were not linked to a treatment effect of Relugolix CT.

.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 34: Summary of serious adverse events by system organ class and preferred term (safety population) in SPIRIT 1 and SPIRIT 2 (37, 62, 63)

Preferred Term	SPIRIT 1						SPIRIT 2					
	Relugolix CT		Placebo		Relugolix + delayed CT		Relugolix CT		Placebo		Relugolix + delayed CT	
No. Of patients with at least one serious AE n (%)	3	1%	5	2%	3	1%	9	4%	4	2%	6	3%
Gastrointestinal disorders	0		2	0.9%	0		3	1.5%	0		0	
Abdominal adhesions	0		1	0.5%	0		0		0		0	
Peptic ulcer	0		1	0.5%	0		0		0		0	
Abdominal pain	0		0		0		2	1.0%	0		0	
Abdominal pain lower	0		0		0		1	0.5%	0		0	
Intestinal obstruction	0		0		0		1	0.5%	0		0	
Hepatobiliary disorders	0		0		1	0.5%	2	1.0%	0		0	
Cholecystitis	0		0		1	0.5%	1	0.5%	0		0	
Cholelithiasis	0		0		0		1	0.5%	0		0	
Infections and infestations	1	0.5%	0		0		0		0		0	
Pneumonia	1	0.5%	0		0		0		0		0	
Injury, poisoning and procedural complications	1	0.5%	2	0.9%	0		0		0		2	1.0%
Cartilage injury	0	0	1	0.5%	0		0		0		0	
Hand fracture	0	0	1	0.5%	0		0		0		0	
Ligament rupture	1	0.5%	0		0		0		0		0	
Neck injury	0		1	0.5%	0		0		0		0	
Clavicle fracture	0		0		0		0		0		1	0.5%
Ulnar nerve injury	0		0		0		0		0		1	0.5%
Nervous system disorders	0		0		1	0.5%	0		1	0.5%	0	
Hemiparesis	0		0		0		0		1	0.5%	0	
Migraine	0		0		1	0.5%	0		0		0	
Psychiatric disorders	0		1	0.5%	1	0.5%	2	1.0%	3	1.5%	2	1.0%
Suicidal ideation	0		1	0.5%	1	0.5%	2	1.0%	1	0.5%	2	1.0%
Anxiety	0		0		0		0		1	0.5%	0	

© Gedeon Richter UK Ltd (2023). All rights reserved

Depression	0		0		0		0		1	0.5%	0	
Generalised anxiety disorder	0		0		0		0		1	0.5%	0	
Reproductive system and breast disorders	2	0.9%	1	0.5%	0		2	1.0%	0		1	0.5%
Endometriosis	1	0.5%	0		0		0		0		0	
Ovarian cyst	1	0.5%	1	0.5%	0		0		0		1	0.5%
Pelvic pain	1	0.5%	0		0		1	0.5%	0		0	
Uterine haemorrhage	0		0		0		1	0.5%	0		0	
Blood and lymphatic system disorders	0		0		0		1	0.5%	0		0	
Anaemia	0		0		0		1	0.5%	0		0	
Cardiac disorders	0		0		0		0		0		1	0.5%
Palpitations	0		0		0		0		0		1	0.5%
Endocrine disorders	0		0		0		1	0.5%	0		0	
Goitre	0		0		0		1	0.5%	0		0	
Neoplasms benign, malignant and unspecified (incl cysts and polyps)	0		0		0		1	0.5%	0		0	
Non-small cell lung cancer stage IIIA	0		0		0		1	0.5%	0		0	
Pregnancy, puerperium and perinatal conditions	0		0		0		1	0.5%	0		0	
Abortion missed	0		0		0		1	0.5%	0		0	
Renal and urinary disorders	0		0		0		1	0.5%	0		0	
Urinary retention							1	0.5%	0		0	

© Gedeon Richter UK Ltd (2023). All rights reserved

SPIRIT OLE (64)

Safety was evaluated by monitoring adverse events, clinical laboratory data, 12-lead ECGs, vital signs and weight, physical examinations, menstrual bleeding patterns, pregnancy, overdose, endometrial biopsies, mammograms, and BMD.

Patients who were treated with placebo had shorter exposure to Relugolix CT, and their data are supportive. Data from patients who were treated with relugolix + delayed CT and transitioned to Relugolix CT therapy have been provided for completeness; however, the incidence of adverse events is confounded by the initial 12 weeks of relugolix + delayed CT.

While all patients in the OLE study received Relugolix CT, all data in this report are presented based on the randomised treatment received in one of the pivotal phase 3 studies (i.e., original treatment assignment). Due to differences in the duration of exposure to Relugolix CT treatment, no cross-comparisons across groups have been performed.

Table 35:Key safety endpoints of SPIRIT OLE

Objective	Endpoint
Adverse events	Incidence of AEs through Week 104
Change in BMD at Week 52 and Week 104	To determine the percent change from pivotal study baseline to Week 52 and Week 104 in BMD measured by DXA

Treatment-emergent adverse events (64)

A cumulative summary of adverse events reported for patients enrolled in this extension study is presented in Table 36. For each treatment group, adverse events are summarised in two columns:

- One for adverse events reported since randomisation in one of the parent studies (“Cumulative”)
- One for adverse events reported since initiation of open-label study treatment in this open-label extension study (“Extension”)

© Gedeon Richter UK Ltd (2023). All rights reserved

The frequency of subjects who reported TEAEs in the placebo group was similar to the Relugolix CT treatment groups. Regarding cumulative adverse events in the SPIRIT parent study and SPIRIT OLE, 249 (90.5%) subjects who were in the placebo group in the parent study experienced at least one TEAE. This applies to 258 (93.1%) subjects who were treated with Relugolix CT throughout the SPIRIT parent study and SPIRIT OLE and to 224 (90.7%) subjects who were treated with relugolix + delayed CT at the parent study.

Table 36: Overall Summary of Adverse Events (Extension Safety Population)

Characteristics	SPIRIT OLE					
	Relugolix CT		Placebo → Relugolix CT		Relugolix + delayed CT→ Relugolix CT	
	Cumulative	Extension	Cumulative	Extension	Cumulative	Extension
Any	258 (93.1%)	204 (73.6%)	249 (90.5%)	215 (78.2%)	224 (90.7%)	177 (71.7%)
Leading to discontinuation	19 (6.9%)	15 (5.4%)	23 (8.4%)	22 (8.0%)	23 (9.3%)	17 (6.9%)
Leading to drug interruption	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Related to study drug	172 (62.1%)	94 (33.9%)	177 (64.4%)	135 (49.1%)	175 (70.9%)	93 (37.7%)
Grade 3 or above	30 (10.8%)	15 (5.4%)	42 (15.3%)	30 (10.9%)	34 (13.8%)	23 (9.3%)
Grade 3 or above related to study drug	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Serious	11 (4.0%)	7 (2.5%)	20 (7.3%)	18 (6.5%)	20 (8.1%)	19 (7.7%)
Serious and related to study drug	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Serious, leading to discontinuation	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Serious, leading to drug interruption	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Fatal outcome	0	0	0	0	0	0

© Gedeon Richter UK Ltd (2023). All rights reserved

In the Relugolix CT group:

- Cumulatively over the 104-week treatment, at least 1 adverse event was reported for 258 patients (93.1%). During participation in this 80-week OLE study, at least 1 adverse event was reported for 204 patients (73.6%)

- The overall incidence of adverse events in the cumulative experience was [redacted] during the periods Day 1 to Week 24 [24 weeks]; [redacted] during >Week 24 to Week 52 [28 weeks]; and [redacted] during >Week 52 to Week 78 [26 weeks], and [redacted] during >Week 78 to Week 104 [26 weeks]

- Relatively few patients were reported to have grade 3 or higher events (5.4% during the OLE study, 10.8% cumulatively) and few patients were reported to have serious adverse events (2.5% during the OLE study, 4.0% cumulatively)

- Serious events assessed by the investigator as related to study drug (0 during the OLE study and [redacted] cumulatively), or events leading to treatment discontinuation (5.4% during the OLE study, [redacted] cumulatively), as shown in Table 36 above.

In the placebo group:

- At least 1 adverse event was reported for 249 patients (90.5%) over the 104-week treatment encompassing the pivotal and OLE studies

- During participation in this 80-week OLE study, at least 1 adverse event was reported for 215 patients (78.2%). The overall incidence of adverse events in the cumulative experience during the periods Day 1 to Week 24 [24 weeks]; >Week 24 to Week 52 [28 weeks]; and >Week 52 to Week 78 [26 weeks], and >Week 78 to Week 104 [26 weeks] was [redacted] and [redacted]

- Relatively few patients were reported to have grade 3 or higher events (10.9% during the OLE and 15.3% cumulatively) and few patients were reported to have serious adverse events (6.5% during the OLE and 7.3% cumulatively), serious events assessed by the investigator as related to study drug remains [redacted] during the OLE study and [redacted] cumulatively), or events leading to

© Gedeon Richter UK Ltd (2023). All rights reserved

treatment discontinuation ([redacted] during the OLE study, [redacted] cumulatively), as shown in Table 36.

In the relugolix + delayed CT group:

- At least 1 adverse event was reported for 224 patients (90.7%) over the 104-week treatment encompassing the pivotal and OLE studies
- During participation in this 80-week OLE study, at least 1 adverse event was reported for 177 patients (71.7%). The overall incidence of adverse events in the cumulative experience during the periods Day 1 to Week 24 [24 weeks]; >Week 24 to Week 52 [28 weeks]; and >Week 52 to Week 78 [26 weeks], and >Week 78 to Week 104 [26 weeks] was [redacted] and [redacted]
- Relatively few patients were reported to have grade 3 or higher events with increased exposure to Relugolix CT (9.3% during the OLE study, 13.8% cumulatively) and few patients were reported to have serious adverse events (7.7% during the OLE study, 8.1% cumulatively)
- Serious events assessed by the investigator as related to study drug remains similar [redacted] during the OLE study and [redacted] cumulatively), or events leading to treatment discontinuation ([redacted] during the OLE study, [redacted] cumulatively), as shown in Table 36.

A cumulative summary of treatment-emergent adverse events reported since the time of first dose of study drug in one of the parent studies, by preferred term reported in at least 5% of patients in any parent treatment group, is presented in Table 37. The incidence of adverse events with previous enrolment in any of the relugolix groups in the parent study was similar to that observed with previous enrolment in the placebo group.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 37: Cumulative summary of adverse events reported for at least 5% of patients in any treatment group by preferred term (extension safety population) (64)

Preferred Term	SPIRIT OLE					
	Relugolix CT		Placebo→ Relugolix CT		Relugolix + delayed CT→ Relugolix CT	
No. of patients with at least one AE n (%)	258	93.1%	249	90.5%	224	90.7%
Headache	146	52.7%	121	44.0%	119	48.2%
Nasopharyngitis	63	22.7%	46	16.7%	30	12.1%
Hot flush	41	14.8%	40	14.5%	106	42.9%
Urinary tract infection	31	11.2%	24	8.7%	25	10.1%
Vulvovaginal mycotic infection	31	11.2%	16	5.8%	12	4.9%
Toothache	30	10.8%	14	5.1%	12	4.9%
Back pain	28	10.1%	24	8.7%	18	7.3%
Nausea	28	10.1%	22	8.0%	13	5.3%
Vaginal infection	[redacted]					
Bone density increased	[redacted]					
Vulvovaginal dryness	[redacted]					
Influenza	[redacted]					
Arthralgia	[redacted]					
Libido decreased	[redacted]					
Depressed mood	[redacted]					
Upper respiratory tract infection	[redacted]					
Alopecia	[redacted]					
Bronchitis	[redacted]					
Fatigue	[redacted]					
Metrorrhagia	[redacted]					
Mood swings	[redacted]					
Constipation	[redacted]					
Corona virus infection	[redacted]					
Sinusitis	[redacted]					
Diarrhoea	[redacted]					
Acne	[redacted]					
Vitamin D decreased	[redacted]					
Cystitis	[redacted]					
Menorrhagia	[redacted]					
Weight increased	[redacted]					

The most frequently reported adverse events (>5% incidence cumulatively in the OLE study) over the cumulative 104-week treatment period (and approximate 30-day safety follow-up period) of the Relugolix CT group are summarised in Table 37. For most of these preferred terms, the first onset of the adverse event was reported during the pivotal studies and generally there was no evidence of an incremental time-dependent increase in events (i.e., more than what would be expected given the longer follow-up), including for events that may be related to a hypoestrogenic state or treatment with the add-back therapy.

© Gedeon Richter UK Ltd (2023). All rights reserved

Cumulatively in the Relugolix CT group, during the pivotal and OLE studies, hot flush was reported for 41 patients (14.8%) (Table 37); bone density decreased was reported for [redacted] patients [redacted]; vulvovaginal dryness was reported for [redacted] patients [redacted] arthralgia was reported for [redacted] patients [redacted] libido decreased was reported for [redacted] patients [redacted] depressed mood was reported for [redacted] patients [redacted] alopecia was reported for [redacted] patients [redacted] metrorrhagia was reported for [redacted] patients [redacted] and mood swings were reported for [redacted] patients [redacted]

In the placebo group, the distribution of events between cumulative experience and OLE study was generally proportional. That indicates no evidence of an increase in common adverse events following transition from placebo to Relugolix CT, including for hot flush (14.5% cumulatively vs. 8.0% in the OLE study), vulvovaginal dryness ([redacted] vs. [redacted] respectively), libido decreased ([redacted] vs. [redacted] respectively), depressed mood [redacted] vs[redacted] respectively), mood swings [redacted] vs[redacted] respectively), alopecia ([redacted] vs. [redacted] respectively), metrorrhagia ([redacted] vs. [redacted] respectively), and menorrhagia ([redacted] vs. [redacted] respectively)

In the relugolix + delayed CT group, the higher incidence of hot flush (42.9%), relative to the Relugolix CT group (14.8%) during the cumulative experience, is consistent with the hypoestrogenic state associated with relugolix + delayed CT during the first 12 weeks of treatment. The only common event (>5% incidence cumulatively) that increased in incidence disproportionately during the OLE study, compared with the cumulative experience was coronavirus infection for which all events ([redacted] patients [[redacted] were reported only during the OLE study; this was consistent with the timing of the global COVID-19 pandemic relative to the timing of the pivotal and OLE studies.

Deaths and Serious Adverse Events (SAE)

A summary of the SAEs reported during SPIRIT OLE is reported in Table 38. No deaths were reported during the study.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 38: Summary of serious adverse events by system organ class and preferred term (safety population) (64, 65)

Preferred Term	Relugolix CT		Placebo→ Relugolix CT		Relugolix + delayed CT→ Relugolix CT	
	Cumulative	Extension	Cumulative	Extension	Cumulative	Extension
No. of patients with at least one serious AE n (%)	11 (4.0%)	7 (2.5%)	20 (7.3%)	18 (6.5%)	20 (8.1%)	19 (7.7%)
Endocrine disorders	[redacted]	0	[redacted]	2 (0.7%)	[redacted]	0
Goitre	[redacted]	0	[redacted]	2 (0.7%)	[redacted]	0
Eye disorders	[redacted]	0	[redacted]	1 (0.4%)	[redacted]	0
Eye pain	[redacted]	0	[redacted]	1 (0.4%)	[redacted]	0
Vision blurred	[redacted]	0	[redacted]	1 (0.4%)	[redacted]	0
Gastrointestinal disorders	[redacted]	0	[redacted]	1 (0.4%)	[redacted]	1 (0.4%)
Abdominal pain lower	[redacted]	0	[redacted]	0	[redacted]	1 (0.4%)
Vomiting	[redacted]	0	[redacted]	1 (0.4%)	[redacted]	0
General disorders and administration site conditions	[redacted]	0	[redacted]	1 (0.4%)	[redacted]	1 (0.4%)
Fatigue	[redacted]	0	[redacted]	0	[redacted]	1 (0.4%)
Non-cardiac chest pain	[redacted]	0	[redacted]	1 (0.4%)	[redacted]	0
Hepatobiliary disorders	[redacted]	0	[redacted]	1 (0.4%)	[redacted]	3 (1.2%)
Cholecystitis	[redacted]	0	[redacted]	0	[redacted]	1 (0.4%)
Cholelithiasis	[redacted]	0	[redacted]	1 (0.4%)	[redacted]	2 (0.8%)
Immune system disorders	[redacted]	0	[redacted]	0	[redacted]	1 (0.4%)
Anaphylactic reaction	[redacted]	0	[redacted]	0	[redacted]	1 (0.4%)
Infections and infestations	[redacted]	2 (0.7%)	[redacted]	5 (1.8%)	[redacted]	2 (0.8%)
Cellulitis	[redacted]	1 (0.4%)	[redacted]	0	[redacted]	0
Sinusitis	[redacted]	1 (0.4%)	[redacted]	0	[redacted]	0
Appendicitis	[redacted]	0	[redacted]	1 (0.4%)	[redacted]	0
Corona virus infection	[redacted]	0	[redacted]	1 (0.4%)	[redacted]	2 (0.8%)
Influenza	[redacted]	0	[redacted]	1 (0.4%)	[redacted]	0
Laryngitis	[redacted]	0	[redacted]	1 (0.4%)	[redacted]	0
Vestibular neuronitis	[redacted]	0	[redacted]	1 (0.4%)	[redacted]	0
Injury, poisoning and procedural complications	[redacted]	0	[redacted]	0	[redacted]	2 (0.8%)
Ligament rupture	[redacted]	0	[redacted]	0	[redacted]	0
Cartilage injury	[redacted]	0	[redacted]	0	[redacted]	0
Clavicle fracture	[redacted]	0	[redacted]	0	[redacted]	0

© Gedeon Richter UK Ltd (2023). All rights reserved

Preferred Term	Relugolix CT		Placebo→ Relugolix CT		Relugolix + delayed CT→ Relugolix CT	
	Cumulative	Extension	Cumulative	Extension	Cumulative	Extension
Fibula fracture		0		0		1 (0.4%)
Hand fracture		0		0		0
Neck injury		0		0		0
Tibia fracture		0		0		2 (0.8%)
Neoplasms benign, malignant and unspecified (incl cysts and polyps)		2 (0.7%)		2 (0.7%)		3 (1.2%)
Diffuse large B-cell lymphoma		1 (0.4%)		0		0
Papillary thyroid cancer		1 (0.4%)		0		0
Hepatic adenoma		0		0		1 (0.4%)
Ovarian adenoma		0		0		1 (0.4%)
Thyroid adenoma		0		1 (0.4%)		0
Uterine leiomyoma		0		1 (0.4%)		1 (0.4%)
Nervous system disorders		1 (0.4%)		0		0
Generalised tonic-clonic seizure		1 (0.4%)		0		0
Pregnancy, puerperium and perinatal conditions		1 (0.4%)		0		1 (0.4%)
Abortion missed		1 (0.4%)		0		0
Abortion spontaneous				0		1 (0.4%)
Psychiatric disorders		1 (0.4%)		1 (0.4%)		6 (2.4%)
Personality disorder		1 (0.4%)		0		0
Suicide threat		1 (0.4%)		0		0
Anxiety disorder		0		1 (0.4%)		0
Borderline personality disorder		0		0		1 (0.4%)
Depression		0		0		1 (0.4%)
Drug dependence		0		0		1 (0.4%)
Panic disorder		0		1 (0.4%)		0
Persistent depressive disorder		0		1 (0.4%)		0
Suicidal ideation		0		1 (0.4%)		3 (1.2%)
Suicide attempt		0		0		1 (0.4%)
Renal and urinary disorders		0		1 (0.4%)		0
Nephrolithiasis		0		1 (0.4%)		0
Reproductive system and breast disorders		1 (0.4%)		4 (1.5%)		2 (0.8%)
Ovarian cyst		1 (0.4%)		0		0

© Gedeon Richter UK Ltd (2023). All rights reserved

Preferred Term	Relugolix CT		Placebo→ Relugolix CT		Relugolix + delayed CT→ Relugolix CT	
	Cumulative	Extension	Cumulative	Extension	Cumulative	Extension
Broad ligament tear		0		1 (0.4%)		0
Endometrial hyperplasia		0		1 (0.4%)		0
Endometriosis		0		2 (0.7%)		1 (0.4%)
Metrorrhagia		0		1 (0.4%)		0
Pelvic pain		0		0		1 (0.4%)
Respiratory, thoracic and mediastinal disorders		0		1 (0.4%)		0
Pulmonary embolism		0		1 (0.4%)		0
Skin and subcutaneous tissue disorders		1 (0.4%)		0		0
Urticaria		1 (0.4%)		0		0
Vascular disorders		0		1 (0.4%)		0
Deep vein thrombosis		0		1 (0.4%)		0

© Gedeon Richter UK Ltd (2023). All rights reserved

In the Relugolix CT group:

- Relative to the duration of follow-up, the proportion of patients with serious adverse events did not increase disproportionately during continued treatment with Relugolix CT during the OLE study relative to the pivotal studies
- The cumulative percentage of patients with serious adverse events in the pivotal studies and OLE study (up to 104 weeks of treatment) and OLE study (80 weeks of treatment) were 4.0% and 2.5%, respectively.

In the placebo group:

- Patients were treated with placebo for 24 weeks and then with Relugolix CT for up to 80 weeks
- The cumulative percentage of patients with serious adverse events was 7.3%; 6.5% of patients had serious adverse events in the OLE study.

In relugolix + delayed CT group:

- Patients were treated with delayed Relugolix CT for 12 weeks and then with Relugolix CT for up to 92 weeks
- The cumulative percentage of patients with serious adverse events was 8.1%. In the OLE study, 7.7% of patients had serious adverse events.

Serious adverse events with first onset during the OLE study were reported at low frequency in all treatment groups (2.5% in the Relugolix CT group, 7.7% in the relugolix + delayed CT group, and 6.5% in the placebo group) with no overall pattern as to the types of events reported.

B.2.11 Ongoing studies

None.

© Gedeon Richter UK Ltd (2023). All rights reserved

B.2.12 Interpretation of clinical effectiveness and safety evidence

Clinical effectiveness

Summary: SPIRIT 1, SPIRIT 2 and SPIRIT OLE

The efficacy of Relugolix CT has been demonstrated through a series of two multicentre Phase 3 trials (SPIRIT 1 and SPIRIT 2) and one Phase 3 open-label extension study of SPIRIT 1 and SPIRIT 2 (SPIRIT OLE).

In the Relugolix CT group, 158 (75%) of 212 patients in SPIRIT 1 and 155 (75%) of 206 patients in SPIRIT 2 met the dysmenorrhoea responder definition compared with 57 (27%) patients receiving placebo in SPIRIT 1 and 62 (30%) patients in SPIRIT 2. The difference in dysmenorrhoea responder rates between Relugolix CT and placebo was 47·6% (95% CI 39·3–56·0) in SPIRIT 1 and 44·9% (36·2–53·5) in SPIRIT 2, both p<0·0001 (37).

For non-menstrual pelvic pain, 124 (59%) patients in SPIRIT 1 and 136 (66%) patients in SPIRIT 2 met the responder definition in the Relugolix CT group compared with 84 (40%) patients receiving placebo in SPIRIT 1 and 87 (43%) in SPIRIT 2. The difference in non-menstrual pelvic pain responders between placebo and Relugolix CT was 18·9% (95% CI 9·5–28·2) in SPIRIT 1 and 23·4% (14·0–32·8) in SPIRIT 2, both p<0·0001. The response rates in the relugolix + delayed CT group were similar in both studies; for dysmenorrhoea, 151 (72%) of 211 patients in SPIRIT 1 and 150 (73%) of 206 patients in SPIRIT 2 had a response, and for non-menstrual pelvic pain 122 (58%) patients in SPIRIT 1 and 109 (53%) patients in SPIRIT 2 had a response. The results of five sensitivity analyses for both co primary endpoints were consistent with the primary analysis for each endpoint (37).

For women originally randomised to Relugolix CT in SPIRIT 1 and SPIRIT 2, the reduction in dysmenorrhoea and NMPP NRS scores were maintained for up to 104 weeks, and for those originally randomised to placebo, a reduction in their endometriosis-associated pain was observed after receiving Relugolix CT during an 80-week, open-label, single-arm extension study (SPIRIT OLE) (64).

© Gedeon Richter UK Ltd (2023). All rights reserved

Safety

The overall incidence of adverse events, both serious and non-serious, was similar among treatment groups. The most common adverse events were headache and nasopharyngitis. Hot flushes were reported more frequently in the relugolix + delayed CT group than in the Relugolix CT or placebo groups, and mostly occurred during the first 12 weeks of treatment. There were nine reports of suicidal ideation across both studies including the run in period, all in women with a self-reported psychiatric history (placebo run in, placebo, Relugolix CT, and relugolix + delayed CT); all patients who had suicidal ideation discontinued study participation (37).

Least squares mean percentage changes from baseline to week 12 and 24 in bone mineral density at the lumbar spine and total hip were less than 1% in patients treated with Relugolix CT in both studies. In the relugolix + delayed CT groups, bone mineral density at the lumbar spine and total hip substantially declined at week 12 with relugolix monotherapy, which stabilised with transition to Relugolix CT (37).

No clinically important differences were evident in vital signs including blood pressure or laboratory parameters including liver function tests and lipids. Most women treated with Relugolix CT or relugolix + delayed CT reported no bleeding or infrequent bleeding compared with the placebo group, in which most women reported normal bleeding or irregular or infrequent bleeding (37).

In patients who did not continue into the long-term study extension, menses resumed after cessation of Relugolix CT or relugolix + delayed CT, other than in those patients with a known reason for non recovery (e.g. pregnancy, medications, or surgery). The median time of menses return was 31 days for both the relugolix combination groups (IQR 21–36) and relugolix + delayed CT groups (24–36). 90 (94%) of the 96 patients with menstruation status follow up from the Relugolix CT group and 120 (91%) of the 132 patients with menstruation status follow up from the relugolix + delayed CT resumed menses within 2 months of stopping treatment. There were 14 pregnancies during the study. period (placebo- 8, Relugolix CT- 4, and relugolix + delayed CT- 20). Of the six pregnancies in the relugolix groups, three occurred during the first month of treatment, and 2 patients who were pregnant had poor compliance by eDiary entry. No congenital anomalies were reported in

© Gedeon Richter UK Ltd (2023). All rights reserved

pregnancies in which the outcome is known. No cases of endometrial hyperplasia or endometrial cancer were reported (37).

Data from the SPIRIT OLE shows that Relugolix CT was generally well-tolerated with a mean decrease < 1 % in bone density that did not progress during long-term treatment. No endometrial safety concerns were identified. Resumption of menses was prompt following treatment discontinuation in the majority of women, even after 104 weeks of continuous treatment with Relugolix CT (64).

Bone mineral density

Although GnRH receptor analogues are approved for the treatment of endometriosis-associated pain, they have either suboptimal efficacy at low doses, require injections, or are associated with undesirable hypoestrogenic adverse effects of hot flushes and bone density loss at high doses. In a phase 2 dose-ranging study in women with endometriosis- associated pain, 24-week treatment with relugolix 40 mg monotherapy was associated with significant reduction in pelvic pain versus placebo, with efficacy similar to leuprolide. However, dose-dependent decreases in bone mineral density and increases in vasomotor symptoms limited the duration of use. Relugolix CT (consisting of 40 relugolix, 1 mg estradiol, and 0·5 mg norethisterone acetate) was developed as a once-daily treatment to achieve efficacy and minimise vasomotor symptoms and bone mineral density loss by maintaining oestradiol concentrations within a therapeutic range consistent with those in the early follicular phase of the menstrual cycle (37).

For the purposes of the economic model, the risk of major osteoporotic fracture was estimated with the percentage BMD changes at different anatomical locations for Relugolix CT and BSC based on the weighted average of values in the SPIRIT 1 and 2 trials, and sources in the literature (62, 63, 69-71). Further information on this can be found in Section B.3.3.

It is worth highlighting that it is unclear if the previously mentioned bone mineral density loss associated with longer-term GnRH agonist therapy is recoverable after cessation of therapy. Whereas some studies in women receiving GnRH agonist suggested that bone loss is recovered when treatment is discontinued (72, 73). Others reported a sustained decrease without recovery (73-76). A further study by

© Gedeon Richter UK Ltd (2023). All rights reserved

Pierce et al. (2000) (77) showed that in a population of women with an average age of approximately 40 years that even 6 years after completion of a course of agonist treatment the bone had not fully recovered, and that overall prolonged use may increase the future risk of osteoporosis.

This lack of bone recovery, particularly in this age group of the population can have considerable effects on their long-term risk of trauma fracture and osteoporosis. It is estimated that on average the rate of normal premenopausal bone loss is between 0.7-1.3% at the lumbar spine (78, 79). It is estimated that having a BMD that is 2.5 standard deviations below the mean of the adult reference population increases the risk of osteoporosis by approximately 20% (80). Therefore, if the normal level of bone loss is further increased by the use of products such as GnRH agonists, which even up to 6 years post treatment is not fully recoverable, then this group of the population will have a substantial potential for increased risk.

Comparatively, SPIRIT trial data shows that even up to 104 weeks of continuous treatment Relugolix CT was associated with a mean bone loss of less than 1%, with those transitioning from monotherapy to CT trending towards recovery (64).Thus it would appear that Relugolix CT has the potential benefit to preserve BMD even when used without interruption for extended periods of time.

Strengths and limitations of the Relugolix CT clinical evidence base

Strengths

The clinical evidence base described in this submission is derived principally from the SPIRIT studies: SPIRIT 1, SPIRIT 2 and SPIRIT OLE. Data from these studies capture evidence on dysmenorrhoea, non-menstrual pelvic pain, pain, pelvic pain, opioid use, analgesic use, dyspareunia, and AEs.

The studies met the primary efficacy endpoint of demonstrating superiority in dysmenorrhoea and non-menstrual pelvic pain response compared to placebo. In the Relugolix CT group, 158 (75%) of 212 patients in SPIRIT 1 and 155 (75%) of 206 patients in SPIRIT 2 met the dysmenorrhoea responder definition compared with 57 (27%) patients receiving placebo in SPIRIT 1 and 62 (30%) patients in SPIRIT 2. The difference in dysmenorrhoea responder rates between Relugolix CT and placebo was 47·6% (95% CI 39·3–56·0) in SPIRIT 1 and 44·9% (36·2–53·5) in

© Gedeon Richter UK Ltd (2023). All rights reserved

SPIRIT 2, both p<0·0001 (37). For non-menstrual pelvic pain, 124 (59%) patients in SPIRIT 1 and 136 (66%) patients in SPIRIT 2 met the responder definition in the Relugolix CT group compared with 84 (40%) patients receiving placebo in SPIRIT 1 and 87 (43%) in SPIRIT 2. The difference in non-menstrual pelvic pain responders between placebo and Relugolix CT was 18·9% (95% CI 9·5–28·2) in SPIRIT 1 and 23·4% (14·0–32·8) in SPIRIT 2, both p<0·0001 (37).

In the SPIRIT 1 and SPIRIT 2 studies, the robustness of the efficacy analyses were supported by 5 sensitivity analyses and subgroup analyses, the results of which were consistent with the primary analysis for each endpoint in nearly all cases. in SPIRIT 1 for NMPP, the lower and higher BMI groups favoured Relugolix CT over placebo, however, the lack of a trend and the relatively small sample in the subgroup make this likely related to chance (37). For the SPIRIT OLE some subgroup analyses yielded subgroups < 30 patients within a treatment group and greater variability. Nevertheless, in the Relugolix CT group, all subgroups, for both primary endpoints (dysmenorrhoea and NMPP) showed consistent point estimates and confidence intervals for the subgroups, overlapping with those of the overall population (64).

Limitations

The clinical evidence from the SPIRIT trials do not provide a head-to-head direct comparison between Relugolix CT and GnRH agonists or standard of care (e.g. first line oral contraceptives). Despite this, the SPIRIT 1 and SPIRIT 2 trials provide the pivotal RCT efficacy and safety data for Relugolix CT and are the most appropriate evidence base. An ITC has been conducted to provide evidence that was not captured via RCTs.

Many subjects in the SPIRIT 1 and SPIRIT 2 trials did not meet the minimum pelvic pain threshold to participate due to strict entry criteria, which could limit generalisability. Most women enrolled were white, potentially reflecting under-recognition or under diagnosis of endometriosis or suboptimal clinical trial engagement among other races and ethnicities. Treatment duration was also only 6 months, impairing long term efficacy data collection (37). The SPIRIT OLE trial does provide longer-term data beyond the SPIRIT 1 and SPIRIT 2 trials.

© Gedeon Richter UK Ltd (2023). All rights reserved

Limitations of the ITC are described in Section B.2.9.

Validity of the study results (SPIRIT 1 and SPIRIT 2)

The eligibility criteria for this study were selected to ensure that the study population was representative of the population of women with symptomatic endometriosis who are likely to be treated in clinical practice. All patients were confirmed to have endometriosis either by direct visualisation or surgical confirmation (37).

The robustness of the primary efficacy analysis results was supported by sensitivity and subgroup analyses, the results of which confirmed the results of the primary endpoints, demonstrating a significant improvement in both NMPP and dysmenorrhoea (62, 63).

Validity of the study results (SPIRIT OLE)

The baseline characteristics and demographics of the study population (and each of the treatment groups) are consistent with the populations analysed in the parent studies and are representative of patients who suffer with symptoms associated with endometriosis who would seek treatment in the community setting, and who have significant disease burden (64).

Despite the consistency with the parent studies in baseline characteristics and demographics, it must be acknowledged that there could be some selection bias among the patients who enrolled in this open-label extension study; however, the risk of this potential bias to meaningfully affect the study conclusions is considered small. Reasons for early termination in the parent study, patients' perceptions regarding parent study treatment assignment and treatment response, and patient motivation to continue or initiate open-label treatment could all play a role in decision making to continue into the open-label extension study (64).

The robustness of the primary efficacy analysis result was supported by subgroup analyses. Some subgroup analyses (e.g. 5 categories of BMI) with many categories yielded subgroups <30 patients within a treatment group and greater variability. Nevertheless, in the Relugolix CT group, all subgroups, for both primary endpoints (dysmenorrhoea and NMPP) showed consistent point estimates and confidence intervals for the subgroups, overlapping with those of the overall population (64).

© Gedeon Richter UK Ltd (2023). All rights reserved

The increased incidence of adverse events observed in the placebo group may have been related to ascertainment bias associated with the open-label nature of the extension study. Investigators and patients were aware that all patients were receiving Relugolix CT during this study and may have been more inclined to report adverse events, particularly when those potentially associated with hormonal changes were observed (64).

Although there is an inherent selection bias driven by the need to complete one of the parent studies to be eligible to enter this extension study, the fact that most patients completed the studies and most of those who completed entered the extension, makes this potential bias less likely to affect significantly the interpretation of the results (64).

Generalisability of SPIRIT trials to women with endometriosis in England

A clinical expert currently practicing in England who was consulted during the development of this submission stated that overall, the baseline demographics of patients in SPIRIT 1 and 2 were representative of the patients seen in clinical practice (81). However, the expert also noted that in some regions of the country, there may be a larger proportion of black patients than was included in the trials.

Most patients enrolled in SPIRIT 1 and 2 were white (>90%), potentially reflecting under recognition or under diagnosis of endometriosis, or suboptimal clinical trial engagement among other races and ethnicities. This demographic makeup is consistent with the generally described epidemiology of endometriosis although recent studies suggest there may be an ascertainment bias due to differences in the odds of endometriosis diagnosis by race and ethnicity (37, 57, 62, 63).

Relugolix CT in general clinical practice

GnRH agonists are licensed for endometriosis use up to a maximum of 6 months and use beyond this point is considered off license. However, many women with endometriosis wish to avoid surgery, either due to fear of infertility, busy lifestyle or cultural/religious reasons, and surgery other than hysterectomy is often not curative (45-47). Feedback sought from HCPs in the development of this submission informs us that pre-surgical GnRH agonist use is typically within the licensed 6-month

© Gedeon Richter UK Ltd (2023). All rights reserved

timeframe, however, due to long waiting lists for surgery and in the absence of other medical therapy options, long-term use beyond 6 months exists (82).

It is unclear which GnRH agonists are more commonly used in England as they are also licensed for other conditions (e.g. prostate cancer) therefore the volumes prescribed in prescription cost analyses cannot be used to estimate shares in this indication. Feedback sought from HCPs suggests that leuprorelin (Prostap) is a common option due to local price tendering/discounts (82). However, the available GnRH agonist formulations are considered equivalent in terms of efficacy.

Life expectancy

Endometriosis is not associated with increased mortality. There are no data to suggest that endometriosis affects life expectancy and fatalities associated with endometriosis are typically related to surgical procedure risks rather than the condition itself.

Patient numbers

An estimated 1.5 million women are affected by endometriosis in the UK (9). Relugolix CT is positioned for a subset of these patients who have failed or are unsuitable for surgery or hormonal contraceptives/oral progestogens. These patients would thus be eligible for treatment with GnRH analogues, however, estimation of the symptomatic patient numbers in England receiving GnRH agonist treatment is more difficult to estimate and is not reported.

By applying incidence figures from Soriano et al (2017) (45) to the different age bands of women in England, Gedeon Richter estimates that there are 1,031 women in England aged 18-52 years who have failed first-line therapy and would be eligible for treatment with Relugolix CT. A full explanation of these calculations is given in the accompanying budget impact analysis submission.

End of life criteria

Gedeon Richter considers that this technology does not meet the end-of-life criteria.

© Gedeon Richter UK Ltd (2023). All rights reserved

B.3 Cost effectiveness

B.3.1 Published cost-effectiveness studies

A systematic literature review (SLR) was conducted to identify published economic models, available economic evidence including economic evaluations, costs, and resource use, as well as relevant utility data for patients with endometriosis-associated pain. A detailed description of the SLR is provided in Appendix G. The relevant studies identified from the SLR are summarised in Table 39.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 39: Summary list of published cost-effectiveness studies

Study	Year	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Grand	2019	A cost-utility analysis that compares oral contraceptives vs no hormonal therapy. It uses a Markov sate transition model structure with five health states. Results are discounted at 3.5%. The model used a time horizon of 1 month with a cycle length of 1 month. The analysis was conducted from the perspective of the NHS England.	Hypothetical cohort: 1000, Starting age of cohort: 32	No hormonal treatment: 9.88 Oral contraceptives: 10.31 Mean difference: 0.43	£1707 for no hormonal treatment £1113 for oral contraceptives £594 is the mean difference:	NR
Bohn	2020	A cost-utility analysis that compares Strategy 4: Proceeding directly to surgery without attempting medical management first.vs Strategy 1: NSAIDs followed by surgery if there was no improvement, Strategy 2: NSAIDs, then a short-acting reversible contraceptive or a long-acting reversible contraceptive (LARC) followed by surgery if no improvement, and Strategy 3: NSAIDs, then a shortacting reversible contraceptive or LARC, then a LARC or a GnRH agonist or antagonist, followed by surgery if no improvement. It uses Decision Tree model structure. The model used a time horizon of 3 years. The analysis was conducted from the societal perspective.	Hypothetical cohort: 10,018,400, 18-45 years	Strategy 4: 1.96 Strategy 1: 2.18 Strategy 2: 2.28 Strategy 3: 2.34	* Strategy 4: 3,980 USD Strategy 1: 2,328 USD Strategy 2: 1,831 USD Strategy 3: 2,842 USD	Strategy 4: 2027.34 Strategy 1: 1067.91 Strategy 2: 803.27 Strategy 3: 1216.66
Bohn	2021	A cost-utility analysis that compares Strategy 4: Proceeding directly to surgery without attempting medical management first.vs Strategy 1: NSAIDs	Hypothetical cohort: 4,817,894	Strategy 4: 9.7 million	** Strategy 4: 42.1 billion	Strategy 3: $1,352/QALY

© Gedeon Richter UK Ltd (2023). All rights reserved

		followed by surgery if there was no improvement, Strategy 2: NSAIDs, then a short-acting reversible contraceptive or a long-acting reversible contraceptive (LARC) followed by surgery if no improvement, and Strategy 3: NSAIDs, then a shortacting reversible contraceptive or LARC, then a LARC or a GnRH agonist or antagonist, followed by surgery if no improvement. It uses Decision Tree model structure. Results are discounted at 3%. The model used a time horizon of 3 years. The analysis was conducted from the payor perspective.	18-45 years	Strategy 1: 10.7 million Strategy 2: 11.2 million Strategy 3: 11.4 million	Strategy 1: 22.6 billion Strategy 2: 12.9 billion Strategy 3: 13.2 billion	

Abbreviations: GnRH= Gonadotropin-releasing Hormone, ICER= Incremental Cost-effectiveness Ratio, LARC= Long-acting Reversible Contraception, NR= Not Reported, NSAIDs= Non-Steroidal Anti-inflammatory Drugs, QALY= Quality-adjusted Life-year, USD= United States Dollar.
*: Results reported for an individual.
**: Results reported for the entire cohort

© Gedeon Richter UK Ltd (2023). All rights reserved

B.3.2 Economic analysis

No cost-effectiveness studies for Relugolix CT (Relugolix 40 mg in combination with oestradiol 1 mg and norethisterone acetate 0.5 mg) were identified in the SLR. A de novo cost-effectiveness model was thus constructed to evaluate the cost-effectiveness of Relugolix CT for treating endometriosis symptoms. Furthermore, there were no previous NICE technology appraisals identified for endometriosis treatment.

Patient population

The patient population considered in the analysis is adult pre-menopausal women (average age of 33.9 years) with moderate to severe endometriosis-related pain who have a history of previous medical or surgical treatment. The model patient population is informed by the characteristics of the population enrolled in the SPIRIT 1 & 2 clinical trial (37).

The patient population demographics applied in the base-case settings of the model are summarised in Table 40 below.

Table 40: Patient population demographics

Baseline characteristics	Value	Source
Age (years)	33.88	SPIRIT 1 and 2
Body surface (m^2)	1.71	
Weight (kg)	70.4	
Total cholesterol (mg/dL)	182.36	
High-density lipoprotein (mg/dL)	29.72	
Systolic blood pressure (mmHg)	115.72	
Smoker (%)	17.1%	
Diabetes (%)	7.1%	

© Gedeon Richter UK Ltd (2023). All rights reserved

Model structure

The model simulates a cohort of premenopausal women with symptomatic endometriosis eligible for treatment with Relugolix CT over time. It includes 12 unique health states based on the response of medical therapies and surgical intervention. The model compares treatment with Relugolix CT to alternative treatments that are currently available from a healthcare perspective over a 16-year time horizon. The model time horizon spans until the age menopause, which is set to the UK average age of 50 years (83). The cost-effectiveness model takes the form of a semi-Markov cohort model. The choice of a Markov model to evaluate the cost effectiveness of Relugolix CT is largely in line with previous models that have evaluated various interventions in the treatment of endometriosis (83-91).

Markov chain models (MCMs) are well-suited to health conditions that have an ongoing “risk” with recurrent health events, such as endometriosis. In a MCM, the patient is assumed to be in a Markov state and events are modelled as transitions between states, which are assigned a health utility and costs. In this case, a MCM was used because of the relatively long time-horizon required to capture the effects of treatment on outcomes and costs and the need to model a relatively complex set of interrelated health states to accurately represent the treatment of symptomatic endometriosis. A MCM was also used here due to the short follow-up of the SPIRIT trials relative to the modelling time horizon, and the need to estimate long-term outcomes for subsequent health states (i.e., those that patients transit to following treatment initiation) from a variety of sources (92).

To reflect the clinically and economically important aspects of the treatment decisions of patients with endometriosis, the model was based on Markov states corresponding to response to treatment.

The model structure for the MCM was developed based on reviews of the study designs for the Phase 3 SPIRIT trials of Relugolix CT (62, 63), prior economic models of treatments for symptomatic endometriosis identified from a targeted review of the literature, and clinical practice guidelines for endometriosis. (29, 83-91), The model structure was also validated with clinical experts during the global advisory board.

© Gedeon Richter UK Ltd (2023). All rights reserved

States of the model were defined on the following characteristics:

- Response status
- Subsequent medical therapy for endometriosis
- Type of surgery (i.e., conservative, hysterectomy)
- Post-surgical recurrence
- Vital status (i.e., alive, or dead)

The model, shown in Figure 38, consists of the following possible Markov states:

- Initial treatment (Relugolix CT, GnRH agonist, best supportive care (BSC))
    - Response
    - Partial response
    - Non-response
- BSC
    - Response to BSC
    - Non-response to BSC
- Waiting time before surgery
- Post-hysterectomy stable
- Post-hysterectomy recurrence
- Post-hysterectomy reoperation
- Post-conservative surgery response
- Post-conservative surgery recurrence
- Dead

© Gedeon Richter UK Ltd (2023). All rights reserved

The model cycle length is 3 months. All patients start in the health state “Initial treatment” (A). Treatment response may be evaluated after three months or six months. Six months (corresponding to two model cycles) is the timepoint for evaluation selected in the model base-case, aligned with the time of evaluation for the two co-primary endpoints in the SPIRIT trials which were evaluated after six months (24 weeks). At the 6-month time point, treatment response of all patients is evaluated. Those with complete response move into the “Complete Response” health state and continue treatment until the end of the model horizon if response is maintained. Patients who do not fulfil the complete response criteria move to “Non-response”. The model also includes a partial response health state however this is only active if response is assessed at 3 months and is thus not utilized in the base-case.

Only patients in the complete response state remain on active treatment. Patients with no response terminate treatment, and after one cycle spent in the “Non-response” health state, reflecting the time for their clinicians to assess which subsequent treatment they will receive, then switch treatment (C) to BSC or undergo surgery to manage endometriosis. BSC includes hormonal therapy with or without analgesics. Surgical options include conservative surgery (laparoscopy) or hysterectomy. Some patients may prefer or be advised to remove one or both ovaries as part of the hysterectomy (e.g., an oophorectomy). In the case of surgery as subsequent treatment, there is a waiting time of six months. During that waiting time, patients are assumed to receive BSC.

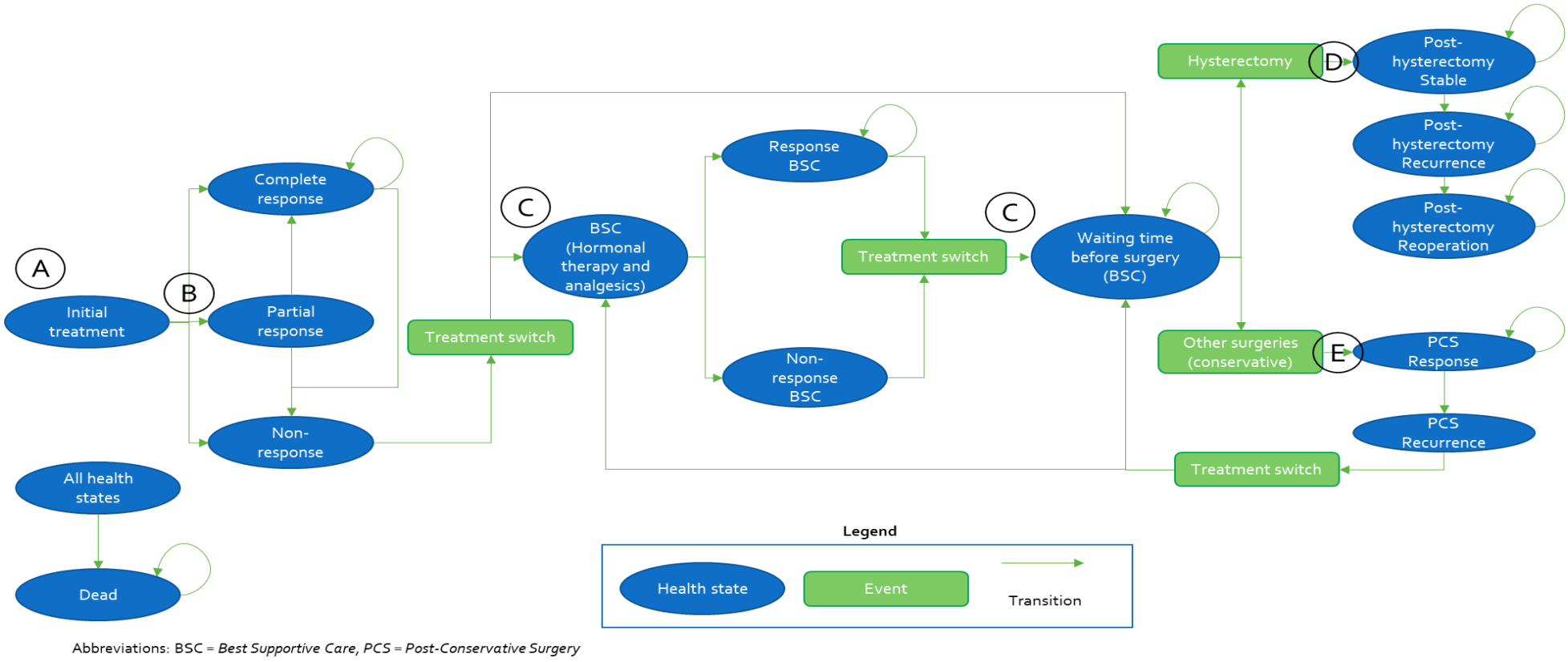
If patients undergo hysterectomy, they move into (D) the “Post-hysterectomy stable” state and remain there unless pain recurs (recurrence post hysterectomy), at which point they either opt for hormonal treatment (e.g., BSC) or an additional surgery (“Post-hysterectomy reoperation”) which in all cases is an oophorectomy. If patients undergo conservative surgery, they move to the “Post-conservative surgery response” state and remain there unless pain recurs. If pain recurs, patients may either undergo an additional surgery or use BSC. Patients may transition to the “Dead” state from all model states.

© Gedeon Richter UK Ltd (2023). All rights reserved

The analysis was conducted from the perspective of the NHS in line with current NICE guidelines. (70) The base-case analysis thus considers all costs incurred within the health care sector. Costs and outcomes are discounted at an annual rate of 3.5%, in line with the NICE reference case (70). The societal perspective was adopted in a scenario analysis.

© Gedeon Richter UK Ltd (2023). All rights reserved

Figure 38: Model structure



Abbreviations: BSC = *Best Supportive Care*, *PCS* = *Post-Conservative Surgery*

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 41: Features of the economic analysis

	Current evaluation	
Factor	**Chosen values**	**Justification**
Time horizon	16 Years	The starting age is 34 as this is the average age in the SPIRIT 1 & 2 trials. Therefore, the time horizon is set to 16 years which accounts for the starting age of 34 until 50 years old, the average age for a woman to reach menopause in the UK (93)
Treatment waning effect?	No	There is a lack of data from key clinical studies that would support a treatment waning effect for either Relugolix CT or any of the GnRH agonist comparators.
Source of utilities	EQ-5D-5L questionnaire from SPIRIT trials were first mapped to 3L using the NICE DSU age-sex based mapping (94, 95). Utilities were then derived using UK value set published by Dolan for EQ-5D-3L (96)	In line with the NICE reference case
Source of costs	British National Formulary (BNF) (97), NHS England national tariff 2022/23 (98), the literature, and KOL (key opinion leader) expert opinion.	Cost inputs were sourced from the British National Formulary (BNF), NHS England national tariff 2022/23 (99), and the literature. Where possible, costs were obtained from UK national resources to reflect the UK NHS perspective. Due to lack of published healthcare resource use (HRU) data specific to the population of interest, HRU frequencies for disease management and regular monitoring and tests or examinations was informed by KOL expert opinion.

Note: there has been no previous NICE technology appraisals for treating pain associated with endometriosis, thus the columns pertaining to previous evaluations have been omitted

© Gedeon Richter UK Ltd (2023). All rights reserved

Intervention technology and comparators

The modelled intervention is Relugolix CT (relugolix 40 mg in combination with oestradiol 1 mg and norethisterone acetate 0.5 mg). Relugolix CT is administered orally once daily. As Relugolix CT maintains estradiol and progestogen concentrations in a range that maintains BMD and endometrial health, it can be used continuously for as long as required. Thus, no maximum treatment duration is implemented in the model other than cessation at menopause.

The modelled comparator is GnRH agonist. The GnRH agonists that are licensed for the treatment of endometriosis in the UK are leuprorelin acetate, goserelin, triptorelin, nafarelin and buserelin. Leuprorelin acetate, goserelin and triptorelin are administered as subcutaneous injection, either in a short-acting (monthly) formulation or long-acting (3-monthly) formulation. Nafarelin and buserelin are administered as daily intranasal treatments. To inform this submission, Gedeon Richter sent out an email survey to UK-based healthcare professionals to provide information for a range of model parameters. Responses to the query pertaining to which GnRH agonist is used for the treatment of endometriosis associated pain were very heterogenous, with no consensus as to which is the most commonly-used GnRH agonist. However, none of the consulted KOLs prescribed intranasal GnRH agonists and they stated that all patients who are currently receiving GnRH agonist treatment opt for the subcutaneous formulations. No patients in the model are therefore allocated to receive nafarelin and buserelin.

Given that the efficacy of GnRH agonists is assumed equal in this population, and the choice of GnRH agonist impacts the cost of treatment only, we assumed a 50/50 split of patients in the GnRH agonist arm amongst the cheapest short-acting GnRH agonist and the cheapest long-acting GnRH agonist. This was considered appropriate given that as there is no difference in efficacy between the GnRH agonists, the more costly GnRH agonists would be dominated by the less costly GnRH agonists in an incremental analysis.

Based on clinical experts' opinion during a global advisory board, treatment with GnRH agonist includes add-back therapy for all patients in the model. Based on the

© Gedeon Richter UK Ltd (2023). All rights reserved

advisory board, patients treated with GnRH agonist are initiated on add-back therapy typically at three months, which is also applied in the base case. GnRH agonists may be administered for up to 12 months when given in combination with add-back therapy (100). This is in line with current recommendations for the treatment with GnRH agonists (100) (10). The European Society of Human Reproduction and Embryology emphasises the limited evidence regarding the use of GnRH agonists in combination with add-back therapy. That is, for adolescents the use of GnRH agonists should be limited to one year due to uncertainty regarding long-term consequences (10). The restricted use is assumed to be applicable for adult women as well, which is corroborated by treatment guidelines published in the US (100). Therefore, the treatment duration of GnRH agonists was restricted to one year (four treatment cycles) in the model.

For add-back therapy, two treatments are included (tibolone and raloxifene). Both treatments are assumed to be prescribed in equal shares, i.e., 50% tibolone and 50% raloxifene.

B.3.3 Clinical parameters and variables

The principal sources of data used to inform the effectiveness of Relugolix CT are the Phase 3 SPIRIT 1 and SPIRIT 2 trials. Both were replicate, multinational, 24-week, randomised, double-blind, placebo-controlled studies in premenopausal women aged 18 - 50 with moderate-to-severe pain associated with endometriosis.

The clinical inputs include probabilities of events occurring in the model, such as withdrawal from treatment, choice of surgical interventions, re-surgeries and treatment schedules.

Treatment response

Three different treatment responses are included in the model: complete response, partial response, and non-response. The model allows for the selection of two possible definitions of complete response to treatment:

- Change from baseline: Numerical rating scale (NRS) score reduction from baseline of both 2.8 for dysmenorrhea and 2.1 for NMPP and no increase of analgesic use; and

© Gedeon Richter UK Ltd (2023). All rights reserved

- Threshold: Achieving or maintaining a threshold below 4 in NRS scale (mild pain) for both NMPP and dysmenorrhea and no increase of analgesic use.

The first definition of response (hereafter referred to as “change from baseline” response) is informed by the co-primary endpoints of the SPIRIT 1 & 2 trials and is selected in the base-case. The thresholds (2.8 and 2.1) were established from an anchor-based approach which used the Patient Global Assessment (PGA) measure as an anchor to correlate with changes in NRS. PGA was collected at baseline and every fourth week at study visits during the SPIRIT 1 & 2 trials (37). The second definition of response (hereafter referred to as “threshold” response) was suggested by clinical experts during an advisory board, as the treatment aims to minimise the level of pain experienced by patients, and measuring response by achieving a certain threshold may be more feasible in clinical practice.

Partial response is defined against the same definitions as complete response (change from baseline and threshold) but relates to when patients have responded in either dysmenorrhea or NMPP after 3 months. At 6 months, patients can only be complete responders (i.e., response in both dysmenorrhea and NMPP). Partial response is thus not accounted for in the model base-case, where treatment response is assessed at 6 months.

Table 43 and Table 44 show the base-case probabilities of complete response (at three and six months) and partial response (at three months) for both Relugolix CT and GnRH agonist and by both definitions of response. Response rates for GnRH agonist are set equal to the response rates of Relugolix CT as per the conclusion of the indirect treatment comparison, as detailed in Section B.2.9.

Table 43 represents the probabilities of response for change from baseline response (A NRS score reduction of both 2.8 for dysmenorrhea and 2.1 for NMPP and no increase of analgesic use) as well as partial response (responded in either dysmenorrhea or NMPP). Table 44 represents the probabilities of response for threshold response (Achieving or maintaining a threshold below 4 in NRS scale [mild pain] in both NMPP and dysmenorrhea and no increase of analgesic use) as well as partial response (responded in either dysmenorrhea or NMPP).

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 42: Probability of response, change from baseline response

Response type	Relugolix CT	GnRH agonist	Source
Complete response: three months	40.4%	40.4%	(62, 63)
Complete response: six months	58.9%	58.9%	(62, 63)
Partial response: three months	30.6%	30.6%	(62, 63)

Table 43: Probability of response, threshold response

Response type	Relugolix CT	GnRH agonist	Source
Complete response: three months	47.4%	47.4%	(62, 63)
Complete response: six months	63.4%	63.4%	(62, 63)
Partial response: three months	25.8%	25.8%	(62, 63)

Treatment distributions

As described previously, in the base case analysis, the maximum duration of GnRH agonist treatment is set to 12 months, i.e., four model cycles (100).

Relugolix CT, BSC, GnRH agonist, and surgery may all be used with or without analgesics. Table 44 and Table 45 show the proportion of patients using analgesics before and after response for each medical treatment option and surgery in the base case analysis. Due to a lack of data, it is assumed that use of analgesics with GnRH agonist and surgery would be equivalent to the use of analgesics in the Relugolix CT arm of the SPIRIT trials. Likewise, the proportion of patients who use analgesics in the surgery arm is also assumed to be equal to that of the Relugolix CT arm of the SPIRIT trials.

Table 44: Proportion of patients using analgesics, change from baseline response

Treatment	Before response	After response	Source
Relugolix CT	90.0%	28.9%	(62, 63)
BSC	72.0%	49.1%	(62, 63)
GnRH agonist	90.0%	28.9%	Assumption
Surgery	90.0%	28.9%	(62, 63)

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 45: Proportion of patients using analgesics, threshold response (applied in model scenario)

Treatment	Before response	After response	Source
Relugolix CT	90.0%	28.3%	(62, 63)
BSC	72.0%	50.0%	(62, 63)
GnRH agonist	90.0%	28.3%	Assumption
Surgery	90.0%	28.3%	(62, 63)

Add-back therapy is standardly prescribed in addition to GnRH agonists for longer term use (apart from situations in which GnRH agonists are used short-term prior to surgery). Additionally, add-back therapy is used after oophorectomy. As confirmed by clinical experts during an advisory board, patients treated with GnRH agonists are initiated on add-back therapy after three months, while add-back therapy is initiated immediately after oophorectomy. In the base case analysis, add-back therapy is prescribed to all patients that use GnRH agonists or undergo oophorectomy.

Following discontinuation (C in Figure 38) of the intervention or comparator, patients' endometriosis symptoms are managed through subsequent treatment (BSC) or by surgery (conservative surgery or hysterectomy). Table 8 shows the base case probabilities for switching to specific treatments. For example, following discontinuation of Relugolix CT due to not achieving response following treatment initiation or loss of response over time, 16.7% of patients switch to BSC, 38.4% undergo conservative surgery, and 45.0% undergo hysterectomy. These proportions were based on patient-level information derived from the SPIRIT extension study of patients who discontinued due to either lack of efficacy (four cases) or AEs (22 cases). Of these 26 cases, five patients (16.7%) initiated BSC and 21 patients (83.3%) underwent surgery. The split between conservative surgery and hysterectomy was estimated based on a real-world evidence study by Soliman et al. which followed patients receiving treatment with leuprolide acetate for endometriosis (101). The same distribution of subsequent treatment strategy as in Relugolix CT was applied to the GnRH agonist arm.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 46: Distribution of subsequent strategy to manage endometriosis following discontinuation of intervention/comparator

Treatment switch	Relugolix CT	GnRH agonist
BSC	16.7%	16.7%
Conservative surgery	38.4%	38.4%
Hysterectomy	45.0%	45.0%
Source	(65, 101); Validated by UK KOL	

Patients in the Relugolix CT arm or the GnRH agonist arm who switched to BSC following discontinuation of the respective treatments may opt for surgery in case they do not obtain response or lose response over time from BSC. The split between the proportion of patients undergoing conservative surgery and hysterectomy is presented below.

Table 47: Distribution of subsequent surgery following discontinuation from BSC

Surgery	Proportion (%)	Source
Conservative surgery	46.0%	(101)
Hysterectomy	54.0%	(101)

Following hysterectomy (D in Figure 38), some patients may undergo reoperation. The probability of reoperation following hysterectomy in the base case analysis is set to 10% based on input from a UK-based clinical expert for the duration of the analysis. In the base case analysis, the three months probability of death following conservative surgery is 0.003% and 0.038% following hysterectomy/oophorectomy (102, 103) and is applied in conjunction with the respective surgery, i.e., during the model cycle when surgery is performed.

Treatment discontinuation

During each cycle where patients are treated with either Relugolix CT or any medical treatment (BSC or GnRH agonist) there is a probability that they will discontinue treatment. Discontinuation of treatment may signal loss of efficacy or intolerability with the treatment. The discontinuation rates for Relugolix CT and BSC were based on post-hoc analysis of discontinuation data from the SPIRIT OLE study. Time-to-discontinuation (TTD) in the SPIRIT OLE study was estimated based on Kaplan-Meier (K-M) analyses, with patients who did not discontinue censored at the last date

© Gedeon Richter UK Ltd (2023). All rights reserved

of contact (see Figure 39). The discontinuation rate over time was estimated based on the hazard rate for TTD at 3-month intervals, consistent with the cycle length employed in the model. It included events of discontinuation due to any reason. As a next step, the estimated discontinuation rate was adjusted for events such as protocol deviation which would not lead to discontinuation in clinical practice. The employed discontinuation rate was also adjusted for those cases where pregnancy or wish to get pregnant was stated as a reason for discontinuation. These patients were excluded from the discontinuation rate since treatment discontinuation in the model leads to either BSC or surgery which are not feasible options for pregnant patients or patients who wish to get pregnant.

Figure 39: Time-to-Discontinuation in the SPIRIT Extension Study

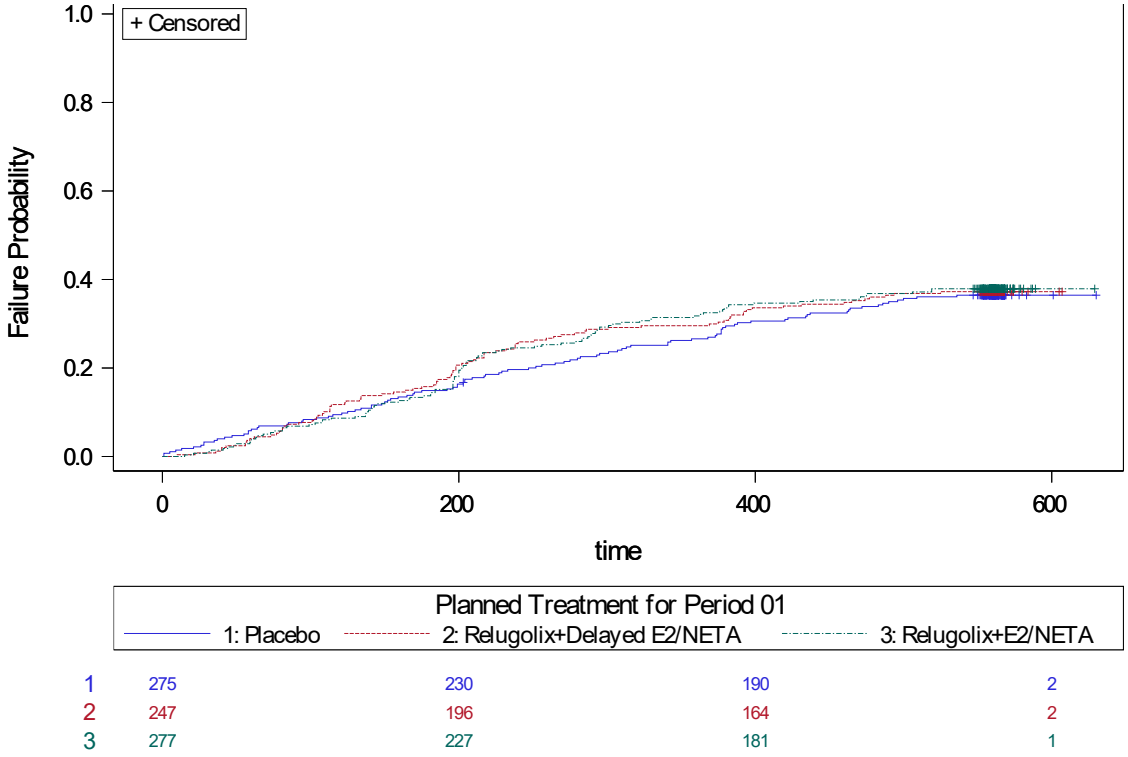


Table 48 shows the discontinuation rates of patients for Relugolix CT and GnRH agonist on a quarterly basis. It is assumed that the discontinuation rate is equal across treatments (based on the discontinuation rates for Relugolix CT). The model base-case assumes that the treatment duration of GnRH agonists is capped at one year, hence the rates presented below for GnRH agonists are only partially

© Gedeon Richter UK Ltd (2023). All rights reserved

applicable. The discontinuation rate beyond 24 months is assumed to be same as the rate at 24 months.

Table 48: Discontinuation rate over time

Treatment	Number of months since treatment response								Source
	3	6	9	12	15	18	21	24	
Relugolix CT	0.017	0.017	0.033	0.021	0.012	0.012	0.012	0.012	(62, 63)
GnRH agonist	0.017	0.017	0.033	0.021	0.012	0.012	0.012	0.012	Assumption

Pain recurrence following surgery

Patients who undergo surgery in the model face a risk of recurrence of pain. This will lead to initiation of a subsequent treatment, either with BSC or additional surgeries. Probabilities of pain recurrence in patients who underwent conservative surgery or hysterectomy were estimated from a study of post-surgery treatment outcomes among patients undergoing hysterectomy or laparoscopy for endometriosis using data on healthcare claims from the Truven Health MarketScan Commercial Claims and Encounters Database between 2004-2013 (104). Table 49 shows the pain recurrence rates on a quarterly basis by type of surgery. The pain recurrence rate beyond 24 months is assumed to be same as the rate at 24 months.

Table 49: Rate of pain recurrence by type of surgery

Type of surgery	Number of months								Source
	3	6	9	12	15	18	21	24	
Post-conservative surgery	0.322	0.073	0.079	0.086	0.094	0.104	0.116	0.131	(104)
Post-hysterectomy	0.021	0.002	0.002	0.002	0.002	0.002	0.002	0.002	(104)

Table 50: Distribution of strategies to manage endometriosis in case of pain recurrence following conservative surgery

Treatment	Proportion (%)	Source
Conservative surgery	8.9%	(104); Validated by UK KOL
BSC	80.0%	(104); Validated by UK KOL
Hysterectomy	11.1%	(104); Validated by UK KOL

© Gedeon Richter UK Ltd (2023). All rights reserved

Adverse events (AEs) and complications

Only AEs with an expected material impact on quality-of-life and/or costs (e.g., grade 3+ based on NCI-CTCAE version 5) and which would affect more than 1% of the patient population for at least one treatment of interest (Relugolix CT, BSC or GnRH agonist) were included in the model. In the SPIRIT 1 & 2 trials, no other grade 3+ AE was observed for more than 1% of the patient population with the highest probability observed for “Headache”. “Hot flush”, “Decreased libido”, “Depression”, “Increased blood pressure”, and “Hair loss” were additional adverse events highlighted by clinical experts during the global advisory board.

For all AEs, two different risk types may be selected (acute vs. constant). The risk type determines whether the probability (3-monthly probability; see Table 51) is applied throughout the treatment duration (constant) or only at treatment start, i.e., during the first model cycle (acute; total probability; see Table 52).

The adverse event profiles of Relugolix CT and BSC were derived from the Relugolix CT and placebo arms in the SPIRIT trials, respectively (62, 63).

The AE profile differs for GnRH agonist alone and GnRH agonist in combination with add-back therapy. Generally, GnRH agonists are assumed to be given in combination with add-back therapy. However, add-back therapy may not be given at GnRH agonist treatment initiation, but later. In the model base-case, it is assumed that 100% of patients who receive GnRH agonist receive add-back therapy when having been on treatment with GnRH agonist alone for three months based on the global advisory board HCP feedback.

The AE profile for GnRH agonist was derived by applying risk ratios to the probabilities for AEs linked to BSC. The risk ratios were derived from a Cochrane review on GnRH analogues for the treatment of endometriosis (see Table 51) (105). A Bucher approach (106) was used to determine the risk ratio for BSC vs. GnRH agonist in combination with add back therapy as no such analysis was available in the review (105). First, the risk ratios for GnRH agonist alone vs. placebo (i.e., BSC) were applied to the placebo data from the SPIRIT 1 & 2 trials. As a next step, the risk ratios of GnRH agonist alone vs. GnRH agonist in combination with add-back

© Gedeon Richter UK Ltd (2023). All rights reserved

therapy were applied to the calculated probabilities for GnRH agonist alone. For decreased libido, hypertension and hair loss, the same probabilities as for Relugolix CT were assumed for GnRH agonist (alone and in combination with add-back therapy) as the probability was 0% for BSC.

If the “acute” risk type is selected it is assumed that the adverse events will occur at treatment initiation and are therefore linked to GnRH agonist alone (as add-back therapy is initiated at a later time point). If the “constant” risk type is selected, the risk ratios for GnRH agonist vs. GnRH agonist in combination with add-back therapy are applied upon treatment initiation with add-back therapy and throughout the treatment duration. Additionally, the AEs are weighted according to the proportion of patients that receive add-back therapy.

Table 51: Overview of relative risk for AEs linked to treatment with GnRH agonist[a]

Adverse event	Risk ratio (BSC vs. GnRH agonist)[b]	Risk ratio (GnRH agonist vs. GnRH agonist in combination with add back therapy)
Hot flush	3.08	1.59
Headache	3.55	1.00†
Depression*	5.21‡	3.13

*No values for depression were available in the Cochrane analysis. Instead, emotional changes were used as a proxy.
†The risk ratio for headache reported for GnRH agonist vs. GnRH agonist in combination with add-back therapy was not statistically significant, hence, a risk ratio of one was applied (i.e., no difference). ‡Data for depression was not available for placebo. Instead, the relative risk of GnRH agonist vs. oral or injectable progestogens was used.
[a]The risk ratios account for all severity grades of AEs but are assumed to apply to grade 3+ events only as well.
[b]The risk ratios for BSC vs. GnRH agonists were based on the comparison of GnRH agonists vs. placebo in line with the SPIRIT 1&2 trials. One exception was the risk ratio for depression (see above).

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 52: Total probability for AEs related to Relugolix CT, BSC, and GnRH agonist

Adverse event	Relugolix CT	BSC	GnRH agonist (monotherapy)	GnRH agonist (in combination with add-back therapy)	Source
Hot flush	0.24%	0.24%	0.74%	0.47%	(62, 63, 105)
Headache	1.67%	0.48%	1.71%	1.71%	(62, 63, 105)
Depression	0.00%	0.24%	1.25%	0.40%	(62, 63, 105)
Increased blood pressure	0.24%	0.00%	0.24%	0.24%	(62, 63)
Decreased libido	0.00%	0.00%	0.00%	0.00%	(62, 63)
Hair loss	0.00%	0.00%	0.00%	0.00%	(62, 63, 107)

Table 53: 3-monthly probability for AEs related to Relugolix CT, BSC, and GnRH agonist

Adverse event	Relugolix CT	BSC	GnRH agonist (monotherapy)	GnRH agonist (in combination with add-back therapy)	Source
Hot flush	0.11%	0.11%	0.34%	0.22%	(62, 63, 105)
Headache	0.78%	0.22%	0.79%	0.79%	(62, 63, 105)
Depression	0.00%	0.11%	0.58%	0.18%	(62, 63, 105)
Increased blood pressure	0.11%	0.00%	0.11%	0.11%	(62, 63)
Decreased libido	0.00%	0.00%	0.00%	0.00%	(62, 63)
Hair loss	0.00%	0.00%	0.00%	0.00%	(62, 63)

© Gedeon Richter UK Ltd (2023). All rights reserved

Cardiovascular events

Statistically significant changes in lipid profiles have been observed with GnRH antagonists and increased low-density lipoprotein (LDL) cholesterol is a known risk factor for CV disease (69, 88). These events (i.e., excess risk of CV events) are therefore captured by the model for all treatments using risk functions from the Framingham Heart Study (FHS) based on treatment-specific changes from baseline in lipid levels at 6 months (88),(108). The FHS risk function is outlined in Equation 1.

Equation 1: The Framingham Heart Study risk function

$$10\ year\ absolute\ CVD\ risk =$$
$$0.95012 +$$
$$2.32888 * \log age +$$
$$1.20904\ * \log total\ cholesterol - 0.70833 * \log HDL\ cholesterol +$$
$$2.76157 * \log SBP,\ untreated +\ 2.82263 * \log SBP,\ treated +$$
$$0.52873 * Smoker +$$
$$0.69154 * Diabetes$$

The baseline 3-monthly probability of a CV event for patients with symptomatic endometriosis was calculated based on the mean age, mean total cholesterol, mean high-density lipoprotein (HDL) cholesterol, mean systolic blood pressure, proportion of patients with diabetes, and proportion of smokers at baseline from the SPIRIT trials using the Framingham Heart Study risk equation for CV events (108). No change in total cholesterol or HDL-levels were assumed for patients receiving BSC treatment, as was confirmed by the placebo-arm in the SPIRIT trials. The 3-monthly CV risk was adjusted as the model population aged over the projection period (e.g., in year 1, the 3-monthly risk was calculated based on age of 34, in year 2 age of 35, etc.). Treatment-specific risks of CV events were then estimated using the Framingham risk function (outlined in Table 54) and the treatment specific change from baseline in total cholesterol levels and high-density cholesterol levels observed in the Relugolix CT and placebo arms of SPIRIT 1 & 2 (see Equation 1 and Table 55). A potential treatment-related risk for CV events is applied while patients are on treatment and is removed upon treatment discontinuation.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 54: Framingham risk function for cardiovascular events

Predictor	Coefficient
Log of age	2.329
Log of total cholesterol	1.209
Log of HDL	-0.708
Log of systolic blood pressure (SBP)	2.762
Treated (with statin) SBP	2.823
Smoker	0.529
Diabetes	0.692
Source	(108)

Table 55: Treatment-specific changes in total cholesterol and HDL from baseline

Treatment	Total cholesterol	HDL	Source
Relugolix CT	2.092	-1.800	(62, 63)
BSC	0.000	0.000	(62, 63)
GnRH agonist	2.092	-1.800	Assumed to be the same as Relugolix CT
Surgery	0.000	0.000	Assumption

Change in bone mineral density and risk of fracture

Women with endometriosis may have lower bone density as a result of their disease (109). Use of GnRH agonists can also cause an immediate decrease in BMD, which may not always recover after long term use (110). A study showed that the use of GnRH agonist has been associated with a decrease in BMD at hip (-1.1% over 12 months) (111). However, add back therapy is assumed to reduce or even eliminate the risk of decreasing BMD during treatment with GnRH agonists. In the base case analysis, all patients are assumed to be initiated on add-back therapy treatment and thus no decrease in BMD and no excess fracture risk is applied.

Excess risk of fracture associated with treatments for endometriosis can be modelled based on the percentage change in BMD measured at the lumbar spine from clinical trials (62, 63, 69) as outlined in Equation 2.

© Gedeon Richter UK Ltd (2023). All rights reserved

Equation 2: Risk of Major Osteoporotic Fracture

$$Risk\ of\ major\ osteoporotic\ fracture = \left(\%\Delta BMD_{tx}\frac{BMD_{mean}}{SD_{BMD}}\right) \times RR_{MOF} \times Risk\ of\ fracture_{annual}$$

BMD = Bone Mass Density, SD = Standard Deviation, RR = Relative Risk

The model includes the option to factor in the additional risk of fractures associated with treatment. To calculate the annual risk of fracture for women aged 34 years (baseline age), the model relies on a prospective study involving 15,000 adults. This study observed the occurrence of fractures taking into account age and gender as relevant factors (112). The risk of fracture at the spine, forearm, hip, or humerus was 0.0018% annually (see Table 56). The input for an average peak bone mass (g/cm2) is based on a cohort study of Canadian women (70). Authors reported that average peak bone mass at the lumbar spine (mean = 1.05 g/cm^2, standard deviation [SD] = 0.12 g/cm^2) occurred between ages 33 to 40 years. Percentage BMD changes at the hip for Relugolix CT and BSC are estimated from the weighted average of values from the SPIRIT 1 & 2 trials (62, 63) and are presented in Table 57.

If relevant, the excess risk of fracture associated with treatment for endometriosis can be calculated by deriving the percentage change in BMD from average peak bone mass and by multiplying the estimated relative risk (RR) of major osteoporotic fracture per unit (SD) change in BMD and by the baseline annual risk of fracture. The RR of major osteoporotic fracture per unit change in BMD is based on a longitudinal study of BMD measurements and incident fractures among postmenopausal women from the Women's Health Initiative (RR = 1.12; 95% confidence interval [CI] 1.04, 1.21) (97). The treatment-related excess risk of fracture can be applied for patients that are on active treatment and is removed upon treatment discontinuation.

Table 56: Baseline fracture and peak bone mass

Indicator	Value	Source
Annual fracture risk	0.0018%	(112)
RR per unit decrease in BMD	1.120	(70)
Average peak bone mass (g/cm^2)	1.046	(70)
SD bone mass (g/cm^2)	0.123	(70)

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 57: Treatment specific changes in BMD from baseline

Treatment	Change in BMD at hip	Source
Relugolix CT	0.00%	(62, 63)
BSC	0.00%	(62, 63)
GnRH agonist	0.00%	Assumed to be the same as Relugolix CT
Surgery	0.00%	Assumption

Complications related to surgery

Complications related to surgery were identified during an advisory board. The risk of complications for urinary tract infection, fistula and urinary retention/complication was derived from a prospective Finnish study on complications following 5,279 hysterectomies (113) and are presented in Table 58. Complications related to surgery are assumed to be of acute nature and persist for a period of three months (one model cycle) in the base-case analysis, based on collected data on intra-operative complications (113).

Table 58: 3-monthly risk of complications related to surgery

Complication	Risk type	Conservative surgery	Hysterectomy	Oophorectomy	Source
Urinary tract infection	Acute	0.00%	1.42%	1.42%	(113, 114)
Fistula	Acute	0.00%	0.04%	0.04%	(113, 115)
Urinary retention/ complication	Acute	0.00%	0.99%	0.99%	(113, 114)
Impact of surgery on other organs (e.g., bowel problems)	Acute	3.00%	3.00%	3.00%	UK KOL input

B.3.4 Measurement and valuation of health effects

The clinical benefit of Relugolix CT is evaluated based on health-related quality of life (HRQoL) and is measured in quality-adjusted life years (QALYs) as per the NICE reference case. Each health state in the model is associated with a utility weight specific to that state. The utility weight was derived from prospective data collected in the SPIRIT 1 & 2 trials. Disutilities associated with surgeries, complications following surgery, and AEs were derived from the published literature.

© Gedeon Richter UK Ltd (2023). All rights reserved

Health-related quality-of-life data from clinical trials

EQ-5D-5L data was measured at baseline and at week 24 in the SPIRIT 1 and 2 trials. The EQ-5D-5L data were first mapped to 3L using the NICE DSU age-sex based mapping (94) (95). Utilities were then derived using UK value set published by Dolan for EQ-5D-3L (96).

Table 59 below displays EQ-5D utilities that were estimated for each patient at baseline and at Week 24 for the pooled modified intent-to-treat (mITT) population from the SPIRIT 1 & 2 trials with treatment arms combined (i.e., all patients).

Table 59: Summary of EQ-5D-3L utility values at baseline and Week 24 for the pooled mITT population

Timepoint	n	Mean (95% CI)	Standard Deviation
Baseline	821	0.58 (0.57, 0.60)	0.24
Week 24	684	0.80 (0.78, 0.81)	0.20

n: Number of subjects included in analysis; CI: Confidence interval

Utility at Week 24 for responders and non-responders was estimated using data from the SPIRIT 1 & 2 trials through OLS regression models with robust standard errors. The SPIRIT 1 & 2 trials reported EQ-5D questionnaires at baseline and week 24 only, thus no repeated measure model could be used. Only non-missing assessments for the baseline and Week 24 visit were included in the regression analyses.

Mapping

Three regression models were run on the pooled mITT population (treatment arms combined), and all models included an intercept term and covariates for baseline EQ-5D-5L utility value and indicator variable for response status (coded “1” for responder and “0” for non-responder as reference) at Week 24. Since baseline utility was already included as a covariate, baseline mean NRS score and age were not considered as covariates to prevent any multicollinearity. The three OLS regression models are described below:

a) One model without a treatment arm covariate

© Gedeon Richter UK Ltd (2023). All rights reserved

$$EQ\text{-}5D_{24\,weeks} = Intercept + \alpha.EQ\text{-}5D_{baseline} + \beta.Response$$

b) One model with a binary treatment arm covariate (coded “1” for relugolix + E2/NETA arm and “0” for placebo arm as reference)

$$EQ\text{-}5D_{24\,weeks} = Intercept + \alpha.EQ\text{-}5D_{baseline} + \beta.Response + \gamma.Treatment$$

c) One model with a binary treatment arm covariate (coded “1” for relugolix + E2/NETA arm and “0” for placebo arm as reference) and with an interaction term “treatment_arm * response_status”

$$EQ\text{-}5D_{24\,weeks} = Intercept + \alpha.EQ\text{-}5D_{baseline} + \beta.Response + \gamma.Treatment + \delta.Treatment \times Response$$

The OLS models were run for the two types of response definition (see section Treatment response for definition), presented in Table 60 and Table 61.

As the treatment arm covariate and the “treatment x response” interaction term were not statistically significant in models b) and c), the most parsimonious regression model a), including an intercept term and covariates for baseline EQ-5D utility value and indicator variable for response status (with non-responder as reference) at Week 24 was used as the final model.

Outputs from the regression model are summarized in Table 60 for the change from baseline response definition, and Table 61 for the threshold response definition for the final model.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 60: Output from Ordinary Least Squares (OLS) regression of mapped EQ-5D-3L utility values at Week 24, change from baseline response

Indicator	Estimate	SE	95% CI Low	95% CI High	Z statistic	P-Value
Intercept	0.5845	0.02547	0.5346	0.6344	22.95	<.0001
Baseline utility	0.2292	0.03374	0.1631	0.2953	6.79	<.0001
Responder	0.1650	0.01299	0.1396	0.1905	12.71	<.0001
Non-responder*	-	-	-	-	-	-

*Referent group; SE: Standard error; CI: Confidence interval; Note: 95% CIs are calculated based on the normal distribution. Only subjects with non-missing responses for all five dimensions of the EQ-5D-5L questionnaire at both baseline and Week 24 are included in analysis.

Table 61: Output from Ordinary Least Squares (OLS) regression of mapped EQ-5D-3L utility values at Week 24, threshold response

Indicator	Estimate	SE	95% CI Low	95% CI High	Z statistic	P-Value
Intercept	0.5928	0.02422	0.5454	0.6403	24.48	<.0001
Baseline utility	0.2002	0.03272	0.1360	0.2643	6.12	<.0001
Responder	0.1714	0.01304	0.1459	0.1970	13.15	<.0001
Non-responder*	-	-	-	-	-	-

*Referent group; SE: Standard error; CI: Confidence interval; Note: 95% CIs are calculated based on the normal distribution. Only subjects with non-missing responses for all five dimensions of the EQ-5D-5L questionnaire at both baseline and Week 24 are included in analysis.

Health state utilities

Health state utilities were derived from this regression using the mean baseline utility value across both treatment arms (0.5838). For instance, responder utility value for the change from baseline response definition was estimated using outputs displayed as:

$$EQ\text{-}5D_{24\,weeks}(responder) = Intercept + \alpha.EQ\text{-}5D_{baseline} + \beta.Response$$
$$EQ\text{-}5D_{24\,weeks}(responder) = 0.5845 + 0.2292 \times 0.5838 + 0.1650 \times 1 = 0.8839$$

Utility of initial treatment corresponds to the baseline value from the SPIRIT trials whereas the utility values of partial response is assumed to be the average of response and non-response. Health state utilities are outlined in Table 62.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 62: Health state utilities

Response type	Utility	95% CI Low	95% CI High	P-Value	Source
Initial treatment	0.5838	0.5676	0.5999	N/A	(62, 63)
Change from baseline response					
Responder	0.8839	0.8697	0.8981	<.0001	(62, 63)
Partial response	0.8014	0.7761	0.8267	N/A	Average of utility input for responder and non-responder
Non-responder	0.7189	0.6979	0.7399	<.0001	(62, 63)
Threshold response					
Responder	0.8816	0.8672	0.8960	<.0001	(62, 63)
Partial response	0.7959	0.7703	0.8215	N/A	Average of utility input for responder and non-responder
Non-responder	0.7102	0.6891	0.7313	<.0001	(62, 63)

Note: Depending on the timepoint at which complete response is evaluated, the initial treatment utility for surgery is applied to all patients during the first (if response evaluated at three months) or the first two (if evaluated at six months) model cycles before undergoing either conservative surgery or hysterectomy.

As presented above, utility for initial treatment is applied to the health state “Initial treatment”. Utility of response is applied to health states “Response”, “Response BSC”, “Post-hysterectomy stable”, and “Post-conservative (PCS) response”. The utility of partial response is applied to the health state “Partial response”. Utility of non-response is applied to health states “BSC”, “Non-response”, “Non-response BSC”, “Post-hysterectomy recurrence”, and “PCS recurrence”. The significant difference in utility between non-responder (0.7189 or 0.7102 depending on the response definition) and the baseline utility (0.5838) in the SPIRIT trials suggests that symptoms in these patients improve but not sufficiently to meet the criteria of treatment response. It is likely that the utility in patients who experience repetitive episodes of non-response regress back to the levels observed at baseline in the SPIRIT trials. The model does not account for the number of failures to respond to different strategies (medical treatment and surgeries), and thus the utility of non-response is thus used equally for both health states “Non-response”, following initial

© Gedeon Richter UK Ltd (2023). All rights reserved

treatment, and “Non-response BSC”. To not decrease further the utility after failing another line of treatment is likely a conservative approach.

Table 63: Utilities applied to health states

Utility applied	Health states
Utility for initial treatment	• Initial treatment
Utility of response	• Response • Response BSC • Post-hysterectomy stable • Post-conservative surgery (PCS) • Response
Utility of partial response	• Partial response
Utility of non-response	• Non-response • Non-response BSC • Waiting time before surgery • Post-hysterectomy recurrence • PCS recurrence • BSC

Health-related quality-of-life studies

An SLR was conducted to identify HRQoL data and is detailed in Appendix H.

Adverse reactions

Disutilities associated with treatment-related adverse events

The disutility of each AE is applied to the cycle in which the event occurs. The risk type determines whether the probability (3-month probability) is applied throughout the treatment duration (constant) or only at treatment start (acute). The type of event (acute vs. constant) determines how disutilities are accounted for in the model. In the case of AEs that have an immediate nature, AEs are applied at treatment initiation. If an AE is selected to be constant it is applied throughout the treatment period (constant consequence). For the base case analysis all AEs are assumed to be constant and to persist while on treatment. Disutilities from AEs for the base case analysis are outlined in Table 64. The reported values represent annual disutility and were thus adjusted for the 3-month cycle length. The disutility of hot flush was derived from a Canadian study for pre-menopausal women with uterine fibroids and is assumed to be applicable to women with endometriosis (116). Since only grade 3+

© Gedeon Richter UK Ltd (2023). All rights reserved

AEs were included from the SPIRIT 1 & 2 trials, the disutility of headache was assumed to be comparable with migraine. The disutility of headache was sourced from a cost-effectiveness analysis on moderate-to-severe migraine (117). The disutilities for decreased libido and depression were derived from a study by Wang et al. who investigated the cost-effectiveness of elagolix versus leuprolide acetate for treating moderate-to-severe endometriosis pain in the US. No disutility is assumed for hypertension. For hair loss a disutility of -0.045 is assumed based on a study assessing health state utilities for non-small cell lung cancer (118).

Table 64: AE disutilities

AE	Disutility	Source
Hot flush	-0.060	(62, 63, 116)
Headache	-0.340	(117)
Decreased libido	-0.049	(91)
Depression	-0.120	(91)
Hypertension	0	Assumption
Hair loss	-0.045	(118)

Surgery related disutilities

The model also accounts for disutilities related to surgery. The disutility may be acute or long-term. Acute disutilities associated with surgery are assumed to represent the detrimental effect that undergoing the surgical procedure may have on the patient's quality of life. The acute disutility is thus only applied to the subsequent cycle following the surgery. The acute disutility weight for hysterectomy was derived from a randomised clinical trial evaluating the cost-effectiveness of laparoscopic hysterectomy compared with standard hysterectomy (Vaginal hysterectomy: - 0.02; Abdominal hysterectomy: - 0.07; Laparoscopic hysterectomy - 0.04) (119) (120). The disutility of hysterectomy was calculated adjusting for the route of hysterectomy as reported in Maresh et al 2002 (weighted average of disutilities according to the proportion of route of hysterectomy) (102). The proportions of patients allocated to each surgery type are presented in Table 47. The acute disutility of hysterectomy is assumed to be applicable to oophorectomy (Table 65). The disutility of conservative surgery is assumed to be equal to that of laparoscopic hysterectomy (-0.04).

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 65: Acute disutility linked to surgical procedures

Type of surgery	Disutility	Source
Conservative surgery	-0.040	(119, 120)
Hysterectomy	-0.0541	(119, 120)
Oophorectomy	-0.0541	(119, 120)

If the disutility is of a long-term nature, it reflects the permanent consequences on HRQoL and it is assumed to apply to all subsequent model cycles following the surgery. Since conservative surgery is limited to removing endometrial tissue while preserving the uterus, conservative surgery is assumed to not have a negative long-term impact. Hence, no long-term disutility is applied (Table 66). The long-term disutility of hysterectomy is applied to the time spent in the post-hysterectomy health states (Stable, Recurrence and Reoperation). Following hysterectomy, women are no longer able to become pregnant, which is assumed to have an impact on their quality of life in the long term. A disutility of 0.180 is applied to account for this (Table 66), representing the disutility linked to infertility and was derived from a global burden of disease report published by the World Health Organization (121).

Table 66: Long-term disutility following surgery

Post-surgery	Disutility	Source
Post-hysterectomy	-0.180	(121)
Post-conservative surgery	0.000	Assumption

In addition to the disutilities directly related to surgery, disutilities linked to surgery complications were applied (Table 67). The disutilities for the urinary tract infection and urinary retention were derived from a cost-effectiveness study of surgical treatment for benign prostatic enlargement (114). The disutility for fistula was derived from a cost-effectiveness study on compare prostate cryotherapy to androgen deprivation therapy for treatment of radiation recurrent prostate cancer (115). Despite including only men, the disutility inputs from these two studies were assumed to be representative for women as well. The disutility for the impact of surgery on other organs was based on a cost-effectiveness analysis on methylnaltrexone bromide for the treatment of opioid-induced constipation in patients with advanced illness (122). Depending on the nature of the complication, the disutilities linked to the respective complication may be applied in conjunction with

© Gedeon Richter UK Ltd (2023). All rights reserved

the surgery (acute) or from the time of surgery and beyond (constant). For the base case analysis, complications were applied as acute and are accounted for only at the time of surgery but not beyond.

Table 67: Disutilities of long-term complications from surgery

Complication	Disutility	Source
Urinary tract infection	-0.006	(114)
Fistula	-0.150	(115)
Urinary retention/ complication	-0.006	(114)
Impact of surgery on other organs (e.g., bowel problems)	-0.017	(122)

Health-related quality-of-life data used in the cost-effectiveness analysis

Table 68: Summary of utility values for cost-effectiveness analysis

State	Utility value: mean (standard error)	95% confidence interval	Reference in submission (section and page number)	Justification
Initial treatment	0.5838	(0.5676, 0.5999)	Mapping, Health state utilities (page 160)	Analysis of prospective EQ-5D data taken from trials
Change from baseline response (model base-case)				
Responder	0.8839	(0.8697, 0.8981)	Mapping, Health state utilities (page 160)	Analysis of prospective EQ-5D data taken from trials
Partial response	0.8014	(0.7761, 0.8267)	Mapping, Health state utilities (page 160)	Analysis of prospective EQ-5D data taken from trials
Non-responder	0.7189	(0.6979, 0.7399)	Mapping, Health state utilities (page 160)	Analysis of prospective EQ-5D data taken from trials
Threshold response (applied in scenario analysis)				
Responder	0.8816	(0.8672, 0.8960)	Mapping, Health state utilities (page 160)	Analysis of prospective EQ-5D data

© Gedeon Richter UK Ltd (2023). All rights reserved

				taken from trials
Partial response	0.7959	(0.7703, 0.8215)	Mapping, Health state utilities (page 160)	Analysis of prospective EQ-5D data taken from trials
Non-responder	0.7102	(0.6891 - 0.7313)	Mapping, Health state utilities (page 160)	Analysis of prospective EQ-5D data taken from trials
Disutilities associated with treatment-related AEs				
Hot flush	-0.06	(-0.05, -0.07)	Hux et al., 2015 (62, 63, 116) (Adverse reactions, page 162)	Literature
Headache	-0.34	(-0.31, -0.37)	Xu et al., 2011 (117) (Adverse reactions, page 162)	Literature
Decreased libido	-0.05	(-0.04, -0.05)	Wang et al., 2019 (91) (Adverse reactions, page 162)	Literature
Depression	-0.12	(-0.11, -0.13)	Wang et al., 2019 (91) (Adverse reactions, page 162)	Literature
Hypertension	0	NA	(Adverse reactions, page 162)	Assumption
Hair loss	-0.05	(-0.04, -0.05)	Nafees et al., 2008 (118) (Adverse reactions, page 162)	Literature
Surgery-related disutilities				
Conservative surgery	-0.04	(-0.04, -0.04)	Geale et al., 2017; Sculpher et al., 2004 (119, 120) (Adverse reactions, page 163)	Literature
Hysterectomy	-0.05	(-0.05, -0.06)		
Oophorectomy	-0.05	(-0.05, -0.06)		

© Gedeon Richter UK Ltd (2023). All rights reserved

Long-term disutility following surgery				
Post-hysterectomy	-0.18	(-0.16, -0.20)	World Health Organization (WHO), 2004 (121) (Adverse reactions, page 163)	Literature
Post-conservative surgery	0	NA	(Adverse reactions, page 163)	Assumption
Disutilities of long-term complications from surgery				
Urinary tract infection	-0.01	(-0.01, -0.01)	Armstrong et al., 2009 (114) (Adverse reactions, page 164)	Literature
Fistula	-0.15	(-0.15, -0.17)	Boyd et al., 2015 (115) (Adverse reactions, page 164)	Literature
Urinary retention/ complication	-0.01	(-0.01, -0.01)	Armstrong et al., 2009 (114) (Adverse reactions, page 164)	Literature
Impact of surgery on other organs (e.g., bowel problems)	-0.02	(-0.02, -0.02)	Earnshaw et al., 2010 (122) (Adverse reactions, page 164)	Literature

B.3.5 Cost and healthcare resource use identification, measurement and valuation

An SLR was conducted to identify cost and HRU data and is detailed in Appendix I.

Intervention and comparators' costs and resource use

Drug acquisition costs

The costs of medical treatment options, including Relugolix CT, GnRH agonist and BSC is applied each cycle that patients are on treatment. Drug acquisition costs are not applied whilst patients are on treatment breaks. The drug cost calculations are outlined in Table 67. As discussed previously, a 50/50 split of patients amongst the

© Gedeon Richter UK Ltd (2023). All rights reserved

least costly short-acting and long-acting GnRH agonist is assumed in the GnRH agonist treatment arm. Thus, only the costs of short and long-acting triptorelin are included for the GnRH agonist arm.

No drug wastage is accounted for in the drug cost calculations.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 69: Drug costs and dosing

Treatment	Package cost (£)	Doses per package	Administrations per cycle	Total drug cost per cycle (£)	Source
Relugolix CT	72.00	28.00	91.31	234.80	(123)
BSC					
Hormonal treatment (Dienogest)	20.50	28.00	91.31	66.85	(124)
Estrogen-progestin oral contraceptive	0.85	21.00	91.31	3.70	(125)
Medroxyprogesterone acetate	2.47	10.00	273.94	67.66	(126)
Levonorgestrel-releasing intrauterine system	71.00	1.00	-	-	(127, 128)
GnRH agonist					
Short-acting GnRH agonist					
Triptorelin (3.75 mg)	69	1	3	225.02	
Long-acting GnRH agonist					
Triptorelin (11.25 mg)	207.00	1.00	1.00	207.00	(129)

© Gedeon Richter UK Ltd (2023). All rights reserved

Administration costs

The relevant forms of administration for the intervention and comparator, as well as subsequent treatment (BSC) encompass intramuscular injection, subcutaneous injection, intranasal administration, and oral administration. Relugolix CT is an orally administered tablet, thus no administration costs are assumed. For GnRH agonist, it was assumed that these treatments would be administered by a nurse based in a general practitioner (GP) surgery. This assumption was validated with KOLs who were asked to provide information regarding who administers GnRH agonist treatment, the duration required for treatment administration and the setting in which they would be administered (hospital or GP surgery).

The unit costs of treatment administration are presented below.

Table 70: Administration costs

Mode of administration	Administration - Resource use	Unit cost (£)	Source
Oral/ intranasal administration	Self-administered	0	Assumption
Intramuscular/subcutaneous injection	GP practice/ Specialty care Nurse-administered	26	Cost of qualified nurse for 30 minutes, Unit Costs of Health & Social Care 2022 (130)

Concomitant medication

Concomitant medication is taken by patients in combination with medical treatment. Analgesics, i.e., NSAIDs, are included for pain management and the frequency of use is based on the SPIRIT trials and is driven by patients' response status as discussed in section Treatment distributions. The model assumes that patients do not require opioids, as these treatments are rarely prescribed in Europe.

The model assumes that patients use dienogest as a post operative hormonal treatment, which was identified through a targeted literature review and confirmed/supplemented by clinical experts at an advisory board. The costs of analgesics and hormonal treatment are presented below.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 71: Concomitant medication cost

Concomitant medication	Package cost (£)	Doses per package	Administrations per cycle	Total cost per cycle (£)	Source
NSAIDs (Ibuprofen 400mg)	4.90	60	273.94	22.37	(124)
Hormonal treatment (Dienogest)	20.50	28	91.31	66.85	(131)

Add-back therapy

As detailed previously, add-back therapy is standardly prescribed in addition to GnRH agonists for longer term use, apart from situations in which GnRH agonists are used short-term prior to surgery. Additionally, add-back therapy is used after oophorectomy. With GnRH agonists, patients are initiated on add-back therapy typically at three months whereas after oophorectomy, add-back therapy is initiated directly following surgery. This assumption was confirmed with clinical experts at an advisory board. In the base case analysis, an equal split (50/50%) between tibolone and raloxifene as add-back therapy is considered.

Table 72: Cost of add-back therapy

Add-back therapy	Cost per package (£)	Tablets per package	Total drug cost per cycle (£)	Source
Tibolone	14.13	84	15.36	(132)
Raloxifene	4.55	28	14.84	(98)

Table 73: Proportion of patients using add-back therapy (Tibolone)

Treatment	Proportion of patients	Source
GnRH agonist	100.0%	KOL input
Oophorectomy	100.0%	KOL input

© Gedeon Richter UK Ltd (2023). All rights reserved

Visits to health care professionals, tests, and procedures

Administration of treatment, monitoring of patients on active medical treatment, follow-up of patients who discontinue treatment or those who undergo surgery require visits with health care professionals as well as certain tests or procedures. Table 74 lists the unit costs of visits to healthcare professionals.

Table 75 outlines the resource use linked to administration and monitoring until the time of evaluation of treatment response (6 months). Table 76 shows the long-term follow-up (i.e., at subsequent model cycles) for each comparator, outlining the resource use beyond the first three cycles. Resource utilisation is calculated quarterly (per model cycle), for example, if one annual visit is required, the quarterly utilisation is ¼, i.e., 0.25. The resource use frequencies are assumed to be equal for Relugolix CT and GnRH agonist.

Table 74: Unit costs of healthcare professional visits

Healthcare provider	Cost per visit (£)	Source
Gynaecologist	181.26	(99, 133)
General Practitioner	42.00	(134)
Nurse	7.99	(134)

Table 75: Resource use linked to administration and monitoring until treatment response evaluation

Healthcare provider	Administration	
	Treatment initiation	6-month follow up
Gynaecologist	1	1
General Practitioner	0	0
Nurse	0	0
Source	UK KOL input	

Table 76: Resource use linked to long-term follow-up, by treatment arm

Healthcare provider	Per cycle frequencies for long-term follow-up	
	Relugolix CT	GnRH agonist
Gynaecologist	0	0
General Practitioner	0	0
Nurse	1	1
Source	UK KOL input	

© Gedeon Richter UK Ltd (2023). All rights reserved

The use of medical tests and procedures such as ultrasound, MRI, DEXA scan and blood tests was queried with KOLs who attended an advisory board. The KOLs explained that patients who are being treated with a pharmacological treatment such as GnRH agonist are not subject to any additional monitoring with tests and procedures. Patients who undergo surgery however, are subject to this type of monitoring both prior to and post-surgery. The pre-surgery resource use frequency is provided in Table 77. The unit costs of tests and procedures is provided in Table 78. An annual ultrasound is also incurred as part of the follow-up for patients who have surgery.

Table 77: Average number of tests per surgery

Test/procedure	Laparoscopy	Hysterectomy	Oophorectomy	Source
Ultrasound	1	1	1	UK KOL input
Magnetic Resonance Imaging Scan	1	1	1	UK KOL input

Table 78: Cost of tests and procedures

Test/procedure	Cost (£)	Source
Ultrasound	181.00	(135)
Blood test	2.92	(136)
Magnetic Resonance Imaging (MRI) Scan	114.00	(135)
Dexa scan	61.00	(135)

Cost of surgery

Costs of different surgical procedures were sourced from pricelists available through the NHS and were validated with KOLs. The cost of laparoscopy is based on the NHS England 2022/23 national tariff workbook (Annex A) unit costs for Major, Intermediate and Minor Laparoscopic or Endoscopic, Upper Genital Tract Procedures, where an average of the three unit costs has been taken (135). The cost of conservative surgery is assumed to be the same as the cost of laparoscopic hysterectomy (£3,337.00). The cost of hysterectomy represents different routes of hysterectomy, namely the vaginal, abdominal, and laparoscopic route, and the respective proportion of patients (102). The costs and proportions are presented below.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 79: Cost of hysterectomy

Route of hysterectomy	Cost (£)	Proportion of patients	Source
Vaginal	4,414.00	30%	Major Open Upper Genital Tract Procedures, average of CC scores 0-5+, currency codes MA07G, MA07F, MA07E; weighted average of elective, day case, and outpatient unit costs (102, 135)
Abdominal	4,414.00	67%	Major Open Upper Genital Tract Procedures, average of CC scores 0-5+, currency codes MA07G, MA07F, MA07E; weighted average of elective, day case, and outpatient unit costs (102, 135)
Laparoscopic	3,337.00	3%	Major, Laparoscopic or Endoscopic, Upper Genital Tract Procedure, average of CC scores 0-2+, currency codes MA08B, MA08A; weighted average of elective, day case, and outpatient unit costs (102, 135)

The cost of oophorectomy is not available in the NHS England 2022/23 national tariff workbook. It was instead calculated using the unit price for hysterectomy (£4,381.69) from the NHS England 2022/23 national tariff workbook (as presented above) and weighting this by multiplying it with the proportion (0.61=2,275/3,703) of hysterectomy (3,703) and oophorectomy (2,275) costs from the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS) (135) (137). The costs are presented below.

Table 80: Cost of surgery

Type of surgery	Cost (£)	Procedure	Source
Conservative surgery (Laparoscopy)	3,337.00	Major, Laparoscopic or Endoscopic, Upper Genital Tract Procedure, average of CC scores 0-2+, currency codes MA08B, MA08A; weighted average of elective, day case, and outpatient unit costs	(135)
Oophorectomy	2,691.96	Unit price for hysterectomy and oophorectomy from the NHS England 2022/23 national tariff workbook multiplied with the proportion of hysterectomy and oophorectomy.	(135, 137)
Hysterectomy	4,381.69	See table below	See Table 80

© Gedeon Richter UK Ltd (2023). All rights reserved

Health-state unit costs and resource use

The frequency of monitoring and disease management related healthcare resource use (HRU) is not driven by health states in the model but instead by whether patients are on active pharmacological treatment (Relugolix CT or GnRH agonist) or if they have undergone surgery. The HRU assumptions applied in the model are summarized above in section Visits to health care professionals, tests, and procedures.

Adverse reaction unit costs and resource use

The costs of AEs and complications are applied in the cycle during which they occur. Costs for AEs, complications from surgery and cardiovascular events and fractures are presented below.

Table 81: Cost of AEs related to medical treatment

Adverse event	Cost (£)	Cost Detail	Source
Hot flush	0	No cost incurred as it is assumed that this will be self-managed and no treatment sought	Assumption
Headache			
Decreased libido	42.00	Based on unit cost for a GP, per surgery consultation lasting 9.22 minutes, excluding travel	(134)
Depression			
Blood pressure			
Hair loss			

Table 82: Cost of complications related to surgery

Complication	Cost (£)	Cost Detail	Source
Urinary tract infection	457.35	Based on the non-elective short stay unit cost for Kidney or Urinary Tract Infections, without Interventions, with CC Score 0-1 (LA04S)	(99)
Fistula	4,039.00	Based on cost of fistula	(115)
Urinary retention/ complication	612.62	Based on the non-elective short stay unit cost for Kidney or Urinary Tract Infections, without Interventions, with CC Score 4-7 (LA04Q)	(99)
Impact of surgery on other organs (e.g., bowel problems)	1,020.76	Based on the non-elective short stay unit cost for Diagnostic Colonoscopy, 19 years and over (FE32Z)	(99)

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 83: Cost of cardiovascular events and fractures

Event	Cost (£)	Cost Detail	Source
Cardiovascular event	2,648.00	Based on an average of the costs for Angina (EB13A-D), Actual or Suspected Myocardial Infarction (EB10A-E), Stroke (AA35A-F), Heart failure or Shock (EB03A-E), Transient Ischaemic Attack (AA29C-F), Peripheral Vascular Disorders (YQ50A-E)	(99)
Hip fracture	8,686.10	Based on an average of the elective costs for Hip Fracture without Interventions, with CC Score 0-3 to 12+ (HE11H to HE11E)	(99)

Societal costs

There is a societal burden associated with endometriosis, predominantly due to absenteeism and presenteeism. Lost productivity or lost work time associated with endometriosis are costly for the society. These societal costs were explored in a scenario analysis.

In the scenario where the societal perspective is adopted, the model estimates the value of lost production due to the number of days absent from work (absenteeism), and the reduction in daily productivity (presenteeism). For absenteeism, the model takes as input number of days absent from work per month (e.g., due to illness or healthcare visits) as well as the number of days absent from work per month following surgery. For presentism, the model also includes the number of days per month lost due to lower productivity. The model estimates the value of lost production by multiplying the number of days (due to absenteeism and presenteeism) by the value of lost productivity (per day). The value of lost productivity (143.40 GBP) is derived from the average national gross income (32,300 GBP (138)) and divided by the number of working days per year (225.25). Absenteeism and presenteeism in patients undergoing medical therapy is split between patients with and without response (Table 42). The estimates were derived from a cross-sectional study including an endometriosis cohort of 745 women and a symptomatic control cohort of 587 women. The estimates from the endometriosis cohort were used in the non-response group whereas the estimates from the control cohort were assumed to reflect absenteeism and presenteeism in patients who respond to treatment in the model (19).

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 84: Number of days absent from work (absenteeism) and lost productivity per month (presenteeism)

Response following surgery	Absenteeism (days)	Presenteeism (days)	Source
With response	1.8	2.8	(19)
Without response	2.4	3.5	(19)

Absenteeism following surgery was derived from a Dutch prospective cohort study. The time to return to work following diagnostic, minor, intermediate, and major surgery was assessed in 148 women aged 18–65 years scheduled for gynaecological surgery for benign indication. Conservative surgery (laparoscopy) was assumed to be a minor surgery while hysterectomy and oophorectomy were assumed to be a major surgery. The time to return to work was adjusted for number of working days per week (139). Table 85 shows the number of days absent from work per surgery.

Table 85: Number of days absent from work per surgery

Type of surgery	Absenteeism (days)	Source
Conservative surgery	10	(139)
Hysterectomy	49	(139)
Oophorectomy	49	(139)

The model estimates the number of days per month with reduced productivity, calculated as a percentage decrease from 100% productivity and applied per model cycle. The number of days lost due to absenteeism and presenteeism are then added together and multiplied by the average annual gross income to produce the societal cost for endometriosis.

B.3.6 Severity

The QALY shortfall for Relugolix CT was calculated using the online calculator tool published by Schneider et al., 2021. Relugolix CT does not meet the criteria for a severity weight as it achieves a QALY weighting of 1.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 86: summary features of QALY shortfall analysis

Factor	Value (reference to appropriate table or figure in submission)	Reference to section in submission
Sex distribution	100% female	Patient population
Starting age	33.9	Patient population

There have been no prior NICE evaluations for interventions for treating pain associated with endometriosis thus we are not able to provide a summary list of QALY shortfall from previous evaluations.

B.3.9 Summary of base-case analysis inputs and assumptions

A tabulated summary of the base-case analysis inputs is provided in Appendix M.

© Gedeon Richter UK Ltd (2023). All rights reserved

Assumptions

A summary of the assumptions applied in the model is presented below.

Table 87: Overview of the base case assumptions

Variable	Assumption	Rationale
Discontinuation	The same discontinuation rate is assumed for all medical treatments (except for GnRH agonists; those are assumed to be terminated after one year). Data on discontinuation were derived from the SPIRIT 1 & 2 trials for Relugolix CT.	Data on discontinuation from the SPIRIT 1 & 2 trials is assumed to be most relevant for the cost-effectiveness analysis of Relugolix CT for the symptomatic treatment of endometriosis in adult women of reproductive age with a history of previous medical or surgical treatment for their endometriosis. An indirect treatment comparison showed no statistically significant difference in terms of treatment effect between Relugolix CT and GnRH agonist. Hence, the same discontinuation rate is assumed.
Use of analgesics	The same use of analgesics is assumed for Relugolix CT, GnRH agonist and surgery. It is applied in accordance with the response definition.	Data on use of analgesics from the SPIRIT 1 & 2 trials is assumed to be most relevant for the cost-effectiveness analysis of Relugolix CT for the symptomatic treatment of endometriosis in adult women of reproductive age with a history of previous medical or surgical treatment for their endometriosis. An indirect treatment comparison showed no statistically significant difference in terms of treatment effect between Relugolix CT and GnRH agonist. Hence, the same use of analgesics is assumed in the GnRH agonist arm. Due to a lack of data, the same use of analgesics was also assumed for patients who undergo surgery.
Treatment duration	A treatment duration function allows treatment to be discontinued after a prespecified period of time, independent of response. In the base case analysis, patients on Relugolix CT and BSC will continue treatment as long as they	The growth of the endometriotic tissue is estrogen-dependent; therefore, endometriosis commonly occurs only until menopause. It is assumed that, if response is achieved, patients may continue treatment until discontinuation due to other reasons (e.g., adverse events). For GnRH agonists, a maximum treatment duration of one year is assumed based on treatment recommendations and available evidence.

© Gedeon Richter UK Ltd (2023). All rights reserved

	respond to treatment or until menopause. Patients treated with GnRH agonist will discontinue treatment after 12 months.	
Surgery	Options for surgery include conservative surgery, hysterectomy, and oophorectomy. There is no limitation on the number of conservative surgeries, although in the base case analysis, patients receive on average less than one conservative surgery during the time horizon. A waiting time of six months is assumed before undergoing surgery.	Patients may receive oophorectomy after hysterectomy and only undergo a hysterectomy and oophorectomy once, whereas patients may have several conservative surgeries. Based on feedback from an advisory board, patients are assumed to experience pain recurrence after surgery that may require additional surgeries. It is assumed that patients wait six months before undergoing surgery. During that time, patients are assumed to receive BSC.
Population	No diagnosis test and associated cost have been included in the model.	Relugolix CT is indicated for symptomatic treatment of endometriosis in adult women of reproductive age with a history of previous medical or surgical treatment for their endometriosis. Hence, all patients are assumed to have been diagnosed at baseline.
	No difference in the pain symptoms is assumed for patients who either received prior surgical or medical treatment and patients who did not receive prior treatment.	Pain symptoms are assumed to be of similar severity across patients irrespective of treatment history. This reflects the patient population that may be eligible for the treatment with Relugolix CT i.e., in adult women of reproductive age with a history of previous medical or surgical treatment for their endometriosis.
	Patients are assumed to not become pregnant while being on treatment.	Patients are assumed to not experience pregnancy while being on treatment. Pregnancy is generally considered a contraindication for the use

© Gedeon Richter UK Ltd (2023). All rights reserved

		of GnRH agonists. In the case of the SPIRIT 1&2 trials, pregnant patients were excluded at baseline.
Mortality rate	No increased mortality linked to endometriosis is considered. Increased mortality due to surgery, however, is applied.	Endometriosis is assumed to be a disease without impact on mortality. However, surgical treatment may be linked to an increased mortality risk which is accounted for in the base case analysis.
Change in BMD	No change in BMD is assumed for any of the treatments included in the analysis.	No change in BMD from baseline is assumed. Relugolix CT is a combination treatment that prevents potential changes in BMD. For GnRH agonists, it is assumed that 100% of patients receive add-back therapy (see below). Hence, no change in BMD is assumed. BSC does not impact BMD, hence, no change in BMD is assumed. For surgery, if relevant, patients are assumed to receive add-back therapy which prevents changes in BMD.

© Gedeon Richter UK Ltd (2023). All rights reserved

B.3.10 Base-case results

Base-case incremental cost-effectiveness analysis results

The base-case cost-effectiveness model results are presented in

Table 88. Relugolix CT is more effective compared to GnRH agonists, with an incremental QALY gain of 0.71 QALYs. Relugolix CT is also associated with a very small increase in total costs compared to GnRH agonists, with an incremental difference of only £1,182. This results in an ICER of £1,670 per QALY which lies considerably below UK cost-effectiveness thresholds of £20,000 to £30,000 per QALY.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 88: Base-case results

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)	ICER incremental (£/QALY)
Relugolix CT	£11,473	11.80	9.75	-	-	-	-	-
GnRH agonist	£10,291	11.54	9.05	£1,182	0.26	0.71	£1,670	£1,670

Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years

© Gedeon Richter UK Ltd (2023). All rights reserved

B.3.11 Exploring uncertainty

A range of sensitivity analyses were conducted to explore the underlying uncertainty in the base-case cost-effectiveness results. These include deterministic and probabilistic sensitivity analyses and are detailed below.

Probabilistic sensitivity analysis

A probabilistic sensitivity analysis (PSA) was conducted to explore the uncertainty around key model parameters. PSA was conducted by varying these parameters using their upper and lower bound values and a distribution was assigned to these parameters. 1,000 simulations were run for the probabilistic sensitivity analysis (PSA), by which time the ICERs had converged to a stable mean, represented by the probabilistic ICERs. The probabilistic ICER (£1,677) lies very closely to the base-case ICER (£1,670) indicating that the cost-effectiveness results are robust.

Output from the PSA iterations is presented as scatter points on the cost-effectiveness plane in

© Gedeon Richter UK Ltd (2023). All rights reserved

Figure 40. All scatter points, which represent the simulated incremental costs and QALYs, are in the northeast quadrant. This indicates that Relugolix CT is associated with both higher costs and QALYs compared to GnRH agonist, that is, Relugolix CT is both more effective and more costly. Overall, the variation in incremental costs and QALYs is limited, indicating little impact of parameter uncertainty on the results and that the analysis is robust.

The PSA results were also plotted in the form of a cost-effectiveness acceptability curve (CEAC), as shown in Figure 41. The CEAC shows the probability of cost effectiveness for Relugolix CT and GnRH agonist given varying willingness to pay thresholds for a QALY. According to the CEAC, the probability of Relugolix CT being cost-effective is 50% at a willingness to pay of £1,600/QALY. The probability is close to 100% at £5,000/QALY.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 89: Probabilistic cost-effectiveness results

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental. costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)	Incremental ICER (£/QALY)
Relugolix CT	£11,440	11.79	9.75	-	-	-	-	-
GnRH agonist	£10,258	11.53	9.04	£1,182	0.26	0.70	£1,677	£1,677

© Gedeon Richter UK Ltd (2023). All rights reserved

Figure 40: PSA output on the cost-effectiveness plane, Relugolix CT vs. GnRH agonist

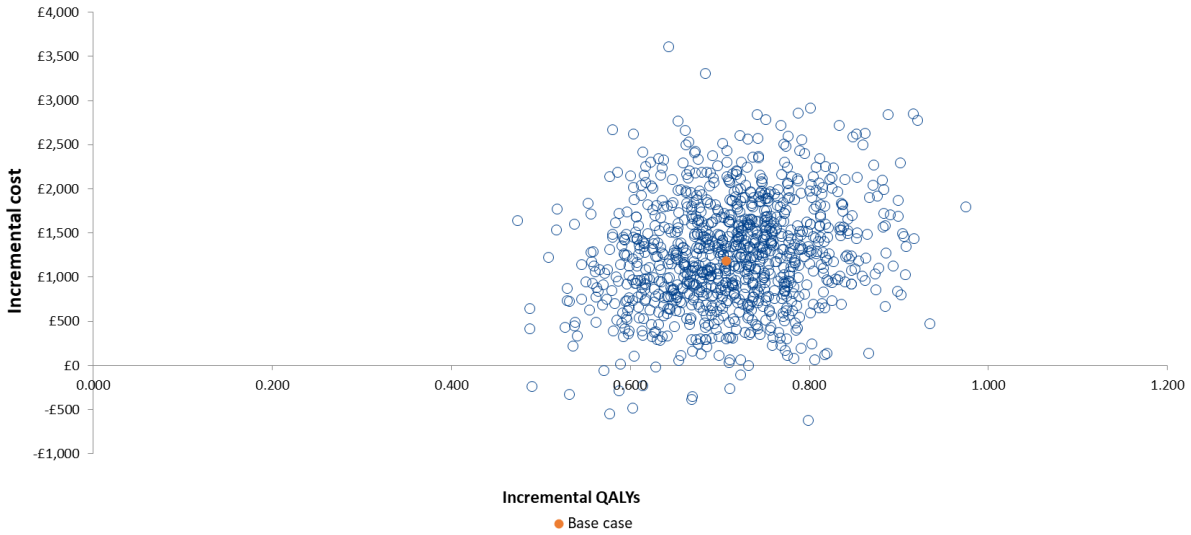


Figure 41: Cost-effectiveness plane – Relugolix CT vs. GnRH agonist

Deterministic sensitivity analysis

One-way deterministic sensitivity analyses (OWSA) were conducted to examine the sensitivity of the model result to lower and upper estimates for parameter values. The results from the OWSA are presented in the form of a tornado diagram where the ten parameters with the largest influence on the ICER are presented Figure 42.

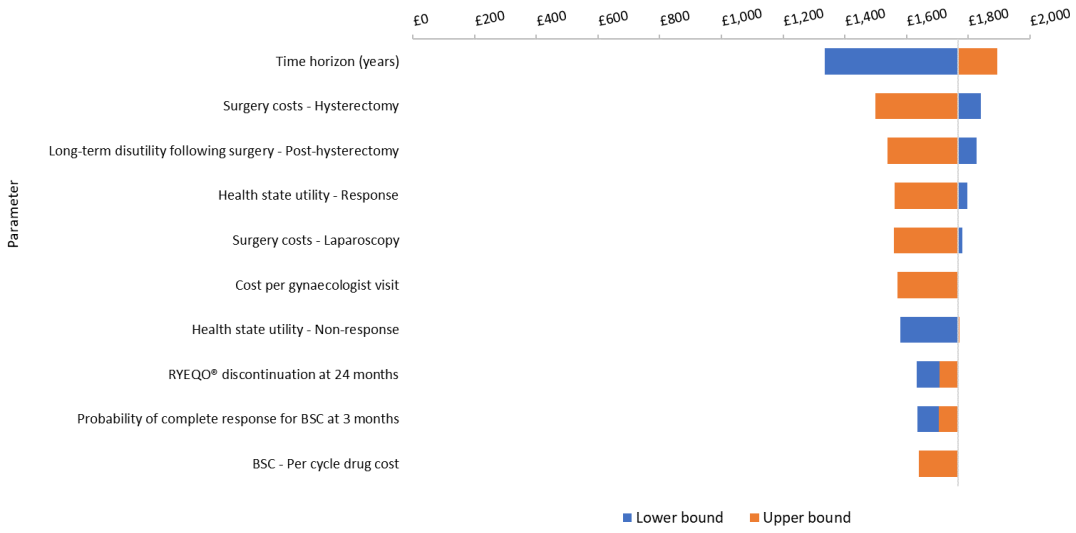
Selected parameters were varied by plus or minus 10% of the base case value except for the annual discount rate (benefits and costs) which was set to 1.5% and 6% respectively according to recommendations from NICE guidelines. Influence on

© Gedeon Richter UK Ltd (2023). All rights reserved

the ICER was defined as the absolute difference between the upper bound (base case +10%) and the lower bound (base case -10%). Parameters that could not be varied without compromising the integrity of the Markov model were excluded from the OWSA. These included the distribution of subsequent treatment strategy following discontinuation of medical therapy or recurrence of pain. Binary variables (definition of treatment response and stopping rule) were also excluded from the OWSA.

The tornado diagram below shows that the analysis time horizon has the largest impact on the ICER comparing Relugolix CT with GnRH agonist. Other parameters that have a large impact are linked to surgery, namely the costs for hysterectomy and the long-term disutility following hysterectomy. Overall, the results appear robust and none of the parameters have a considerable impact upon the ICER, as it does not exceed £2,000 per QALY in any of the OWSAs.

Figure 42: Tornado diagram – Relugolix CT vs. GnRH agonist



© Gedeon Richter UK Ltd (2023). All rights reserved

Scenario analysis

The sensitivity of the model results to changes in key assumptions or parameters underpinning the model base-case was examined through several scenario analyses. The scenarios analyses results are presented below, with pairwise ICERs presented for Relugolix CT vs. GnRH agonists. The ICERs estimated in each of the scenario analyses lie closely to the base-case ICERs, as they are typically in the range of £1,600 to £1,750 per QALY, compared to the base-case ICER of £1,670 per QALY against GnRH agonist. None of the scenarios resulted in ICERs above £2,000 per QALY. The scenario that had the largest impact upon the ICERs was the adopting a societal perspective for the analysis. In this scenario the incremental costs for Relugolix CT vs. GnRH agonist reduced from £1,182 in the model base-case to £101, leading to a much smaller ICER of £143 per QALY.

The scenario that had the second-largest impact upon the ICER was increasing GnRH agonist treatment duration to 24 months, as opposed to 12 months in the base-case. In this scenario, the incremental costs reduced slightly (£1,182 vs. £803) due to an increase in treatment acquisition costs in the GnRH agonist arm. The incremental QALYs also reduced slightly (0.71 to 0.62) due to responders on GnRH agonist being able to remain on treatment for a longer duration. This led to a lower ICER of £1,288 compared to £1,670 in the base-case.

© Gedeon Richter UK Ltd (2023). All rights reserved

Table 90: Results of scenario analyses

Structural assumption	Base-case scenario	Other scenarios considered	Incremental costs	Incremental QALYs	ICER vs. relugolix CT
Base-case			**£1,182**	**0.71**	**£1,670**
Definition of response	Change from baseline: NRS score reduction from baseline of both 2.8 for dysmenorrhea and 2.1 for NMPP and no increase of analgesic use	Threshold: Achieving or maintaining a threshold below 4 in NRS scale (mild pain) for both NMPP and dysmenorrhea and no increase of analgesic use	£1,315	0.76	£1,742
Timepoint for evaluation of complete response	6 months	3 months	£778	0.48	£1,622
Duration of GnRH agonist treatment	12 months	6 months	£1,270	0.73	£1,739
		24 months	£803	0.62	£1,288
GnRH agonist and HRT dose intensity	100%	50%	£1,205	0.71	£1,703
Waiting time for surgery	6 months	12 months	£1,210	0.71	£1,711
Perspective for analysis	Payer	Societal	£101	0.71	£143

© Gedeon Richter UK Ltd (2023). All rights reserved

B.3.12 Subgroup analysis

No subgroup analyses were included.

B.3.13 Benefits not captured in the QALY calculation

We believe that all benefits associated with Relugolix CT are captured within the QALY calculation.

B.3.14 Validation

Validation of cost-effectiveness analysis

The model has undergone thorough internal validation. The model was developed internally by a team of health economists. The structure and clinical assumptions of the model were discussed and ratified as part of an advisory board which included UK clinical experts and industry representatives. In addition to the advisory board, KOL engagement was enhanced with primary research interviews with consultant gynaecologists where the model assumptions, particularly those pertaining to HRU were discussed in more detail before finalisation. Feedback was also elicited from a sample of 5 KOLs via email. All feedback and external ratification went into the final model and this written submission.

B.3.15 Interpretation and conclusions of economic evidence

Relugolix CT is a highly cost-effective treatment when considering the NICE cost-effectiveness threshold of £20,000 to £30,000 per QALY. Relugolix CT is associated with increased QALYs at a very small increase in costs. Furthermore, Relugolix CT offers an important treatment to patients without a limit on the maximum duration of treatment, unlike the comparator in this population, GnRH agonists.

© Gedeon Richter UK Ltd (2023). All rights reserved

B.4 References

1. Dunselman GA, Vermeulen N, Becker C, Calhaz-Jorge C, D'Hooghe T, De Bie B, et al. ESHRE guideline: management of women with endometriosis. Human reproduction (Oxford, England). 2022;29(3).

2. Johnson NP, Hummelshoj L. Consensus on current management of endometriosis. Human reproduction (Oxford, England). 2013;28(6).

3. Johnson NP, Hummelshoj L, Adamson GD, Keckstein J, Taylor HS, Abrao MS, et al. World Endometriosis Society consensus on the classification of endometriosis. Human reproduction (Oxford, England). 2017;32(2).

4. Practice Committee of the American Society for Reproductive Medicine, editor Treatment of pelvic pain associated with endometriosis: a committee opinion. Fertility and sterility; 2014 2014 Apr: Fertil Steril.

5. Zondervan KT, Becker CM, Missmer SA. Endometriosis. The New England journal of medicine. 2020;382(13).

6. Goldstein DP, deCholnoky C, Emans SJ, M. LJ. Laparoscopy in the diagnosis and management of pelvic pain in adolescents. The Journal of reproductive medicine. 1980;24(6).

7. National Institute for Health and Care Excellence. Endometriosis: diagnosis and management guidance NICE: NICE; 2017 [Available from: https://www.nice.org.uk/guidance/ng73/chapter/Context.

8. Parasar P, Ozcan P, Terry K. Endometriosis: Epidemiology, Diagnosis and Clinical Management. Current obstetrics and gynecology reports. 2017;6(1).

9. National Institute for Health and Care Excellence. Clinical Knowledge Base - Endometriosis: How common is it? : NICE; 2020 [Available from: https://cks.nice.org.uk/topics/endometriosis/background-information/prevalence/.

10. European Society of Human Reproduction and Embryology. Information on Endometriosis - Patient Leaflet Based on the ESHRE Guideline on Endometriosis: European Society of Human Reproduction and Embryology; 2022 [Available from: https://www.eshre.eu/Guideline/Endometriosis.

11. Hickey M, Ballard K, Farquhar C. Endometriosis. BMJ (Clinical research ed). 2014;348.

12. Endometriosis UK. THE NATIONAL ENDOMETRIOSIS SOCIETY INFORMATION - Understanding-endometriosis: Endometriosis UK; 2012 [Available from: http://www.endometriosis-uk.org.

13. Ye L, Whitaker LHR, Mawson RL, Hickey M. Endometriosis. BMJ. 2022;379:e068950.

© Gedeon Richter UK Ltd (2023). All rights reserved

14. Coleman L, Overton C. GPs have key role in early diagnosis of endometriosis. The Practitioner. 2015;259(1780).

15. Smolarz B, Szyłło K, Romanowicz H. Endometriosis: Epidemiology, Classification, Pathogenesis, Treatment and Genetics (Review of Literature). International journal of molecular sciences. 2021;22(19).

16. Eskenazi B, Warner M. Epidemiology of endometriosis. Obstetrics and gynecology clinics of North America. 1997;24(2).

17. Kristensen J, Lebovic D, Mueller M, Vigano P, Wullschleger M, D'Hooghe T, et al. The burden of endometriosis: costs and quality of life of women with endometriosis and treated in referral centres. Human reproduction (Oxford, England). 2012;27(5).

18. All Party Parliamentary Group on Endometriosis. Endometriosis in the UK: Time for Change. House of Commons; 2020.

19. Nnoaham KE, Hummelshoj L, Webster P, d'Hooghe T, de Cicco Nardone F, de Cicco Nardone C, et al. Impact of endometriosis on quality of life and work productivity: a multicenter study across ten countries. Fertility and sterility. 2011;96(2).

20. Ghai V, Jan H, Shakir F, Haines P, Kent. Diagnostic delay for superficial and deep endometriosis in the United Kingdom. Journal of obstetrics and gynaecology : the journal of the Institute of Obstetrics and Gynaecology. 2020;40(1).

21. de Sepulcri R, do Amaral VF. Depressive symptoms, anxiety, and quality of life in women with pelvic endometriosis. European journal of obstetrics, gynecology, and reproductive biology. 2009;142(1).

22. Petrelluzzi KF, Garcia MC, Petta CA, Grassi-Kassisse DM, Spadari-Bratfisch R. Salivary cortisol concentrations, stress and quality of life in women with endometriosis and chronic pelvic pain. Stress (Amsterdam, Netherlands). 2008;11(5).

23. Fourquet J, Báez L, Figueroa M, Iriarte RI, Flores I. Quantification of the impact of endometriosis symptoms on health-related quality of life and work productivity. Fertility and sterility. 2011;96(1).

24. Tripoli TM, Sato H, Sartori MG, de Araujo FF, Girão MJ, Schor E. Evaluation of quality of life and sexual satisfaction in women suffering from chronic pelvic pain with or without endometriosis. The journal of sexual medicine. 2011;8(2).

25. Rush G, Misajon R. Examining subjective wellbeing and health-related quality of life in women with endometriosis. Health care for women international. 2018;39(3).

26. Verket NJ, Uhlig T, Sandvik L, Andersen MH, Tanbo TG, Qvigstad E. Health-related quality of life in women with endometriosis, compared with the general population and women with rheumatoid arthritis. Acta obstetricia et gynecologica Scandinavica. 2018;97(11).

© Gedeon Richter UK Ltd (2023). All rights reserved

27. Jia SZ, Leng JH, Shi JH, Sun PR, Lang J. Health-related quality of life in women with endometriosis: a systematic review. Journal of ovarian research. 2012;5(1).

28. Abbott JA, Hawe J, Clayton RD, Garry R. The effects and effectiveness of laparoscopic excision of endometriosis: a prospective study with 2-5 year follow-up. Human reproduction (Oxford, England). 2003;18(9).

29. Becker C, Bokor A, Heikinheimo O, Horne A, Jansen F, Kiesel L, et al. ESHRE guideline: endometriosis. Human Reproduction Open. 2022:1-26.

30. Horne AW, Vincent K, Hewitt CA, Middleton LJ, Koscielniak M, Szubert W, et al. Gabapentin for chronic pelvic pain in women (GaPP2): a multicentre, randomised, double-blind, placebo-controlled trial. Lancet (London, England). 2020;396(10255).

31. National Institute for Health and Care Excellence. Leuprorelin acetate | Drugs | BNF content published by NICE: NICE; 2023 [Available from: https://bnf.nice.org.uk/drugs/leuprorelin-acetate/.

32. National Institute for Health and Care Excellence. Goserelin | Drugs | BNF content published by NICE: NICE; 2023 [Available from: https://bnf.nice.org.uk/drugs/goserelin/.

33. National Institute for Health and Care Excellence. Triptorelin | Drugs | BNF content published by NICE: NICE; 2023 [Available from: https://bnf.nice.org.uk/drugs/triptorelin/.

34. Tang H, Wu R, Li X, Zhou Y, Liu Z, Wang C, et al. Curative effect of 1.88-mg and 3.75-mg gonadotrophin-releasing hormone agonist on stage III-IV endometriosis: Randomized controlled study. The journal of obstetrics and gynaecology research. 2017;43(10).

35. Brown J, Pan A, Hart R. Gonadotrophin-releasing hormone analogues for pain associated with endometriosis. The Cochrane database of systematic reviews. 2010;2010(12).

36. Wu D, Hu M, Hong L, Hong S, Ding W, Min J, et al. Clinical efficacy of add-back therapy in treatment of endometriosis: a meta-analysis. Archives of gynecology and obstetrics. 2014;290(3).

37. Giudice LC, As-Sanie S, Arjona Ferreira JC, Becker CM, Abrao MS, Lessey BA, et al. Once daily oral relugolix combination therapy versus placebo in patients with endometriosis-associated pain: two replicate phase 3, randomised, double-blind, studies (SPIRIT 1 and 2). Lancet (London, England). 2022;399(10343).

38. Bafort C, Beebeejaun Y, Tomassetti C, Bosteels J, Duffy J. Laparoscopic surgery for endometriosis. The Cochrane database of systematic reviews. 2020;10(10).

39. Sutton CJ, Ewen SP, Whitelaw N, Haines P. Prospective, randomized, double-blind, controlled trial of laser laparoscopy in the treatment of pelvic pain

© Gedeon Richter UK Ltd (2023). All rights reserved

associated with minimal, mild, and moderate endometriosis. Fertility and sterility. 1994;62(4).

40. Jarrell J, Mohindra R, Ross S, Taenzer P, Brant R. Laparoscopy and reported pain among patients with endometriosis. Journal of obstetrics and gynaecology Canada : JOGC = Journal d'obstetrique et gynecologie du Canada : JOGC. 2005;27(5).

41. Arcoverde FVL, Andres MP, Borrelli GM, Barbosa PA, Abrão MS, Kho R. Surgery for Endometriosis Improves Major Domains of Quality of Life: A Systematic Review and Meta-Analysis. Journal of minimally invasive gynecology. 2019;26(2).

42. Franck C, Poulsen MH, Karampas G, Giraldi A, Rudnicki M. Questionnaire-based evaluation of sexual life after laparoscopic surgery for endometriosis: a systematic review of prospective studies. Acta obstetricia et gynecologica Scandinavica. 2018;97(9).

43. Chapron C, Bourret A, Chopin N, Dousset B, Leconte M, Amsellem-Ouazana D, et al. Surgery for bladder endometriosis: long-term results and concomitant management of associated posterior deep lesions. Human reproduction (Oxford, England). 2010;25(4).

44. Byrne D, Curnow T, Smith P, Cutner A, Saridogan E, Clark TJ. Laparoscopic excision of deep rectovaginal endometriosis in BSGE endometriosis centres: a multicentre prospective cohort study. BMJ open. 2018;8(4).

45. Soriano LC, López-Garcia E, Schulze-Rath R, Rodríguez LAG. Incidence, treatment and recurrence of endometriosis in a UK-based population analysis using data from The Health Improvement Network and the Hospital Episode Statistics database. The European journal of contraception & reproductive health care : the official journal of the European Society of Contraception. 2017;22(5).

46. Guo S-W. Recurrence of endometriosis and its control. Human reproduction update. 2009;15(4).

47. Vercellini P, Crosignani PG, Abbiati A, Somigliana E, Viganò P, Fedele L. The effect of surgery for symptomatic endometriosis: the other side of the story. Human reproduction update. 2009;15(2).

48. Domingo S, Pellicer A. Overview of current trends in hysterectomy. Expert Review of Obstetrics & Gynecology. 2009;4(6):673-85.

49. Aarts JWM, Nieboer TE, Johnson N, Tavender E, Garry R, Mol BWJ, et al. Surgical approach to hysterectomy for benign gynaecological disease. The Cochrane database of systematic reviews. 2015;2015(8).

50. British Fibroid Trust. Hysterectomy 2023 [Available from: http://www.britishfibroidtrust.org.uk/Fib_info/hysterectomy.php.

© Gedeon Richter UK Ltd (2023). All rights reserved

51. As-Sanie S, Black R, Giudice LC, Gray Valbrun T, Gupta J, Jones B, et al. Assessing research gaps and unmet needs in endometriosis. American journal of obstetrics and gynecology. 2019;221(2).

52. Becker CM, Gattrell WT, Gude K, Singh S. Reevaluating response and failure of medical treatment of endometriosis: a systematic review. Fertility and sterility. 2017;108(1).

53. Giudice L. Clinical practice. Endometriosis. The New England journal of medicine. 2010;362(25).

54. Shakiba K, Bena JF, McGill KM, Minger J, Falcone F. Surgical treatment of endometriosis: a 7-year follow-up on the requirement for further surgery. Obstetrics and gynecology. 2008;111(6).

55. Abbott J, Hawe J, Hunter D, Holmes M, Finn P, Garry R. Laparoscopic excision of endometriosis: a randomized, placebo-controlled trial. Fertility and sterility. 2004;82(4).

56. Hudson. The missed disease? Endometriosis as an example of 'undone science'. Reproductive biomedicine & society online. 2021;14.

57. Bougie O, Yap MI, Sikora L, Flaxman T, Singh S. Influence of race/ethnicity on prevalence and presentation of endometriosis: a systematic review and meta-analysis. BJOG : an international journal of obstetrics and gynaecology. 2019;126(9).

58. Becker C, As-Sanie S, Abrao MS, Brown E, Arjona Ferreira JC, Wagman RB, et al. SPIRIT long-term extension study: two-year efficacy and safety of relugolix combination therapy in women with endometriosis-associated pain. Human Reproduction. 2022;37(Supplement_1):105-2.

59. D'Hooghe T, Fukaya T, Osuga Y, Besuyen R, López B, Holtkamp GM, et al. Efficacy and safety of ASP1707 for endometriosis-associated pelvic pain: the phase II randomized controlled TERRA study. Human reproduction (Oxford, England). 2019;34(5).

60. Lang J, Yu Q, Zhang S, Li H, Gude K, von Ludwig C, et al. Dienogest for Treatment of Endometriosis in Chinese Women: A Placebo-Controlled, Randomized, Double-Blind Phase 3 Study. Journal of women's health (2002). 2018;27(2).

61. Strowitzki T, Marr J, Gerlinger C, Faustmann T, Seitz C. Dienogest is as effective as leuprolide acetate in treating the painful symptoms of endometriosis: a 24-week, randomized, multicentre, open-label trial. Human reproduction (Oxford, England). 2010;25(3).

62. Myovant Sciences GmbH. SPIRIT 1 (MVT-601-3101) Clinical Study Report. Data on file. 2021.

63. Myovant Sciences GmbH. SPIRIT 2 (MVT-601-3102) Clinical Study Report. Data on file. 2021.

© Gedeon Richter UK Ltd (2023). All rights reserved

64. Myovant Sciences GmbH. SPIRIT OLE (MVT-601-3103) Clinical Study Report. Data on file. 2022.

65. ClinicalTrials.gov. SPIRIT EXTENSION: Efficacy and Safety Extension Study of Relugolix in Women With Endometriosis-Associated Pain - Full Text View - ClinicalTrials.gov: ClinicalTrials.gov; 2023 [Available from: https://clinicaltrials.gov/ct2/show/NCT03654274.

66. Gedeon Richter. Data on file: Global Advisory board report. 2023.

67. Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. Medical decision making : an international journal of the Society for Medical Decision Making. 2013;33(5).

68. Bourdel N, Alves J, Pickering G, Ramilo I, Roman H, Canis M. Systematic review of endometriosis pain assessment: how to choose a scale? Human reproduction update. 2015;21(1).

69. Taylor HS, Giudice LC, Lessey BA, Abrao MS, Kotarski J, Archer DF, et al. Treatment of Endometriosis-Associated Pain with Elagolix, an Oral GnRH Antagonist. The New England journal of medicine. 2017;377(1).

70. Berger C, Goltzman D, Langsetmo L, Joseph L, Jackson S, Kreiger N, et al. Peak bone mass from longitudinal data: implications for the prevalence, pathophysiology, and diagnosis of osteoporosis. J Bone Miner Res. 2010;25(9):1948-57.

71. Looker AC, Sarafrazi Isfahani N, Fan B, Shepherd JA. FRAX-based Estimates of 10-year Probability of Hip and Major Osteoporotic Fracture Among Adults Aged 40 and Over: United States, 2013 and 2014. National health statistics reports. 2017(103).

72. Matta WH, Shaw RW, Hesp R, Evans R. Reversible trabecular bone density loss following induced hypo-oestrogenism with the GnRH analogue buserelin in premenopausal women. Clinical endocrinology. 1988;29(1).

73. Dawood MY, Lewis V, Ramos J. Cortical and trabecular bone mineral content in women with endometriosis: effect of gonadotropin-releasing hormone agonist and danazol. Fertility and sterility. 1989;52(1).

74. Orwoll ES, Yuzpe AA, Burry KA, Heinrichs L, Buttram VC, Hornstein MD. Nafarelin therapy in endometriosis: long-term effects on bone mineral density. American journal of obstetrics and gynecology. 1994;171(5).

75. Caird LE, West CP, Lumsden MA, Hannan WJ, Gow SM. Medroxyprogesterone acetate with Zoladex for long-term treatment of fibroids: effects on bone density and patient acceptability. Human reproduction (Oxford, England). 1997;12(3).

© Gedeon Richter UK Ltd (2023). All rights reserved

76. Taga M, Minaguchi H. Reduction of bone mineral density by gonadotropin-releasing hormone agonist, nafarelin, is not completely reversible at 6 months after the cessation of administration. Acta obstetricia et gynecologica Scandinavica. 1996;75(2).

77. Pierce SJ, Gazvani MR, Farquharson RG. Long-term use of gonadotropin-releasing hormone analogs and hormone replacement therapy in the management of endometriosis: a randomized trial with a 6-year follow-up. Fertility and sterility. 2000;74(5).

78. Recker RR, Lappe JM, Davies KM, Kimmel DB. Change in bone mass immediately before menopause. Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research. 1992;7(8).

79. Riggs BL, Wahner HW, Melton LJ, Richelson LS, Judd HL, Offord KP. Rates of bone loss in the appendicular and axial skeletons of women. Evidence of substantial vertebral bone loss before menopause. The Journal of clinical investigation. 1986;77(5).

80. Looker AC, Orwoll ES, Johnston CC, Lindsay RL, Wahner HW, Dunn WL, et al. Prevalence of low femoral bone density in older U.S. adults from NHANES III. Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research. 1997;12(11).

81. Gedeon Richter. Data on file: Clinical expert opinion - generalisability. 2023.

82. Gedeon Richter. Data on file: Clinical expert survey responses. 2023.

83. Bohn JA, Bullard KA, Rodriguez MI, Ecker AM. Stepwise Approach to the Management of Endometriosis-Related Dysmenorrhea: A Cost-Effectiveness Analysis. Obstetrics and gynecology. 2021;138(4).

84. Arakawa I, Momoeda M, Osuga Y, Ota I, Koga K. Cost-effectiveness of the recommended medical intervention for the treatment of dysmenorrhea and endometriosis in Japan. Cost Effectiveness and Resource Allocation. 2018;16(1):1-10.

85. Wu B, Yang Z, Tobe RG, Wang Y. Medical therapy for preventing recurrent endometriosis after conservative surgery: a cost-effectiveness analysis - Google Search 2018 [Available from: https://www.google.com/search?q=Medical+therapy+for+preventing+recurrent+endometriosis+after+conservative+surgery%3A+a+cost-effectiveness+analysis&rlz=1C1ONGR_enGB1025GB1025&oq=Medical+therapy+for+preventing+recurrent+endometriosis+after+conservative+surgery%3A+a+cost-effectiveness+analysis&gs_lcrp=EgZjaHJvbWUyBggAEEUYOdIBBzUyNmowajSoAgCwAgA&sourceid=chrome&ie=UTF-8.

86. Dai Y, Shi B, Huang X, Duan J, Qiu Y, Ha C, et al. Cost-effectiveness Analysis of Dienogest Compared With Gonadotropin-Releasing Hormone Agonist

© Gedeon Richter UK Ltd (2023). All rights reserved

After Conservative Surgery for Endometriosis in China. Clinical Therapeutics. 2021;43(8):1276-84.e1.

87. Grand TS, Basarir H, Jackson LJ. The cost-effectiveness of oral contraceptives compared to 'no hormonal treatment' for endometriosis-related pain: An economic evaluation. PloS one. 2019;14(1).

88. Institute for Clinical and Economic Review. Final Evidence Report – Elagolix for Treating Endometriosis. 2018.

89. Leonardi M, Martin E, Reid S, Blanchette G, Condous G. Deep endometriosis transvaginal ultrasound in the workup of patients with signs and symptoms of endometriosis: a cost analysis. BJOG : an international journal of obstetrics and gynaecology. 2019;126(12).

90. Sanghera S, Barton P, Bhattacharya S, Horne AW, Roberts TE. Pharmaceutical treatments to prevent recurrence of endometriosis following surgery: a model-based economic evaluation. 2016.

91. Wang ST, Johnson SJ, Mitchell D, Soliman AM, Vora JB, Agarwal SK. Cost-effectiveness of elagolix versus leuprolide acetate for treating moderate-to-severe endometriosis pain in the USA. Journal of comparative effectiveness research. 2019;8(5).

92. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. Med Decis Making. 1993;13(4):322-38.

93. NHS England. Menopause 2023 [Available from: https://www.nhs.uk/conditions/menopause/.

94. Hernández Alava M, Pudney S, Wailoo A. Estimating the Relationship Between EQ-5D-5L and EQ-5D-3L: Results from a UK Population Study. Pharmacoeconomics. 2023;41(2):199-207.

95. National Institute for Health and Care Excellence. NICE health technology evaluations: the manual. 2022.

96. Dolan P. Modeling valuations for EuroQol health states. Med Care. 1997;35(11):1095-108.

97. Crandall CJ, Larson J, Wright NC, Laddu D, Stefanick ML, Kaunitz AM, et al. Serial Bone Density Measurement and Incident Fracture Risk Discrimination in Postmenopausal Women. JAMA Intern Med. 2020;180(9):1232-40.

98. British National Formulary. Raloxifene hydrochloride. 2023.

99. NHS England. 2021/22 National Cost Collection Data Publication 2023 [Available from: https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/.

© Gedeon Richter UK Ltd (2023). All rights reserved

100. Rafique S, Decherney AH. Medical Management of Endometriosis. Clin Obstet Gynecol. 2017;60(3):485-96.

101. Soliman AM, Bonafede M, Farr AM, Castelli-Haley J, Winkel C. Analysis of Adherence, Persistence, and Surgery Among Endometriosis Patients Treated with Leuprolide Acetate Plus Norethindrone Acetate Add-Back Therapy. J Manag Care Spec Pharm. 2016;22(5):573-87.

102. Maresh MJ, Metcalfe MA, McPherson K, Overton C, Hall V, Hargreaves J, et al. The VALUE national hysterectomy study: description of the patients and their surgery. Bjog. 2002;109(3):302-12.

103. Chapron C, Querleu D, Bruhat MA, Madelenat P, Fernandez H, Pierre F, et al. Surgical complications of diagnostic and operative gynaecological laparoscopy: a series of 29,966 cases. Hum Reprod. 1998;13(4):867-72.

104. Soliman AM, Du EX, Yang H, Wu EQ, Haley JC. Retreatment Rates Among Endometriosis Patients Undergoing Hysterectomy or Laparoscopy. J Womens Health (Larchmt). 2017;26(6):644-54.

105. Veth VB, van de Kar MMA, Duffy JMN, van Wely M, Mijatovic V, Maas JWM. Gonadotropin-releasing hormone analogues for endometriosis. Cochrane Database of Systematic Reviews. 2023(6).

106. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol. 1997;50(6):683-91.

107. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. Stat Med. 2010;29(7-8):932-44.

108. D'Agostino RB, Sr., Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation. 2008;117(6):743-53.

109. National Institutes of Health Clinical Center. Effects of Endometriosis on Bone Mineral Density. 2006.

110. Sauerbrun-Cutler M-T, Alvero R. Short- and long-term impact of gonadotropin-releasing hormone analogue treatment on bone loss and fracture. Fertility and Sterility. 2019;112:799-803.

111. Crosignani PG, Luciano A, Ray A, Bergqvist A. Subcutaneous depot medroxyprogesterone acetate versus leuprolide acetate in the treatment of endometriosis-associated pain. Hum Reprod. 2006;21(1):248-56.

112. Singer BR, McLauchlan GJ, Robinson CM, Christie J. Epidemiology of fractures in 15,000 adults: the influence of age and gender. J Bone Joint Surg Br. 1998;80(2):243-8.

© Gedeon Richter UK Ltd (2023). All rights reserved

113. 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.

114. Armstrong N, Vale L, Deverill M, Nabi G, McClinton S, N'Dow J, et al. Surgical treatments for men with benign prostatic enlargement: cost effectiveness study. BMJ. 2009;338:b1288.

115. Boyd KA, Jones RJ, Paul J, Birrell F, Briggs AH, Leung HY. Decision analytic cost-effectiveness model to compare prostate cryotherapy to androgen deprivation therapy for treatment of radiation recurrent prostate cancer. BMJ Open. 2015;5(10):e007925.

116. Hux M, Ng C, Ortega GL, Ferrazzi S, Goeree R. Utility values for pre-menopausal women suffering from symptomatic uterine fibroids. Expert Rev Pharmacoecon Outcomes Res. 2015;15(1):181-9.

117. Xu R, Insinga RP, Golden W, Hu XH. EuroQol (EQ-5D) health utility scores for patients with migraine. Qual Life Res. 2011;20(4):601-8.

118. Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins J. Health state utilities for non small cell lung cancer. Health Qual Life Outcomes. 2008;6:84.

119. 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.

120. Sculpher M, Manca A, Abbott J, Fountain J, Mason S, Garry R. Cost effectiveness analysis of laparoscopic hysterectomy compared with standard hysterectomy: results from a randomised trial. BMJ. 2004;328(7432):134.

121. World Health Organization (WHO). Global Burden of Disease 2004 update: Disability weights for diseases and conditions 2004 [

122. Earnshaw SR, Klok RM, Iyer S, McDade C. Methylnaltrexone bromide for the treatment of opioid-induced constipation in patients with advanced illness--a cost-effectiveness analysis. Aliment Pharmacol Ther. 2010;31(8):911-21.

123. British National Formulary. Myfembree 2023 [Available from: https://bnf.nice.org.uk/drugs/relugolix-with-estradiol-and-norethisterone-acetate/.

124. British National Formulary. Ibuprofen 2023 [Available from: https://bnf.nice.org.uk/drugs/ibuprofen/.

125. British National Formulary. Ethinylestradiol with levonorgestrel 2023 [Available from: https://bnf.nice.org.uk/drugs/ethinylestradiol-with-levonorgestrel/.

126. British National Formulary. Medroxyprogesterone acetate 2023 [Available from: https://bnf.nice.org.uk/drugs/medroxyprogesterone-acetate/.

© Gedeon Richter UK Ltd (2023). All rights reserved

127. Electronic Medicines Compendium. Benilexa One Handed 20 micrograms/24 hours Intrauterine Delivery System, Summary of Product Characteristics 2023 [Available from: https://www.medicines.org.uk/emc/product/13055/smpc.

128. Primary Care Women's Health Forum. Gedeon Richter receives licence for the first and only reloadable, single-handed intrauterine delivery system (IUS) for contraception in the UK 2021 [Available from: https://pcwhf.co.uk/news/gedeon-richter-receives-licence-for-the-first-and-only-reloadable-single-handed-intrauterine-delivery-system-ius-for-contraception-in-the-uk/#.

129. British National Formulary. Triptorelin 2023 [Available from: https://bnf.nice.org.uk/drugs/triptorelin/medicinal-forms/.

130. EndoAct Canada. Discussion: The need for action on endometriosis in Canada 2022 [Available from: https://endoact.ca/wp-content/uploads/2022/02/EndoActCanadaBackgroundMaterial.pdf.

131. British National Formulary. Dienogest 2023 [Available from: https://bnf.nice.org.uk/drugs/dienogest/.

132. British National Formulary. Tibolone 2023 [Available from: https://bnf.nice.org.uk/drugs/tibolone/.

133. NHS England. 2020/21 National Cost Collection Data Publication 2022 [Available from: https://www.england.nhs.uk/publication/2020-21-national-cost-collection-data-publication/.

134. Jones K, H. W, Birch S, Castelli A, Chalkley M, Dargan A, et al. Unit Costs of Health and Social Care 2022 2023 [Available from: https://kar.kent.ac.uk/100519/1/Unit_Costs_of_Health_and_Social_Care_2022%20%287%29.pdf.

135. NHS England. 2022/23 national tariff workbook (Annex A) 2022 [Available from: https://www.england.nhs.uk/publication/national-tariff-payment-system-documents-annexes-and-supporting-documents/.

136. Akhtar W, Chung Y. Saving the NHS one blood test at a time. BMJ Qual Improv Rep. 2014;2(2).

137. Menon U, McGuire AJ, Raikou M, Ryan A, Davies SK, Burnell M, et al. The cost-effectiveness of screening for ovarian cancer: results from the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS). Br J Cancer. 2017;117(5):619-27.

138. Office for National Statistics. Average household income, UK: financial year ending 2022 2023 [Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/bulletins/householddisposableincomeandinequality/financialyearending2022.

139. Vonk Noordegraaf A, Anema JR, Louwerse MD, Heymans MW, van Mechelen W, Brölmann HA, et al. Prediction of time to return to work after

© Gedeon Richter UK Ltd (2023). All rights reserved

gynaecological surgery: a prospective cohort study in the Netherlands. Bjog. 2014;121(4):487-97.

140. As-Sanie S, Mathur V, Mehedintu C, Imm SJ, Warsi QA, Hunsche E, et al. RELUGOLIX COMBINATION THERAPY IMPROVES MULTIPLE DIMENSIONS OF QUALITY OF LIFE IN WOMEN WITH ENDOMETRIOSIS-ASSOCIATED PAIN: RESULTS FROM THE SPIRIT PROGRAM. Fertility and Sterility. 2021;116(3):e51.

141. ClinicalTrials.gov. SPIRIT 1: Efficay and safety study of relugolix in women with endometris-associated pain 2020 [Available from: https://classic.clinicaltrials.gov/ct2/show/NCT03204318.

142. ClinicalTrials.gov. SPIRIT 2: Efficay and safety of relugolix in womed with endometriosis-associated pain 2020 [Available from: https://classic.clinicaltrials.gov/ct2/show/NCT03204331.

143. As-Sanie S, Becker CM, Johnson N, Lessey BA, Abrao MS, Brown EL, et al. EFFICACY AND SAFETY OF RELUGOLIX COMBINATION THERAPY IN WOMEN WITH ENDOMETRIOSIS-ASSOCIATED PAIN: PHASE 3 RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED STUDY (SPIRIT 2). Fertility and Sterility. 2020;114(3, Supplement):e77.

144. As-Sanie S, Giudice LC, Abrao M, Wilk K, Mehedintu C, Becker CM, et al. Sustained Efficacy and Safety of Relugolix Combination Therapy in Women with Endometriosis-Associated Pain: Spirit 52-Week Data. Journal of Minimally Invasive Gynecology. 2021;28(11, Supplement):S10-S1.

145. As-Sanie S, Mehedintu C, Mathur V, Soulban G, Hunsche E, Imm SJ, et al. SUSTAINED IMPROVEMENT IN PHYSICAL FUNCTION AND QUALITY OF LIFE IN WOMEN WITH ENDOMETRIOSIS-ASSOCIATED PAIN TREATED WITH RELUGOLIX COMBINATION THERAPY OVER 104 WEEKS: SPIRIT LONG-TERM EXTENSION STUDY. Fertility and Sterility. 2022;118(4):e1.

146. As-Sanie S, Johnson N, Lukes AS, Arjona Ferreira JC, Li J, Perry JS, et al. TIME TO MINIMAL OR NO PELVIC PAIN WITH RELUGOLIX COMBINATION THERAPY IN WOMEN WITH ENDOMETRIOSIS-ASSOCIATED PAIN: RESULTS FROM THE SPIRIT PROGRAM. Fertility and Sterility. 2022;118(4):e6-e7.

147. McClung M, Santora A, Johnson N, Cummings S, Wang F, Wagman R. Evaluation of relugolix combination therapy (REL-CT) and bone mineral density (BMD) in women with endometriosis-associated pain theough 52 weeks: SPIRIT long-term extension (LTE) study. Aging Clin Exp Res. 2022;34:S249-50.

148. Bergqvist A. A comparative study of the acceptability and effect of goserelin and nafarelin on endometriosis. Gynecol Endocrinol. 2000;14(6):425-32.

149. Granese R, Perino A, Calagna G, Saitta S, De Franciscis P, Colacurci N, et al. Gonadotrophin-releasing hormone analogue or dienogest plus estradiol valerate to prevent pain recurrence after laparoscopic surgery for endometriosis: a multi-center randomized trial. Acta Obstet Gynecol Scand. 2015;94(6):637-45.

© Gedeon Richter UK Ltd (2023). All rights reserved

150. Guzick DS, Huang LS, Broadman BA, Nealon M, Hornstein MD. Randomized trial of leuprolide versus continuous oral contraceptives in the treatment of endometriosis-associated pelvic pain. Fertil Steril. 2011;95(5):1568-73.

151. Hornstein MD, Hemmings R, Yuzpe AA, Heinrichs WL. Use of nafarelin versus placebo after reductive laparoscopic surgery for endometriosis. Fertil Steril. 1997;68(5):860-4.

152. Howell R, Edmonds DK, Dowsett M, Crook D, Lees B, Stevenson JC. Gonadotropin-releasing hormone analogue (goserelin) plus hormone replacement therapy for the treatment of endometriosis: a randomized controlled trial. Fertil Steril. 1995;64(3):474-81.

153. Kiilholma P, Tuimala R, Kivinen S, Korhonen M, Hagman E. Comparison of the gonadotropin-releasing hormone agonist goserelin acetate alone versus goserelin combined with estrogen-progestogen add-back therapy in the treatment of endometriosis. Fertil Steril. 1995;64(5):903-8.

154. Mäkäräinen L, Rönnberg L, Kauppila A. Medroxyprogesterone acetate supplementation diminishes the hypoestrogenic side effects of gonadotropin-releasing hormone agonist without changing its efficacy in endometriosis. Fertil Steril. 1996;65(1):29-34.

155. Moghissi KS, Schlaff WD, Olive DL, Skinner MA, Yin H. Goserelin acetate (Zoladex) with or without hormone replacement therapy for the treatment of endometriosis. Fertil Steril. 1998;69(6):1056-62.

156. Tahara M, Matsuoka T, Yokoi T, Tasaka K, Kurachi H, Murata Y. Treatment of endometriosis with a decreasing dosage of a gonadotropin-releasing hormone agonist (nafarelin): a pilot study with low-dose agonist therapy ("draw-back" therapy). Fertil Steril. 2000;73(4):799-804.

157. van Valkenhoef G, Lu G, de Brock B, Hillege H, Ades AE, Welton NJ. Automating network meta-analysis. Research synthesis methods. 2012;3(4).

158. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ (Clinical research ed). 2011;343.

159. Hu D, O'Connor AM, Wang C, Sargeant JM, Winder CB. How to Conduct a Bayesian Network Meta-Analysis. Front Vet Sci. 2020;7:271.

160. Turner RM, Davey J, Clarke MJ, Thompson SG, Higgins JP. Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews. Int J Epidemiol. 2012;41(3):818-27.

161. Dias S, Welton NJ, Sutton AJ, Ades AE. NICE Decision Support Unit Technical Support Document 2: a generalised linear modelling framewoprk for pairwiswe and network meta-analysis of randomised controlled trials [Internet]. London: National Institute for Health and Care Excellence (NICE); 2011.

© Gedeon Richter UK Ltd (2023). All rights reserved

162. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. Stat Sci,. 1992;7(4):457-72.

163. Harrer M, Cuijpers P, Furukawa T, Ebert D. Doing meta-analysis with R: a hands-on guide. Boca Raton, FL and London: Chapman & Hall/CRC Press; 2021. Available from: https://bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/#welcome.

164. Mbuagbaw L, Rochwerg B, Jaeschke R, Heels-Andsell D, Alhazzani W, Thabane L, et al. Approaches to interpreting and choosing the best treatments in network meta-analyses. Syst Rev. 2017;6(1):79.

165. Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. J Clin Epidemiol. 2011;64(2):163-71.

166. Murad MH, Wang Z, Chu H, Lin L. When continuous outcomes are measured using different scales: guide for meta-analysis and interpretation. Bmj. 2019;364:k4817.

167. Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane Handbook for Systematic Reviews of Interventions version 6.4 (updated August 2023). Cochrane, 2023. 2023.

168. Biberoglu KO, Behrman SJ. Dosage aspects of danazol therapy in endometriosis: short-term and long-term effectiveness. Am J Obstet Gynecol. 1981;139(6):645-54.

169. Rohatgi A. WebPlotDigitizer [Available from: https://automeris.io/WebPlotDigitizer/.

170. Jelicic Kadic A, Vucic K, Dosenovic S, Sapunar D, Puljak L. Extracting data from figures with software was faster, with higher interrater reliability than manual extraction. J Clin Epidemiol. 2016;74:119-23.

171. Daly C, Downing BC, Welton NJ. A practical guide to inconsistency checks in Bayesian network meta-analysis. 2022.

172. Dias S, Sutton AJ, Welton NJ, Ades AE. NICE Decision Support Unit Technical Support Document 3: heterogeneity: subgroups, meta-regression, bias and bias-adjustment. Heterogeneity: Subgroups, Meta-Regression, Bias And Bias-Adjustment. London: National Institute for Health and Care Excellence (NICE) Copyright © 2012 National Institute for Health and Clinical Excellence, unless otherwise stated. All rights reserved.; 2012.

173. Bluett R, Fonseca-Kelly Z, Burke C. Efficacy of surgical intervention for the treatment of pain and improvement of quality of life in patients with endometriosis. Journal of Endometriosis and Pelvic Pain Disorders. 2016;8(3):101-5.

174. Grundström H, Gerdle B, Alehagen S, Berterö C, Arendt-Nielsen L, Kjølhede P. Reduced pain thresholds and signs of sensitization in women with persistent

© Gedeon Richter UK Ltd (2023). All rights reserved

pelvic pain and suspected endometriosis. Acta Obstetricia et Gynecologica Scandinavica. 2019;98(3):327-36.

175. Oppenheimer A, Verdun S, Perot M, Du Cheyron J, Panel P, Fauconnier A. Do high-dose progestins impair sexual function in women treated for endometriosis? A prospective observational longitudinal study. Acta Obstetricia et Gynecologica Scandinavica. 2021;100(5):850-9.

176. Armour M, Lawson K, Wood A, Smith CA, Abbott J. The cost of illness and economic burden of endometriosis and chronic pelvic pain in Australia: a national online survey. PloS one. 2019;14(10):e0223316.

177. Malik A, Sinclair J, Ng CHM, Smith CA, Abbott J, Armour M. Allied health and complementary therapy usage in Australian women with chronic pelvic pain: a cross-sectional study. BMC Women's Health. 2022;22(1):1-9.

178. Sperschneider ML, Hengartner MP, Kohl-Schwartz A, Geraedts K, Rauchfuss M, Woelfler MM, et al. Does endometriosis affect professional life? A matched case-control study in Switzerland, Germany and Austria. BMJ open. 2019;9(1):e019570.

179. As-Sanie S, Soliman A, Evans K, Erpelding N, Lanier R, Katz N. Healthcare utilization and cost burden among women with endometriosis by opioid prescription status in the first year after diagnosis: a retrospective claims database analysis. Journal of Medical Economics. 2020;23(4):371-7.

180. Chibber S, Baker R, Byrd T, Marsh E, editors. Managing Pain: Trends in Endometriosis-Related Emergency Department Visits Using the Nationwide Emergency Department Sample (NEDS) Database (2010-2014). REPRODUCTIVE SCIENCES; 2018: SAGE PUBLICATIONS INC 2455 TELLER RD, THOUSAND OAKS, CA 91320 USA.

181. Hudelist G, Aas-Eng MK, Birsan T, Berger F, Sevelda U, Kirchner L, et al. Pain and fertility outcomes of nerve-sparing, full-thickness disk or segmental bowel resection for deep infiltrating endometriosis—A prospective cohort study. Acta Obstetricia et Gynecologica Scandinavica. 2018;97(12):1438-46.

182. Parra RS, Feitosa MR, Valerio FP, de Camargo HP, Zanardi JVC, Feres O, et al. Laparoscopic bowel resection of deep infiltrating endometriosis. Comparative outcomes of a public teaching hospital and a referral private hospital. Acta Cirúrgica Brasileira. 2020;35.

183. Surrey ES, Soliman AM, Johns B, Vora J, Cross S, Agarwal SK. PDB58 REAL WORLD EVIDENCE ON THE IMPACT OF ELAGOLIX UTILIZATION ON PAIN SYMPTOMS, TREATMENT PATTERNS AND HEALTH RESOURCE USE AMONG ENDOMETRIOSIS PATIENTS IN THE UNITED STATES. Value in Health. 2020;23:S119.

© Gedeon Richter UK Ltd (2023). All rights reserved

B.5 Appendices

Appendix C: Summary of product characteristics (SmPC) and UK public assessment report

Appendix D: Identification, selection and synthesis of clinical evidence

Appendix E: Subgroup analysis

Appendix F: Adverse reactions

Appendix G: Published cost-effectiveness studies

Appendix H: Health-related quality of life studies

Appendix I: Cost and healthcare resource use identification

Appendix J: Clinical outcomes and disaggregated results from the model

Appendix K: Price details of treatments included in the submission

Appendix M: Summary of model inputs

© Gedeon Richter UK Ltd (2023). All rights reserved

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Relugolix–estradiol–norethisterone acetate for treating symptoms of endometriosis [ID3982; GID-TA10873]

Summary of Information for Patients (SIP)

October 2023

File name	Version	Contains confidential information	Date
ID3982_RelugolixCT_SIP_FINAL_06.10.23_[noCON]	Final	No	6 Oct 2023

Summary of Information for Patients (SIP):

The pharmaceutical company perspective

What is the SIP?

The Summary of Information for Patients (SIP) is written by the company who is seeking approval from NICE for their treatment to be sold to the NHS for use in England. It is a plain English summary of their submission written for patients participating in the evaluation. It is not independently checked, although members of the public involvement team at NICE will have read it to double-check for marketing and promotional content before it is sent to you.

The **Summary of Information for Patients** template has been adapted for use at NICE from the Health Technology Assessment International – Patient & Citizens Involvement Group (HTAi PCIG). Information about the development is available in an open-access IJTAHC journal article

SECTION 1: Submission summary

Note to those filling out the template: Please complete the template using plain language, taking time to explain all scientific terminology. Do not delete the grey text included in each section of this template as you move through drafting because it might be a useful reference for patient reviewers. Additional prompts for the company have been in red text to further advise on the type of information which may be most relevant and the level of detail needed. You may delete the red text.

1a) Name of the medicine (generic and brand name):

Response: Relugolix in combination with oestradiol and norethisterone acetate (or Relugolix combination therapy [Relugolix CT] for short). The brand name is Ryeqo®.

1b) Population this treatment will be used by. Please outline the main patient population that is being appraised by NICE:

Response: Adults with symptoms of endometriosis

1c) Authorisation: Please provide marketing authorisation information, date of approval and link to the regulatory agency approval. If the marketing authorisation is pending, please state this, and reference the section of the company submission with the anticipated dates for approval.

Response: Relugolix CT is currently being evaluated by the European Medicines Agency (the organisation that gives companies the legal right to sell medicines in the European Union). Once they have given marketing authorisation, Relugolix CT will undergo a fast-track approval by the Medicines and Healthcare products Regulatory Authority in the UK. More information is presented in section B.1.2 of the main submission (Document B).

1d) Disclosures. Please be transparent about any existing collaborations (or broader conflicts of interest) between the pharmaceutical company and patient groups relevant to the medicine. Please outline the reason and purpose for the engagement/activity and any financial support provided:

Response: Not applicable

SECTION 2: Current landscape

Note to authors: This SIP is intended to be drafted at a global level and typically contain global data. However, the submitting local organisation should include country-level information where needed to provide local country-level context.

Please focus this submission on the **main indication (condition and the population who would use the treatment)** being assessed by NICE rather than sub-groups, as this could distract from the focus of the SIP and the NICE review overall. However, if relevant to the submission please outline why certain sub-groups have been chosen.

2a) The condition – clinical presentation and impact

Please provide a few sentences to describe the condition that is being assessed by NICE and the number of people who are currently living with this condition in England. Please outline in general terms how the condition affects the quality of life of patients and their families/caregivers. Please highlight any mortality/morbidity data relating to the condition if available. If the company is making a case for the impact of the treatment on carers this should be clearly stated and explained.
Response: Endometriosis is a condition which usually affects those assigned female at birth between puberty and menopause. Endometriosis happens when cells usually found in the womb are found in other parts of the body. These cells react to hormonal changes each month like those in the womb, but when this happens outside the womb it can result in pain, inflammation, and bleeding (1). Symptoms of endometriosis are different for everyone, but they can include chronic pain, fatigue, depression, an inability to conceive, problems in working and social life, and relationship or sexual issues. There is no definitive cure for endometriosis, and it can severely impact quality of life and wellbeing (1). Endometriosis is estimated to affect 1.5 million women in the UK, which is similar to the number affected by diabetes (1).

2b) Diagnosis of the condition (in relation to the medicine being evaluated)

Please briefly explain how the condition is currently diagnosed and how this impacts patients. Are there any additional diagnostic tests required with the new treatment?
Response: The process of getting a diagnosis can be time consuming because the symptoms are often similar to other conditions. The only definitive way to diagnose endometriosis is laparoscopy, a procedure in which a doctor will look inside the tummy through a small cut using a narrow tube with an eyepiece (1).

2c) Current treatment options:

The purpose of this section is to set the scene on how the condition is currently managed:

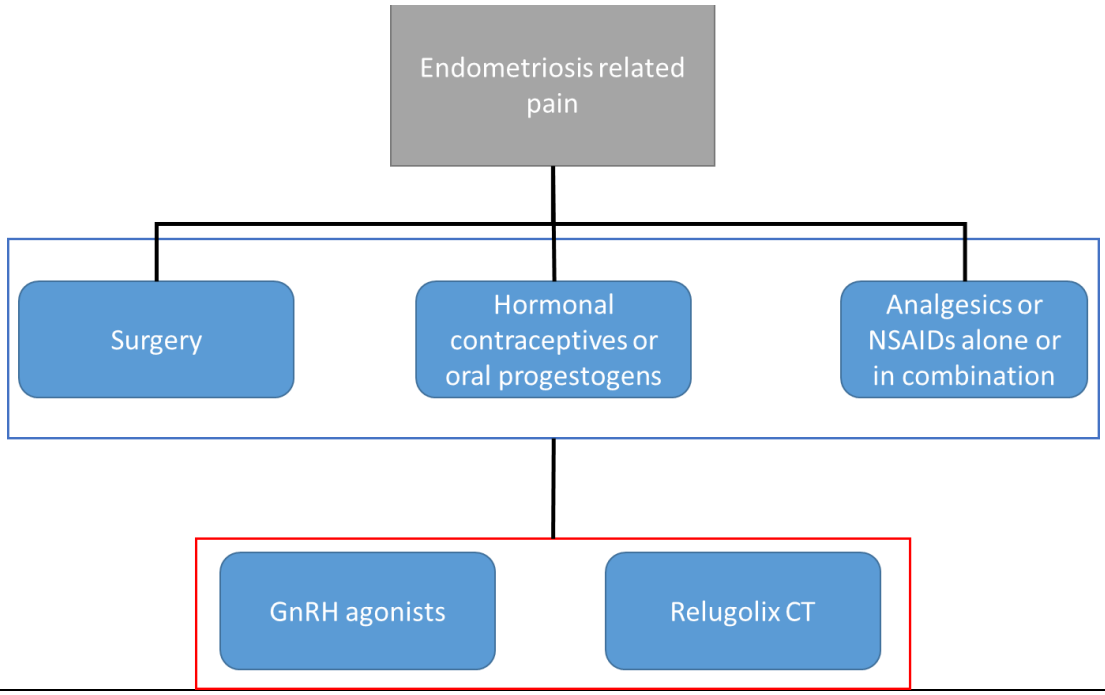
- What is the treatment pathway for this condition and where in this pathway the medicine is likely to be used? Please use diagrams to accompany text where possible. Please give emphasis to the specific setting and condition being considered by NICE in this review. For example, by referencing current treatment guidelines. It may be relevant to show the treatments people may have before and after the treatment under consideration in this SIP.
- Please also consider:
    - if there are multiple treatment options, and data suggest that some are more commonly used than others in the setting and condition being considered in this SIP, please report these data.
    - are there any drug–drug interactions and/or contraindications that commonly cause challenges for patient populations? If so, please explain what these are.

Response:

There is no cure for endometriosis and current treatment options aim to manage the symptoms of endometriosis. Hormone based treatments and surgical options are available, as are pain relief medications. Management of endometriosis can vary significantly, depending on the patient's stage of life and patient choice, and the type and location of the endometrioses (2). For example, hormonal therapies and some types of surgery have a contraceptive effect or pose a risk to fertility and should not be used in women who are trying to conceive or who wish to have children in future.

It is anticipated that Relugolix CT will be used at the same point in the treatment pathway as gonadotrophin releasing hormone (GnRH) agonists, i.e. after surgery, hormonal treatments, and analgesics (pain killers) have failed to control symptoms (Figure 1).

Figure 1: Proposed position of Relugolix CT in the endometriosis treatment pathway



2d) Patient-based evidence (PBE) about living with the condition

Context:

- **Patient-based evidence (PBE)** is when patients input into scientific research, specifically to provide experiences of their symptoms, needs, perceptions, quality of life issues or experiences of the medicine they are currently taking. PBE might also include carer burden and outputs from patient preference studies, when conducted in order to show what matters most to patients and carers and where their greatest needs are. Such research can inform the selection of patient-relevant endpoints in clinical trials.

In this section, please provide a summary of any PBE that has been collected or published to demonstrate what is understood about **patient needs and disease experiences**. Please include the methods used for collecting this evidence. Any such evidence included in the SIP should be formally referenced wherever possible and references included.

Response:

A survey of 10,000 people living in the UK who were diagnosed with endometriosis was conducted as part of an all-party parliamentary group report on the burden of endometriosis in 2020. In the survey, 95% and 81% of respondents said that endometriosis had had a negative, or very negative, impact on their wellbeing and mental health, respectively. Furthermore, 89% felt isolated due to their condition, and 90% would have liked access to psychological support (3).

SECTION 3: The treatment

Note to authors: Please complete each section with a concise overview of the key details and data, including plain language explanations of any scientific methods or terminology. Please provide all references at the end of the template. Graphs or images may be used to accompany text if they will help to convey information more clearly.

3a) How does the new treatment work?

What are the important features of this treatment?

Please outline as clearly as possible important details that you consider relevant to patients relating to the mechanism of action and how the medicine interacts with the body

Where possible, please describe how you feel the medicine is innovative or novel, and how this might be important to patients and their communities.

If there are relevant documents which have been produced to support your regulatory submission such as a summary of product characteristics or patient information leaflet, please provide a link to these.

Response:

Development of endometriosis depends on the hormone oestrogen. One of the active substances in Relugolix CT, **relugolix**, blocks the pituitary gland (a gland that controls many other hormone-producing glands in the body) from releasing luteinising hormone and follicle-stimulating hormone, which in turn prevents the production of progesterone and decreases the production of oestrogen (4).

Another active substance of Relugolix CT, **oestradiol**, is a natural sex hormone that helps to reduce symptoms related to the lowered levels of oestrogen caused by relugolix, such as hot flushes and bone density loss. However, oestradiol used alone can cause hyperplasia (growth) of the endometrium (the lining of the womb), which could lead to endometrial cancer. Relugolix CT, therefore, also contains the active substance **norethisterone acetate**, a synthetic progesterone replacement that blocks the effects of oestradiol on the womb, reducing the risk of endometrial growth (4).

The combination of relugolix with oestradiol and norhisterone acetate has the potential to improve patients' quality of life by providing long-term symptom control without the debilitating side effects caused by blockade of oestrogen production.

As the regulatory process is ongoing, the summary of product characteristics and patient information leaflet are not yet publicly available.

3b) Combinations with other medicines

Is the medicine intended to be used in combination with any other medicines?

- Yes/No

If yes, please explain why and how the medicines work together. Please outline the mechanism of action of those other medicines so it is clear to patients why they are used together.

If yes, please also provide information on the availability of the other medicine(s) as well as the main side effects.

If this submission is for a combination treatment, please ensure the sections on efficacy (3e), quality of life (3f) and safety/side effects (3g) focus on data that relate to the combination, rather than the individual treatments.

No, Relugolix CT is not intended to be used in combination with other medicines

3c) Administration and dosing

How and where is the treatment given or taken? Please include the dose, how often the treatment should be given/taken, and how long the treatment should be given/taken for.

How will this administration method or dosing potentially affect patients and caregivers? How does this differ to existing treatments?

Response:

Relugolix CT is available as a tablet. Each tablet contains 40 mg of relugolix, 1 mg of oestradiol and 0.5 mg of norethisterone acetate. Patients should take one tablet per day at about the same time, with or without food, but with a little liquid.

It is recommended that treatment starts within the first five days after the start of bleeding due to a period. Starting treatment at a different point in the menstrual cycle may result in initial irregular or heavier bleeding.

As Relugolix CT is taken orally, it is not expected to place any burden on patients or their carers. In contrast, GnRH agonists are administered either as a nasal spray several times a day or as injections every one or three months.

3d) Current clinical trials

Please provide a list of completed or ongoing clinical trials for the treatment. Please provide a brief top-level summary for each trial, such as title/name, location, population, patient group size, comparators, key inclusion and exclusion criteria and completion dates etc. Please provide references to further information about the trials or publications from the trials.

Response:

Three clinical trials have assessed Relugolix CT for the treatment of endometriosis: SPIRIT 1, SPIRIT 2 and SPIRIT OLE.

SPIRIT 1 (NCT03204318) and SPIRIT 2 (NCT03204331) had identical study designs and were carried out in Australasia, Europe, North America, South America and South Africa. The studies enrolled women between the ages of 18 and 50 who hadn't yet reached the menopause. To take part, patients had to have had a diagnosis of endometriosis within the last 10 years. They also had to have moderate or severe endometriosis pain, with a mean score of at least 4 out of 10 for period pain and at least 2.5 out of 10 for non-period pelvic pain. They were not allowed to take part if they had poor bone health, chronic pelvic pain not caused by endometriosis, or could not take Relugolix CT for any reason.

In each study, patients were randomly allocated to one of three treatments: Relugolix CT, placebo or delayed Relugolix CT. The allocation of treatments was double-blinded, which means neither the patients nor the people running the study knew which treatment each patient was taking. Treatment lasted for up to 24 weeks.

In total, 638 patients were enrolled in SPIRIT 1: 212 received Relugolix CT, 213 received placebo, and 213 received delayed Relugolix CT. SPIRIT 2 included 623 patients: 208 received Relugolix CT, 208 received placebo, and 207 received delayed Relugolix CT.

Both trials completed in 2021 and have been published in the Lancet (5)
https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(22)00622-5/fulltext

SPIRIT OLE (NCT03654274) was designed to assess long-term treatment with Relugolix CT. Patients could enter this study if they had completed 24 weeks of treatment in SPIRIT 1 or SPIRIT 2. In this study, all patients received Relugolix CT for up to 2 years. The study was open-label, which means that everyone involved knew that the patients were receiving Relugolix CT.

In total, 802 patients were enrolled in SPIRIT OLE.

The SPIRIT OLE study completed in early 2023 and has not been fully published yet; however results have been posted on ClinicalTrials.gov:
https://classic.clinicaltrials.gov/ct2/show/study/NCT03654274

3e) Efficacy

Efficacy is the measure of how well a treatment works in treating a specific condition.

In this section, please summarise all data that demonstrate how effective the treatment is compared with current treatments at treating the condition outlined in section 2a. Are any of the outcomes more important to patients than others and why? Are there any limitations to the data which may affect how to interpret the results? Please do not include academic or commercial in confidence information but where necessary reference the section of the company submission where this can be found.

Response:

Both SPIRIT 1 and 2 trials met their main goals for efficacy. The key efficacy outcomes for both studies are shown below (5). More detailed information on the efficacy data for Relugolix CT can be found in section B.2.6 of the main submission (Document B).

- In SPIRIT 1, 158 (75%) of 212 patients in the Relugolix CT group were considered to have improved period pain compared with 57 (27%) of 212 patients in the placebo group.
- In SPIRIT 1, 124 (59%) of 212 patients in the Relugolix CT group were considered to have improved non-period pelvic pain versus 84 (40%) patients in the placebo group
- In SPIRIT 2, 155 (75%) of 206 patients in the Relugolix CT group were considered to have improved period pain compared with 62 (30%) of 204 patients in the placebo group

- In SPIRIT 2, 136 (66%) of 206 patients were considered to have improved non-period pelvic pain in the Relugolix CT group compared with 87 (43%) of 204 patients in the placebo group

The SPIRIT OLE has not been fully published yet; however results have been posted on ClinicalTrials.gov: https://classic.clinicaltrials.gov/ct2/show/study/NCT03654274
The results show that the reductions in period pain and non-period pelvic pain seen during SPIRIT 1 & 2 were sustained for up to 2 years of treatment with Relugolix CT.

Although there are no trials that directly compare Relugolix with other available treatments, it is possible to compare them indirectly. The manufacturer of Relugolix CT has carried out an indirect comparison. The results are not published, but further information can be found in Section B.2.9 of the main submission (Document B).

3f) Quality of life impact of the medicine and patient preference information

What is the clinical evidence for a potential impact of this medicine on the quality of life of patients and their families/caregivers? What quality of life instrument was used? If the EuroQol-5D (EQ-5D) was used does it sufficiently capture quality of life for this condition? Are there other disease specific quality of life measures that should also be considered as supplementary information?

Please outline in plain language any quality of life related data such as **patient reported outcomes (PROs).**

Please include any **patient preference information (PPI)** relating to the drug profile, for instance research to understand willingness to accept the risk of side effects given the added benefit of treatment. Please include all references as required.

Response:

The EQ-5D-5L was used in the SPIRIT trials. These data are not currently publicly available; however they show that patients treated with Relugolix CT had greater improvements in their quality of life than those who received placebo. It is important to note that the EQ-5D was not designed specifically for use in endometriosis.

The Endometriosis Health Profile-30 (EHP-30) is a survey that is specifically designed for use in patients with endometriosis. It includes questions on pain, control and powerlessness, social support, emotional well-being and self-image (6). The EHP-30 was used to assess quality of life in the SPIRIT trials.

EHP-30 is scored on a scale of 0 to 100 where 0 is a perfect state of health and 100 is the worst possible health status. Reductions in EHP-30 scores were significantly higher in the Relugolix CT groups compared to the placebo groups for both SPIRIT 1 and SPIRIT 2. In SPIRIT 1 EHP-30 scores were reduced on average by 33.8 points for Relugolix CT vs 18.7 points for the placebo group. In SPIRIT 2 the EHP-30 domain score was reduced by 32.2 points on average for Relugolix CT vs 19.9 points for the placebo group.

Painful periods and non-menstrual pelvic pain have been shown to reduce quality of life (7-11). In SPIRIT 1&2, Relugolix CT decreased period pain within 8 weeks and non-period pain within 12 weeks of starting treatment (5).

Studies have also shown that there is a link between quality of life and sexual function, and sexual issues related to endometriosis have a meaningful impact on patients' lives (12). In SPIRIT 1 & 2, Relugolix CT reduced pain during sexual intercourse for patients with endometriosis. On average, in SPIRIT 1, patients reported a reduction in pain of 2.4 points out of 10 in the Relugolix CT group

vs 1.7 points out of 10 in the placebo group. In SPIRIT 2, patients on Relugolix CT reported a reduction of 2.4 points out of 10 vs 1.9 points out of 10 in the placebo group (5).

3g) Safety of the medicine and side effects

When NICE appraises a treatment, it will pay close attention to the balance of the benefits of the treatment in relation to its potential risks and any side effects. Therefore, please outline the main side effects (as opposed to a complete list) of this treatment and include details of a benefit/risk assessment where possible. This will support patient reviewers to consider the potential overall benefits and side effects that the medicine can offer.

Based on available data, please outline the most common side effects, how frequently they happen compared with standard treatment, how they could potentially be managed and how many people had treatment adjustments or stopped treatment. Where it will add value or context for patient readers, please include references to the Summary of Product Characteristics from regulatory agencies etc.

Response:

Studies in endometriosis and another condition called uterine fibroids have shown that the most common side effects with Relugolix CT (which may affect more than 1 in 10 people) are headache and hot flushes.

In the SPIRIT 1 & 2 trials, the percentage of patients who had side effects was similar in the Relugolix CT and placebo groups (71% with Relugolix CT vs 66% with placebo in SPIRIT 1, and 81% with Relugolix CT vs 75% with placebo in SPIRIT 2) (5). The most common side-effects were headache, nasopharyngitis (inflammation of the nasal passages and throat) and hot flushes (see table). Hot flushes mostly occurred within the first 12 weeks of treatment (5).

Table 1: Side effects reported in SPIRIT 1 and 2 trials (5)

	Number (%) of patients			
	SPIRIT 1		**SPIRIT 2**	
	Relugolix CT	**Placebo**	**Relugolix CT**	**Placebo**
Headache	57 (27)	46 (22)	81 (39)	64 (31)
Nasophayngitis	13 (6)	12 (6)	29 (14)	17 (8)
Hot flushes	22 (10)	21 (10)	28 (14)	7 (3)

Very few patients withdrew from the SPIRIT 1 & 2 studies because of side-effects with Relugolix CT: 4% in SPIRIT 1 (vs 2% with placebo) and 5% in SPIRIT 2 (vs 4% with placebo) (5).

Relugolix CT must not be used in women who have, or have had, venous thromboembolism (blood clots in the veins) or those who have had a stroke or a heart attack. It must also not be used in women who have a blood clotting disorder, osteoporosis, migraines or headaches with neurological symptoms, cancers that are influenced by sex hormones (such as breast cancer or genital cancer), liver tumours, or abnormal liver function, or in women who are pregnant, breastfeeding or have genital bleeding of unknown cause (4).

Relugolix CT must not be used together with hormonal contraception (4). After four weeks of use, Relugolix CT provides adequate contraceptive protection.

3h) Summary of key benefits of treatment for patients

Issues to consider in your response:

- Please outline what you feel are the key benefits of the treatment for patients, caregivers and their communities when compared with current treatments.
- Please include benefits related to the mode of action, effectiveness, safety and mode of administration
-

Response:

Endometriosis is a chronic condition for which there is no cure. Without proper management it can lead to years of pain, which can have a debilitating physical and social impact. GnRH agonists are not effective for all patients and can only be taken for a limited time because of their adverse effect on bone health. The combination of relugolix with oestradiol and norethisterone acetate in Relugolix CT has the potential to provide symptom control without impacting bone health, meaning it can be taken for long periods of time.

As an oral formulation, Relugolix CT offers a less invasive route of administration than those GnRH agonists that are administered via injections. This is a benefit for patients who find injections uncomfortable or who experience painful reactions to injections.

3i) Summary of key disadvantages of treatment for patients

Issues to consider in your response:

- Please outline what you feel are the key disadvantages of the treatment for patients, caregivers and their communities when compared with current treatments. Which disadvantages are most important to patients and carers?
- Please include disadvantages related to the mode of action, effectiveness, side effects and mode of administration
- What is the impact of any disadvantages highlighted compared with current treatments

Response:

Unlike GnRH agonist injections, which are given every one or three months, Relugolix CT needs to be taken every day. Some patients may occasionally forget to take a dose. If two or more tables are missed on consecutive days, the contraceptive effect of Relugolix CT may be reduced and the patient will need to use a non-hormonal form of contraception for the next seven days.

3i) Value and economic considerations

Introduction for patients:

Health services want to get the most value from their budget and therefore need to decide whether a new treatment provides good value compared with other treatments. To do this they consider the costs of treating patients and how patients' health will improve, from feeling better and/or living longer, compared with the treatments already in use. The drug manufacturer provides this information, often presented using a health economic model.

In completing your input to the NICE appraisal process for the medicine, you may wish to reflect on:

- The extent to which you agree/disagree with the value arguments presented below (e.g., whether you feel these are the relevant health outcomes, addressing the unmet needs and issues faced by

patients; were any improvements that would be important to you missed out, not tested or not proven?)

- If you feel the benefits or side effects of the medicine, including how and when it is given or taken, would have positive or negative financial implications for patients or their families (e.g., travel costs, time-off work)?
- How the condition, taking the new treatment compared with current treatments affects your quality of life.

Response:

The manufacturer of Relugolix CT built an economic model in Microsoft Excel to explore the cost-effectiveness of Relugolix CT when compared with GnRH agonists in pre-menopausal patients who had previously failed treatment with GnRH agonists. The economic model shows the different ways in which a patient's health can change after medical treatments and surgery. It compared the total costs (drugs and healthcare resource use) generated by Relugolix CT and GnRH agonists as well as the survival and quality of life over their lifetime; these last two are combined to produce a measure called the quality-adjusted life year (QALY). One QALY is equal to one year of life in perfect health.

The model used data from the SPIRIT trials; the key input was response to treatment measured as a reduction in period pain and non-period pain with no increase in use of painkillers. The model also included factors such as withdrawal from treatment, choice of surgery, repeat surgery and treatment schedules.

3j) Innovation

NICE considers how innovative a new treatment is when making its recommendations.
If the company considers the new treatment to be innovative please explain how it represents a 'step change' in treatment and/ or effectiveness compared with current treatments. Are there any QALY benefits that have not been captured in the economic model that also need to be considered (see section 3f)

Response:

There is an unmet need for an effective, non-surgical treatment for endometriosis that can be administered orally and on a long-term basis. Relugolix CT is a novel oral GnRH antagonist that meets this unmet need. There are currently no oral pharmacological treatment options licensed for the long-term treatment of endometriosis symptoms.

3k) Equalities

Are there any potential equality issues that should be taken into account when considering this condition and this treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged. Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics More information on how NICE deals with equalities issues can be found in the NICE equality scheme Find more general information about the Equality Act and equalities issues here
Response: There is evidence to suggest that women from some minority ethnic groups may be underdiagnosed and/or visit their doctor later for help with endometriosis and thus have more severe symptoms (13). The Endometriosis All-Party Parliamentary Group Report (October 2020) also highlights that Black, Asian and minority ethnic communities can receive a lower quality of care. These health inequalities have been thought to be due to socioeconomic factors since Black, Asian, and minority ethnic women are more likely to live in areas of high deprivation, have lower incomes, experience language barriers and have poorer access to women's healthcare services (3).

SECTION 4: Further information, glossary and references

4a) Further information

Feedback suggests that patients would appreciate links to other information sources and tools that can help them easily locate relevant background information and facilitate their effective contribution to the NICE assessment process. Therefore, please provide links to any relevant online information that would be useful, for example, published clinical trial data, factual web content, educational materials etc. Where possible, please provide open access materials or provide copies that patients can access.
Response: European Medicines Agency website: https://www.ema.europa.eu/en/medicines/human/EPAR/ryeqo SPIRIT 1&2 clinical trials: published paper in Lancet: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(22)00622-5/fulltext SPIRIT OLE: results have been posted on ClinicalTrials.gov: https://classic.clinicaltrials.gov/ct2/show/study/NCT03654274 Endometriosis UK: https://www.endometriosis-uk.org/ All parliamentary group report on endometriosis, 2020 https://www.endometriosis-uk.org/sites/default/files/files/Endometriosis%20APPG%20Report%20Oct%202020.pdf NICE guidance for endometriosis diagnosis and management https://www.nice.org.uk/guidance/ng73

Further information on NICE and the role of patients**:**

- Public Involvement at NICE Public involvement | NICE and the public | NICE Communities | About | NICE
- NICE's guides and templates for patient involvement in HTAs Guides to developing our guidance | Help us develop guidance | Support for voluntary and community sector (VCS) organisations | Public involvement | NICE and the public | NICE Communities | About | NICE
- EUPATI guidance on patient involvement in NICE: https://www.eupati.eu/guidance-patient-involvement/
- EFPIA – Working together with patient groups: https://www.efpia.eu/media/288492/working-together-with-patient-groups-23102017.pdf
- National Health Council Value Initiative. https://nationalhealthcouncil.org/issue/value/
- INAHTA: http://www.inahta.org/
- European Observatory on Health Systems and Policies. Health technology assessment - an introduction to objectives, role of evidence, and structure in Europe: http://www.inahta.org/wp-content/themes/inahta/img/AboutHTA_Policy_brief_on_HTA_Introduction_to_Objectives_Role_of_Evidence_Structure_in_Europe.pdf

4b) Glossary of terms

Response:

Analgesic – A drug that reduces pain

Chronic – A health problem that requires ongoing management over a period of years or decades and is one that cannot currently be cured but can be controlled with the use of medication and/or other therapies.

Contraceptive – Any drug, device, or method preventing pregnancy.

Endometriosis – A disease in which tissue similar to the lining of the womb grows outside the womb. It can cause severe pain in the pelvis and make it harder to get pregnant.

Endometrium – The lining of the womb

Follicle stimulating hormone - A hormone that influences the production of the hormones progesterone and oestrogen which are involved in the development of endometriosis.

Gonadotropin releasing hormone (GnRH) agonist/ analogue – A medication which reduces the levels of oestrogen in the body.

Gonadotropin releasing hormone (GnRH) antagonist – A medication which reduces the levels of oestrogen in the body. Relugolix is a GnRH antagonist.

Hormone - Any of various chemicals made by living cells that influence the development, growth, sex, etc. and are carried around the body in the blood.

Hyperplasia – Growth

Inflammation - A response triggered by damage to living tissue that leads to redness, pain, and swelling.

Luteinising hormone – A hormone that influences the production of the hormones progesterone and oestrogen which are involved in the development of endometriosis.

Menopause - Menopause is when periods stop due to a fall in hormone levels. It usually affects women between the ages of 45 and 55, but it can happen earlier.

Nasopharyngitis – Inflammation of the nasal passages, throat, or the area behind the nose and mouth.

Neurological – Relating to the brain, spinal cord, or nerves.

NSAID – Non-steroidal anti-inflammatory drug. A medicine widely used to relieve pain, reduce inflammation, and bring down a temperature.

Oestrogen – A sex hormone that influences the mechanism behind endometriosis related pain.

Oestradiol – Is a type of oestrogen (see above)

Oral – Taken by mouth.

Osteoperosis - a bone disease that develops when bone mineral density and bone mass decreases, or when the structure and strength of bone changes.

Pituitary gland - a gland that controls many other hormone-producing glands in the body

Progesterone – A sex hormone which influences the mechanism behind endometriosis related pain.

Progestogen - A steroid hormone that acts like progesterone, a hormone that prepares the uterus for pregnancy. Progestogen is used in oral contraceptives and to treat gynecological disorders (including endometriosis). Noresthisterone acetate is a progestogen.

QALY – Quality adjusted life year. A measure of how well a treatment improves or lengthens a patient's life. One QALY is equal to one year of life in perfect health.

Relugolix CT – Relugolix combination therapy, comprising 40 mg relugolix, 1 mg estradiol (as hemihydrate), and 0.5 mg norethisterone acetate. The brand name for Relugolix CT is Ryeqo®.

Venous thromboembolism – Blood clots in the veins

4c) References

Please provide a list of all references in the Vancouver style, numbered and ordered strictly in accordance with their numbering in the text:

Response:

1. Endometriosis UK. THE NATIONAL ENDOMETRIOSIS SOCIETY INFORMATION - Understanding-endometriosis: Endometriosis UK; 2012 [Available from: http://www.endometriosis-uk.org.
2. National Institute for Health and Care Excellence. Endometriosis: diagnosis and management guidance NICE: NICE; 2017 [Available from: https://www.nice.org.uk/guidance/ng73/chapter/Context.
3. All Party Parliamentary Group on Endometriosis. Endometriosis in the UK: Time for Change. House of Commons; 2020.
4. European Medicines Agency. Ryeqo [Available from: https://www.ema.europa.eu/en/medicines/human/EPAR/ryeqo.
5. Giudice LC, As-Sanie S, Arjona Ferreira JC, Becker CM, Abrao MS, Lessey BA, et al. Once daily oral relugolix combination therapy versus placebo in patients with endometriosis-associated pain: two replicate phase 3, randomised, double-blind, studies (SPIRIT 1 and 2). Lancet. 2022;399(10343):2267-79.
6. University of Oxford Innovation. The Endometriosis Health Profile (EHP) 2016 [Available from: https://innovation.ox.ac.uk/outcome-measures/endometriosis-health-profile-ehp/#:~:text=The%20purpose%20of%20the%20EHP,to%20100%20worst%20health%20status.
7. Allyn K, Evans S, Seidman LC, Payne LA. "Tomorrow, I'll Be Fine": Impacts and coping mechanisms in adolescents and young adults with primary dysmenorrhoea. J Adv Nurs. 2020;76(10):2637-47.
8. Facchin F, Barbara G, Saita E, Mosconi P, Roberto A, Fedele L, et al. Impact of endometriosis on quality of life and mental health: pelvic pain makes the difference. Journal of Psychosomatic Obstetrics & Gynecology. 2015;36(4):135-41.
9. Fernández-Martínez E, Onieva-Zafra MD, Parra-Fernández ML. The Impact of Dysmenorrhea on Quality of Life Among Spanish Female University Students. Int J Environ Res Public Health. 2019;16(5).
10. Hooker AB, van Moorst BR, van Haarst EP, van Ootegehem NA, van Dijken DK, Heres MH. Chronic pelvic pain: evaluation of the epidemiology, baseline demographics, and clinical variables via a prospective and multidisciplinary approach. Clin Exp Obstet Gynecol. 2013;40(4):492-8.
11. Tripoli TM, Sato H, Sartori MG, de Araujo FF, Girão MJ, Schor E. Evaluation of quality of life and sexual satisfaction in women suffering from chronic pelvic pain with or without endometriosis. J Sex Med. 2011;8(2):497-503.
12. Della Corte L, Di Filippo C, Gabrielli O, Reppuccia S, La Rosa VL, Ragusa R, et al. The Burden of Endometriosis on Women's Lifespan: A Narrative Overview on Quality of Life and Psychosocial Wellbeing. Int J Environ Res Public Health. 2020;17(13).
13. Hudson. The missed disease? Endometriosis as an example of 'undone science'. Reproductive biomedicine & society online. 2021;14.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

[ID3982]

Clarification questions – company response

October 2023

File name	Version	Contains confidential information	Date
ID3982 Relugolix EAG clarification letter_company responses 09Nov23 [REDACTED]	Final	Yes	09/11/23

Notes for company

Highlighting in the template

Square brackets and grey highlighting are used in this template to indicate text that should be replaced with your own text or deleted. These are set up as form fields, so to replace the prompt text in [grey highlighting] with your own text, click anywhere within the highlighted text and type. Your text will overwrite the highlighted section.

To delete grey highlighted text, click anywhere within the text and press DELETE.

Section A: Clarification on effectiveness data

Literature searches

A 1. Priority question: The Evidence Assessment Group (EAG) was concerned that the searches appear to contain a number of limitations, which may account for the low recall of results. Many of the identified weaknesses are carried throughout the clinical and economic searches. Taking the primary Embase search strategy reported for the clinical effectiveness searches in Appendix D as an example, here are some of the key areas of concern:

a) Conditions Facet

i. **This facet only contains subject headings, no free text terms.**

ii. **The main subject heading for endometriosis is not exploded (although this does happen in the update searches)**

iii. **Records containing subject headings for adenomyosis / uterus myoma / and ovary cancer/ are excluded from the search results using the Boolean operator NOT. The use of NOT is generally not recommended. If included in a strategy it should always be used with extreme caution, as it can easily remove relevant records**

containing both terms. In this example, the EAG would request that the strategy be amended to remove this use of the NOT operator.

b) Interventions Facet

i. **Missing subject heading and synonyms for Relugolix, although use of the '.mp.' field tag may negate some loss of recall.**

ii. **Failure to explode subject headings for 'oral contraceptive agent/' (misses trade names of individual products) and 'contraceptive agent/' (misses subheading 'hormonal contraceptive agent/' and subject headings below it in the EMTREE hierarchy).**

iii. **Further missing relevant free text and subject headings for named comparators.**

iv. **Table 91 in Appendix D lists the predefined set of criteria for study selection. In the list of comparators for the original systematic literature review (SLR), the following surgical procedures are listed:**

Surgery:

Conservative procedures:

Surgical ablation/excision

Ovarian cystectomy

Laparoscopy

Definitive procedures:

Removal of endometrioma

Abdominal or vaginal hysterectomy

Salpingectomy/tuboplasty

Oophorectomy

However, the search strategies only contain limited search terms for two of the listed surgeries: endometrium ablation and laparotomy.

c) Pain Facet

i. **Given the low number of results for Embase (n=509) the inclusion of a pain facet feels both unnecessary and overly restrictive. The EAG would recommend removing this.**

Given the lack of relevant papers found, the EAG would request that the main Embase and MEDLINE clinical searches be rerun and expanded with the above points in mind and the resulting new papers screened for includes. Please note that the limitations described in the condition facet and the inclusion of a pain facet are also present in the economics searches, therefore the EAG would request that the main Embase and MEDLINE economics searches also be rerun and screened to check that no potentially relevant papers have been missed.

Company response: As outlined in our e-mail dated 19th October 2023, Gedeon Richter would ask the EAG to reconsider their request for a full re-run of the SLR, given that this would take a significant amount of time and delay the appraisal. We propose an expedited solution that makes use of a recently-published Cochrane review by Veth et al. that assessed the efficacy and safety of GnRH agonists for the treatment of painful symptoms associated with endometriosis (1). This would involve carrying out and reporting on a feasibility assessment of any studies in this review that were not identified in the submitted clinical SLR. We also propose running updated searches based on the Cochrane review search strategies, as the original Cochrane searches were carried out in May 2022.

Gedeon Richter are aware that the EAG has concerns about whether the scope of the Cochrane review matches the submission scope. However, we believe the only significant difference between the two is that the Cochrane review did not include Relugolix CT. Table 1 shows a comparison of the NICE scope with the Cochrane

scope. As stated in our submission, we consider the comparator for Relugolix CT to be GnRH agonists. Non-steroidal anti-inflammatory drugs (NSAIDs) are generally included under analgesics in the majority of trials, and we are not aware of the existence of any randomised controlled trials (RCTs) comparing either GnRH agonists or antagonists with neuromodulators (which in general comprise medical devices such as vagus nerve stimulators).

Table 1 Comparison of the NICE scope with the Cochrane review scope

	Final NICE scope	**Cochrane review**
Population	Adults with symptoms of endometriosis	Endometriosis
Intervention	Relugolix in combination with oestradiol and norethisterone acetate (also known as norethisterone acetate)	GnRH agonists
Comparator(s)	Established clinical management without relugolix in combination with oestradiol and norethisterone, including: • analgesics or non-steroidal anti-inflammatory drug (NSAID) alone or in combination with each other • neuromodulators • hormonal treatment such as combined hormonal contraception (off-label for some combined hormonal contraceptives), oral progestogens, gonadotropin-releasing hormone (GnRH) agonists.	Analgesics Calcium-regulating agents Hormonal treatment (gestrinone, progesterone, danazol, add-back therapy) Placebo
Outcomes	The outcome measures to be considered include: • overall pain • opioid use • analgesic use • recurrence of endometriosis • admission to hospital • subsequent surgical treatment • fertility • adverse effects of treatment	Overall pain associated with endometriosis Adverse effects

	• complications of treatment • health-related quality of life	

The feasibility assessment for the ITC included 92 publications; 67 that were identified in the SLR and 25 that were identified in a pragmatic literature search that involved searching the web using key words related to GnRH agonist therapies used to treat moderate-to-severe pain associated with endometriosis. In addition, we have identified 12 publications in the Cochrane review that were not in the submitted clinical SLR and we are now working on a extraction of their outcomes. As per our email of 2nd November 2023, we will provide a full response on this on Thursday 16th November.

A 2. Please confirm whether any additional searches, other than those reported in Appendix D Section D.1.1, were conducted to retrieve information regarding adverse events (AEs) for Relugolix and, if so, provide full details including date, resource names and search strategies used.

Company response: We can confirm that no additional searches were conducted.

A 3. Whilst searches are reported for Medline and Embase, with additional pragmatic grey literature searches, there are no searches reported for the Cochrane Library (either CDSR or CENTRAL), please explain the rationale behind not searching these resources.

Company response: Cochrane Reviews and Editorials are indexed in PubMed, and searched were run in PubMed. Therefore, no separate search of the Cochrane Library was carried out.

A 4. The EAG noted that the update search of Embase.com (Appendix D) carried the following date limit: Line #37 (#33 AND #34 AND [humans]/lim AND [01-04-2022]/sd NOT [01-11-2022]/sd) Given that this search was run on 1 Dec 2022, this line would appear to discard results added to the database since 1st November 2022. Please confirm if this is the case and, if so, rerun these searches and screen the previously discarded records.

Company response: We can confirm that the Embase searches were run on 1st November 2022, and that this is a typographical error in the submission.

A 5. The syntax for the Pubmed update search appears unusual. Unlike the economics searches there are no field tags and the line combinations are missing the hash tag before the line numbers which affects the ability to rerun the searches. Please can you confirm that the update search for clinical effectiveness was conducted using Pubmed and provide the original search strategy as run. The Cochrane Manual recommends that *"… bibliographic database search strategies should be copied and pasted into an appendix exactly as run and in full, together with the search set numbers and the total number of records retrieved by each search strategy. The search strategies should not be re-typed, because this can introduce errors"* (Cochrane Handbook for Systematic Reviews of Interventions. Version 6.4, 2023; Section 4.5, https://training.cochrane.org/handbook/current/chapter-04#section-4-5).

Company response: For the original SLR, the searches were run in OVID, but for the update they were run in PubMed itself. The search strategy is provided below.

Medline search strategies (Current [First] Update of Report)		
Facet	Terms	Hits
1 - disease	("endometriosis"[MeSH Terms] OR "endometriosis"[All Fields] OR "endometrioses"[All Fields]) NOT ("adenomyosis"[MeSH Terms] OR "adenomyosis"[All Fields] OR "adenomyoses"[All Fields] OR "uterus myoma"[All Fields] OR "ovary cancer"[All Fields] OR "ovarian cancer"[All Fields])	29,399
2 - treatments	"gonadotropine"[All Fields] OR "gonadotropines"[All Fields] OR "gonadotropins"[MeSH Terms] OR "gonadotropins"[All Fields] OR "gonadotropin"[All Fields] OR "gonadotropin-releasing hormone"[All Fields] OR (("gonadotropin-releasing hormone"[MeSH Terms] OR ("gonadotropin releasing"[All Fields] AND "hormone"[All Fields]) OR "gonadotropin-releasing hormone"[All Fields] OR "gnrh"[All Fields]) AND "agonist*"[All Fields]) OR (("gonadotropin-releasing hormone"[MeSH Terms] OR ("gonadotropin releasing"[All Fields] AND "hormone"[All Fields]) OR "gonadotropin-releasing hormone"[All Fields] OR "gnrh"[All Fields]) AND "antag*"[All Fields]) OR (("gonadotropin-releasing hormone"[MeSH Terms] OR ("gonadotropin releasing"[All Fields] AND "hormone"[All Fields]) OR "gonadotropin-releasing hormone"[All Fields] OR "gnrh"[All Fields]) AND "analog*"[All Fields]) OR ("relugolix"[Supplementary Concept] OR "relugolix"[All Fields]) OR ("elagolix"[Supplementary Concept] OR "elagolix"[All Fields]) OR ("linzagolix"[Supplementary Concept] OR "linzagolix"[All Fields]) OR "oral contraceptive"[All Fields] OR ("combin*"[All Fields] AND ("mouth"[MeSH Terms] OR "mouth"[All Fields] OR "oral"[All Fields] OR "hormone*"[All Fields]) AND ("pill*"[All Fields] OR "contracept*"[All Fields])) OR (("non steroid*"[All Fields] OR "nonsteroid*"[All Fields]) AND "anti-inflammatory"[All Fields] AND	645,665

	("agent"[All Fields] OR "agents"[All Fields])) OR "nsaid*"[Title/Abstract] OR "opioid*"[Title/Abstract] OR ("progestinic"[All Fields] OR "progestinics"[All Fields] OR "progestins"[Pharmacological Action] OR "progestins"[MeSH Terms] OR "progestins"[All Fields] OR "progestin"[All Fields]) OR ("dienogest"[Supplementary Concept] OR "dienogest"[All Fields]) OR "medroxyprogesterone acetate"[All Fields] OR ("leuprolide"[MeSH Terms] OR "leuprolide"[All Fields] OR "leuprorelin"[All Fields]) OR ("leuprolide"[MeSH Terms] OR "leuprolide"[All Fields]) OR "intrauterine device*"[All Fields] OR "levonorgestrel"[All Fields] OR "aromatase inhibitor"[All Fields] OR ("androgen s"[All Fields] OR "androgene"[All Fields] OR "androgenes"[All Fields] OR "androgenic"[All Fields] OR "androgenicity"[All Fields] OR "androgenized"[All Fields] OR "androgenizing"[All Fields] OR "androgenous"[All Fields] OR "androgens"[Pharmacological Action] OR "androgens"[MeSH Terms] OR "androgens"[All Fields] OR "androgen"[All Fields] OR "virilism"[MeSH Terms] OR "virilism"[All Fields] OR "androgenization"[All Fields]) OR ("danazol"[MeSH Terms] OR "danazol"[All Fields] OR "danazole"[All Fields]) OR ("laparotomy"[MeSH Terms] OR "laparotomy"[All Fields] OR "laparotomies"[All Fields]) OR "endometrium ablation"[All Fields]	
3 - outcomes	"dysmenorrhea"[MeSH Terms] OR "dysmenorrhea"[All Fields] OR "dysmenorrheas"[All Fields] OR "dysmenorrhoea"[All Fields] OR "pelvic pain"[All Fields] OR (("non-menstrual"[All Fields] OR ("nonmenstrual"[All Fields] OR "nonmenstruating"[All Fields])) AND ("pelvics"[All Fields] OR "pelvis"[MeSH Terms] OR "pelvis"[All Fields] OR "pelvic"[All Fields]) AND ("pain"[MeSH Terms] OR "pain"[All Fields])) OR "nmpp"[All Fields] OR ("dyspareunia"[MeSH Terms] OR "dyspareunia"[All Fields]) OR "endometriosis pain"[All Fields] OR "pain"[All Fields]	900,911
4 – trial terms	"randomized controlled trial"[Publication Type] OR "randomized controlled trials as topic"[MeSH Terms] OR "randomized controlled trial"[All Fields] OR "randomised controlled trial"[All Fields] OR "controlled clinical trial"[Publication Type] OR "controlled clinical trials as topic"[MeSH Terms] OR "controlled clinical trial"[All Fields] OR "randomized"[Title/Abstract] OR "placebo"[Title/Abstract] OR "drug therapy"[MeSH Subheading] OR "trial"[Title/Abstract]	3,579,356
4 - Combined	1 AND 2 AND 3 AND 4	954
5 - + filters	((("endometriosis"[MeSH Terms] OR "endometriosis"[All Fields] OR "endometrioses"[All Fields]) NOT ("adenomyosis"[MeSH Terms] OR "adenomyosis"[All Fields] OR "adenomyoses"[All Fields] OR "uterus myoma"[All Fields] OR "ovary cancer"[All Fields] OR "ovarian cancer"[All Fields])) AND ("gonadotropine"[All Fields] OR "gonadotropines"[All Fields] OR "gonadotropins"[MeSH Terms] OR "gonadotropins"[All Fields] OR "gonadotropin"[All Fields] OR "gonadotropin-releasing hormone"[All Fields] OR (("gonadotropin-releasing hormone"[MeSH Terms] OR ("gonadotropin releasing"[All Fields] AND "hormone"[All Fields]) OR "gonadotropin-releasing hormone"[All Fields] OR "gnrh"[All Fields]) AND "agonist*"[All Fields]) OR (("gonadotropin-releasing hormone"[MeSH Terms] OR ("gonadotropin releasing"[All Fields] AND "hormone"[All Fields]) OR "gonadotropin-releasing hormone"[All Fields] OR "gnrh"[All Fields]) AND "antag*"[All Fields]) OR (("gonadotropin-releasing hormone"[MeSH Terms] OR ("gonadotropin releasing"[All Fields] AND "hormone"[All Fields]) OR "gonadotropin-releasing hormone"[All Fields] OR "gnrh"[All Fields]) AND "analog*"[All Fields]) OR ("relugolix"[Supplementary Concept] OR "relugolix"[All Fields]) OR ("elagolix"[Supplementary Concept] OR "elagolix"[All Fields]) OR ("linzagolix"[Supplementary Concept] OR "linzagolix"[All Fields]) OR "oral contraceptive"[All Fields] OR ("combin*"[All	26

	Fields] AND ("mouth"[MeSH Terms] OR "mouth"[All Fields] OR "oral"[All Fields] OR "hormone*"[All Fields]) AND ("pill*"[All Fields] OR "contracept*"[All Fields])) OR (("non steroid*"[All Fields] OR "nonsteroid*"[All Fields]) AND "anti-inflammatory"[All Fields] AND ("agent"[All Fields] OR "agents"[All Fields])) OR "nsaid*"[Title/Abstract] OR "opioid*"[Title/Abstract] OR ("progestinic"[All Fields] OR "progestinics"[All Fields] OR "progestins"[Pharmacological Action] OR "progestins"[MeSH Terms] OR "progestins"[All Fields] OR "progestin"[All Fields]) OR ("dienogest"[Supplementary Concept] OR "dienogest"[All Fields]) OR "medroxyprogesterone acetate"[All Fields] OR ("leuprolide"[MeSH Terms] OR "leuprolide"[All Fields] OR "leuprorelin"[All Fields]) OR ("leuprolide"[MeSH Terms] OR "leuprolide"[All Fields]) OR "intrauterine device*"[All Fields] OR "levonorgestrel"[All Fields] OR "aromatase inhibitor"[All Fields] OR ("androgen s"[All Fields] OR "androgene"[All Fields] OR "androgenes"[All Fields] OR "androgenic"[All Fields] OR "androgenicity"[All Fields] OR "androgenized"[All Fields] OR "androgenizing"[All Fields] OR "androgenous"[All Fields] OR "androgens"[Pharmacological Action] OR "androgens"[MeSH Terms] OR "androgens"[All Fields] OR "androgen"[All Fields] OR "virilism"[MeSH Terms] OR "virilism"[All Fields] OR "androgenization"[All Fields]) OR ("danazol"[MeSH Terms] OR "danazol"[All Fields] OR "danazole"[All Fields]) OR ("laparotomy"[MeSH Terms] OR "laparotomy"[All Fields] OR "laparotomies"[All Fields]) OR "endometrium ablation"[All Fields]) AND ("dysmenorrhea"[MeSH Terms] OR "dysmenorrhea"[All Fields] OR "dysmenorrheas"[All Fields] OR "dysmenorrhoea"[All Fields] OR "pelvic pain"[All Fields] OR (("non-menstrual"[All Fields] OR ("nonmenstrual"[All Fields] OR "nonmenstruating"[All Fields])) AND ("pelvics"[All Fields] OR "pelvis"[MeSH Terms] OR "pelvis"[All Fields] OR "pelvic"[All Fields]) AND ("pain"[MeSH Terms] OR "pain"[All Fields])) OR "nmpp"[All Fields] OR ("dyspareunia"[MeSH Terms] OR "dyspareunia"[All Fields]) OR "endometriosis pain"[All Fields] OR "pain"[All Fields]) AND ("randomized controlled trial"[Publication Type] OR "randomized controlled trials as topic"[MeSH Terms] OR "randomized controlled trial"[All Fields] OR "randomised controlled trial"[All Fields] OR ("controlled clinical trial"[Publication Type] OR "controlled clinical trials as topic"[MeSH Terms] OR "controlled clinical trial"[All Fields]) OR "randomized"[Title/Abstract] OR "placebo"[Title/Abstract] OR "drug therapy"[MeSH Subheading] OR "trial"[Title/Abstract])) AND ((humans[Filter]) AND (2022/4:3000/12/12[pdat]))	

A 6. The clinical effectiveness SLR (Appendix D) mentions additional searches of grey literature (Trials databases and Google top up searches). Whilst search terms are provided, there is no record of the hits retrieved for each resource and only ClincalTrials.gov appears in the PRISMA flow chart. Please provide full details for each resource, including hits per line and an updated PRISMA flow chart.

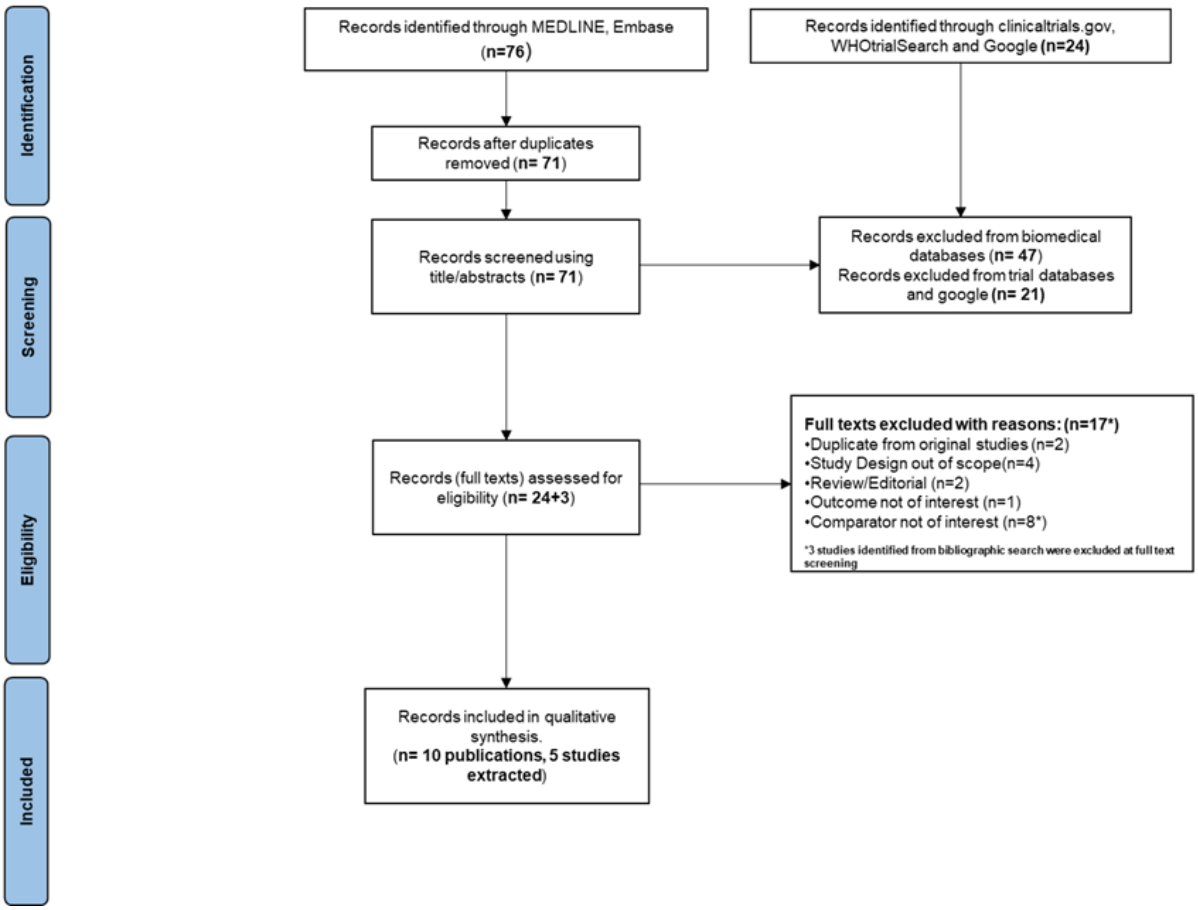
Company response: In the SLR update, searches of the grey literature retrieved the following numbers of hits; however none were deemed relevant for inclusion:

- WHOtrialSearch: 12 hits

- Google_GoogleScholar: 7 hits
- Clinicaltrials.gov: 5 hits

The figure below shows the PRISMA for the SLR update.

Figure 1: PRISMA diagram for the SLR update



A 7. Appendix G reports searches of EconLit, HTA database (Centre for Reviews and Dissemination; CRD) and NHS EED (CRD), however no search strategies are provided. Please confirm if these resources were searched and provide full search strategies (Please note that it is stated that these searches were used to inform all economics searches including health-related quality of life (HRQoL; Appendix H) and Resource use (Appendix I). Please provide updated PRISMA flow charts if required.

Company response: Owing to time constraints, EconLit was not searched, so its inclusion in Appendix G is an error for which we apologise. We can confirm that CRD was searched; 32 hits were retrieved but none were deemed relevant at screening.

The 32 hits were combined for economic evaluation, health care resource use and utility studies.

A 8. Appendix G also report searches of five conference proceedings:

a) International Health Economics Association

b) World Congress on Health Economics, Health Policy, and Healthcare Management

c) European Health Economics Association

d) American Society of Health Economists

e) The Professional Society for Health Economics and Outcomes Research

Whilst some search terms are provided (Table 118), there is no records of how many hits were retrieved per term or per conference. Please provide full details and amend any PRISMA flowcharts as required.

Company response: Both the original and updated SLRs reported number of publications rather than number of hits. PRISMA diagrams showing the study selection process for the economic evaluation, cost and resource use, and utility studies in the SLR update are shown below.

Figure 2: PRISMA flow chart for selection of economic evaluation studies

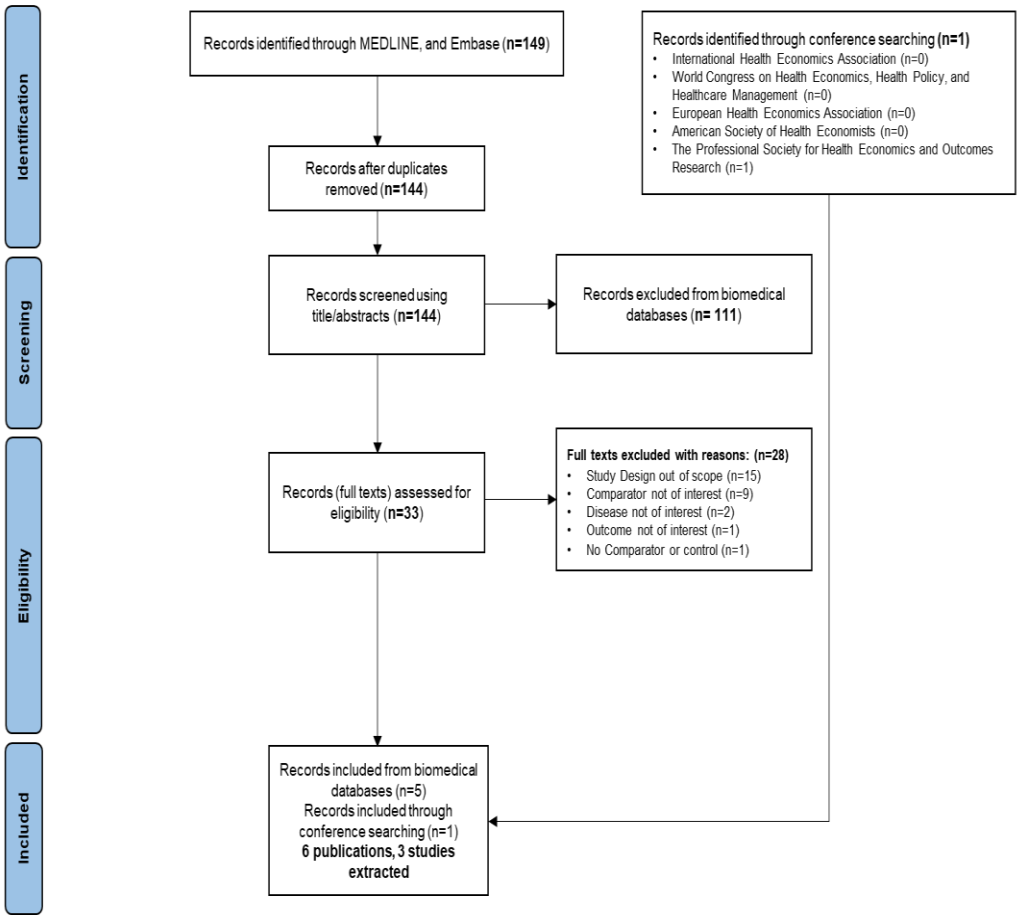


Figure 3: PRISMA flow chart for selection of cost and resource use studies

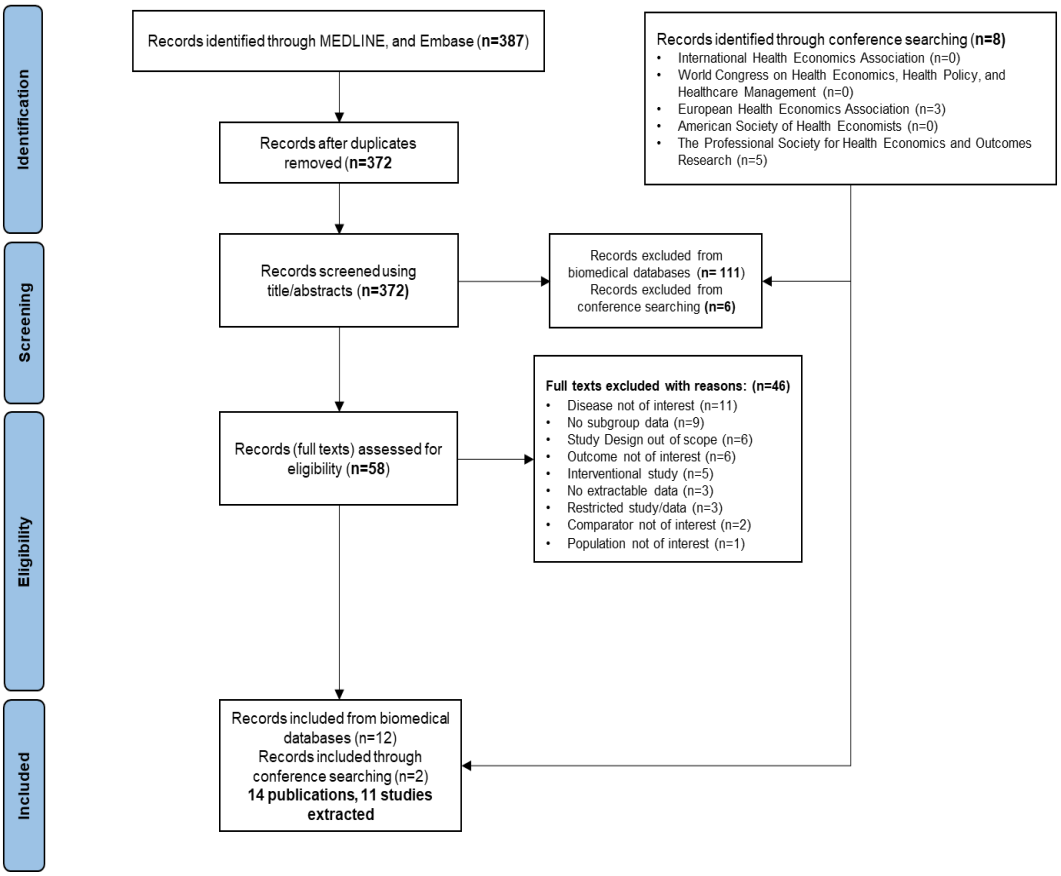
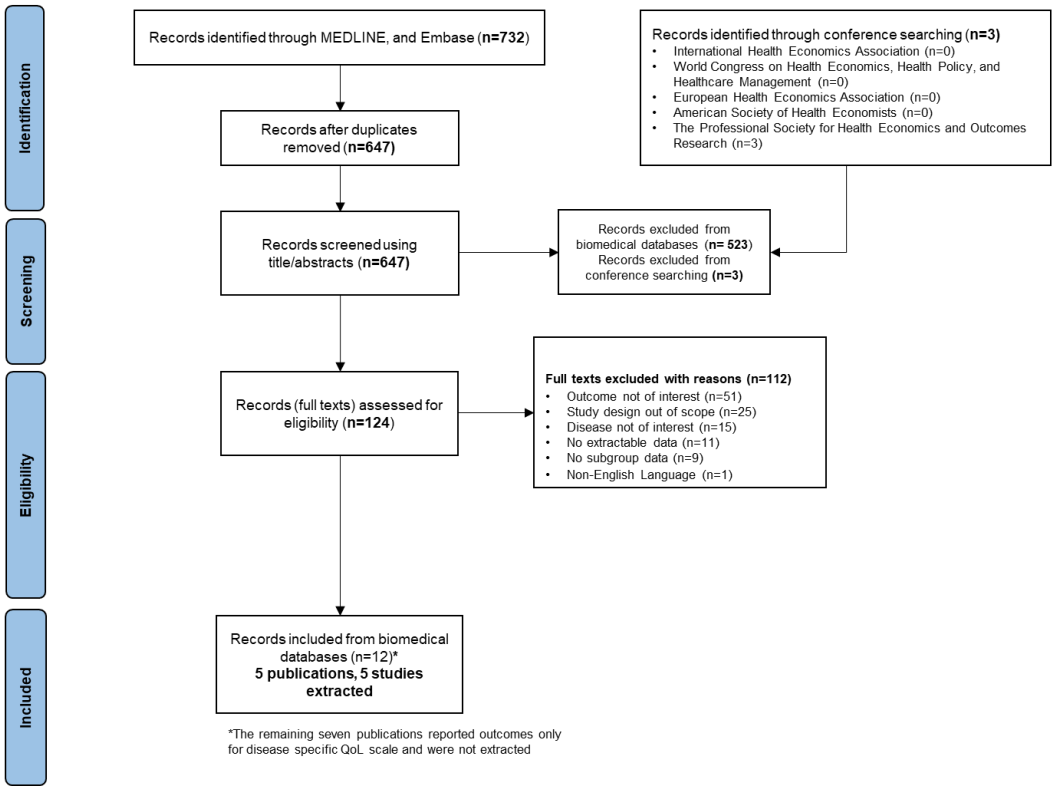


Figure 4: PRISMA flow chart for selection of utility studies



A 9. Please confirm the date span searched for all databases for all sections.

Company response: Date spans for the clinical and economic SLR updates are shown below.

- Clinical SLR update:
  - Embase searches: 01 November 2022 (01 April 2022 to 01 November 2022)
  - PubMed searches: 01 December 2022 (01 April 2022 to 01 December 2022)
  - Clinical pragmatic searches (trial databases and google): 12 December 2022 (01 April 2022 to 01 December 2022)
- Economic SLR update:
  - Embase and PubMed searches were run on 05 December 2022
    - Economic evaluation studies: last 5 years
    - Cost and resource use studies: last 5 years
    - Utility studies: last 10 years
  - Conference searches: 15 December 2022 for past 2 years (2020–2022)

A 10. Please clarify why there were language limitations in other (than PubMed and Embase) databases.

Company response: Language limitations were applied to reflect that the most relevant and high quality research is usually published in English for ease of accessibility to readers.

A 11. Please clarify why the Cochrane CENTRAL was not searched for relevant studies.

Company response: As outlined in our response to question A3, no separate search was carried out because Cochrane Reviews and Editorials are indexed in PubMed.

Decision problem (DP)

A 12. Priority question: The DP population in Table 1 of the company submission (CS) is stated to be the same as the scope. However, given the stated position of relugolix as second line as in the indirect treatment comparison (ITC) and economic analysis, please clarify that the DP population should be narrowed to only second line. If this is not the case, please conduct analyses (ITC and economic) at lines of therapy other than second line with the appropriate comparators.

Company response: Relugolix CT is indicated for symptomatic treatment of endometriosis in women with a history of previous medical or surgical treatment for their endometriosis. In line with this, and with current UK clinical practice, we anticipate that Relugolix CT will be used at second-line and therefore agree that the decision problem population should be narrowed to second-line.

A 13. Priority question: According to Section B.1.1. of the CS, gonadotropin releasing hormone (GnRH) agonists are listed as the relevant comparator for Relugolix combination therapy (CT; relugolix in combination with oestradiol and norethisterone acetate). However, more information on why nonsteroidal anti-inflammatory drugs (NSAIDs), neuromodulators, such as gabapentin and/or pregabalin, were not considered as relevant comparators

should be provided. Similarly, other established clinical management pathways of endometriosis i.e., hormonal treatments e.g., combined hormonal contraception (off-label for some combined hormonal contraceptives) and oral progestogens, were listed in the final scope issued by the National Institute for Health and Care Excellence (NICE), but not addressed in the CS.

If the choice of GnRH agonists as the relevant comparator is related to a restriction of the population to second line (see question A12), please explain why GnRH agonists have been chosen given that NICE guideline NG73 in the algorithm reproduced in the CS restricts use to a *"3 month course…before surgery"*. In fact, according to this algorithm, if initial management with analgesic or hormonal treatment (combined contraceptive pill or progestogen) is not effective, not tolerated or contraindicated, then surgery (excision or ablation) is recommended. Therefore, please conduct all analyses (ITC and economic) with the appropriate comparators, including analgesia or hormonal treatment if first line and surgery if second line.

Company response: As outlined in the response to question A12, Relugolix CT is indicated for symptomatic treatment of endometriosis in women **with a history of previous medical or surgical treatment** for their endometriosis (i.e. second-line). NSAIDs, neuromodulators, and surgical procedures would be used before Relugolix CT, and are therefore not considered relevant comparators.

Relugolix CT is the only oral GnRH antagonist licensed in the UK; there are no direct, licensed comparators. GnRH agonists are the closest comparator at second-line in the clinical pathway of care. As noted in the question, use of GnRH agonists is restricted to three months (or six months if add-back therapy is used); however they are often used (off-label) for longer durations in clinical practice. There is no restriction on duration of treatment with Relugolix CT.

A 14. Priority question: According to Section B.1.1. of the CS, hospital admission and fertility were not collected in the Relugolix CT trials. It is also unclear why other outcomes i.e., overall pain, recurrence of endometriosis, or complications of treatment are missing. Please clarify.

Company response: Gedeon Richter would like to clarify that Relugolix CT is not a disease-modifying drug; it relieves the symptoms of endometriosis rather than removing diseased endometrial tissue, for example. Therefore, it is not possible for endometriosis to 'recur' after treatment with Relugolix CT.

Overall pain associated with endometriosis (i.e. overall pelvic pain) was collected in the Relugolix CT studies (see Section B.2.6 in the company submission); its omission from the list of clinical effectiveness outcome measures in the decision problem table was an oversight.

The Relugolix CT trials included a comprehensive assessment of safety as measured by adverse events, clinical laboratory data, 12-lead ECGs, vital signs, physical examinations, menstrual bleeding patterns, pregnancy, overdose, bone mineral density, and paired endometrial biopsies, which would have captured any complications of treatment.

A 15. Priority question: Please explain the mismatch between different outcomes in clinical effectiveness, cost-effectiveness and ITC sections of the CS (Table 1, column *"Decision problem addressed in the company submission"*).

Company response: The outcomes listed for the clinical effectiveness section were all taken from the SPIRIT trials. Some of the outcomes in the cost-effectiveness section were taken from other sources as they were not available in the trials and are therefore not presented in the clinical effectiveness section. For example, pain recurrence following surgery was obtained from a study of post-surgery treatment outcomes, and complications of surgery were derived from a study on complications following hysterectomy.

The ITC focused on pain associated with endometriosis and included two outcomes: overall pelvic pain (which was an outcome in the SPIRIT trials) and total pelvic pain (a composite of dysmenorrhoea, non-menstrual pelvic pain and dyspareunia, each of which were individual outcomes in the SPIRIT trials).

The table below clarifies the outcomes in the submission.

Table 2: Outcomes used in the submission

Clinical effectiveness	Cost effectiveness	ITC
• Dysmenorrhoea • Non-menstrual pelvic pain • Dyspareunia • EHP-30 pain domain • Overall pelvic pain • Opioid use • Analgesic use • Health-related quality of life (EQ-5D-5L) • Adverse effects	• Dysmenorrhoea • Non-menstrual pelvic pain • Recurrence of pain • Analgesic use • Adverse effects of treatment • Subsequent surgical treatment • Subsequent medical treatment • Complications related to surgery • Health-related quality of life (EQ-5D-5L)	• Overall pelvic pain • Total pelvic pain (a composite of dysmenorrhoea, non-menstrual pelvic pain and dyspareunia)

Systematic review

A 16. Priority question: Appendix D indicates no restriction by line of therapy. However, analgesia is not listed as comparator. Various types of surgery are listed as comparators, but no study of surgery was included. Please conduct a systematic review that is consistent with the population and comparators of the DP as requested in questions A12 and A13 and where the studies included are consistent with the eligibility criteria.

Company response: The list of interventions and comparators was agreed with our Gedeon Richter global colleagues: only key terms related to surgical treatment actually used in the electronic search strategy were reported in the PICOS (i.e., Laparotomy and Endometrial Ablation Techniques); GnRH antagonists were removed from the updated SLR as those treatments are not available in Europe/UK (note that the original SLR was conducted with a US scope).

A 17. Priority question: Please clarify on differences/discrepancies between the original and the updated SLR in terms of interventions, comparators, and outcomes. E.g., an exclusion of GnRH antagonists in the updated SLR.

Company response: please see response to question A16.

A 18. Please clarify who adapted the quality assessment tool for studies included in the SLR and how the adaptation was done. Also, please mention whether one or two independent reviewers did quality appraisals.

Company response: We believe the EAG is referring to Appendix D1.3 of the company submission, which describes the quality assessment of the studies included in the ITC. This was done using the template in Section 2.5 of the user guide to the company evidence submission template as provided by NICE upon invitation to participate. The assessments were carried out by one independent reviewer and checked by a second.

Indirect treatment comparison (ITC)

A 19. Priority question: According to Section B.2.9. of the CS (Tables 26-27), only three studies were used for the ITC i.e., D'Hooghe et al. 2019, Lang 2018 and Strowitzki et al. 2010 for overall pelvic pain and total pelvic pain respectively. However, studies of Elagolix 150 mg and/or 250 mg or Linzagolix 50, 75, 100 and 200 mg are missing as relevant for the ITC. Similarly, it is unclear why all NSAIDs, neuromodulators, surgery (9 different procedures), aromatase inhibitors, androgenic drugs, gestrinone, selective oestrogen (or progesterone) receptor modulators (SORM/SPRM) were not included in the ITC. Please include ITCs with all relevant comparators that are consistent with the population and comparators of the DP as requested in questions A12 and A13.

Company response: As noted in our response to questions A12 and A13, Relugolix CT is indicated for symptomatic treatment of endometriosis in women **with a history of previous medical or surgical treatment** for their endometriosis (i.e. second-line). NSAIDs, neuromodulators, and surgical procedures would be used before Relugolix CT, and are therefore not considered relevant comparators.

Relugolix CT is the only oral GnRH antagonist; there are no direct, licensed comparators. GnRH agonists are the closest comparator at second-line in the clinical pathway of care.

A 20. Priority question: Odds ratios are calculated from standardised mean differences (SMDs) for the ITCs of OPP and TPP. Although this method is mentioned in the Cochrane Handbook, the Handbook states that it is *"…based on the assumption that an underlying continuous variable has a logistic distribution with equal standard deviation in the two intervention groups…The assumption is unlikely to hold exactly and the results must be regarded as an approximation."*

a) Please conduct these ITCs using the SMDs unconverted.

b) Please also conduct the ITCs using MDs for a measure of pelvic pain on a numerical rating scale (NRS) or visual analogue scale (VAS) with conversion between scales with upper limit at 100 or 10 as appropriate.

Company response: We would like to make the EAG aware that the ITC using odds ratios (ORs) has been updated as we have become aware of an abbreviated study report for D'Hooghe 2019 that is available on the study sponsor's website and reports actual values (2). The data inputs now reflect the use of the actual values and do not rely on digitization of figures from the publication as was the case in the original ITC. The full results are available in the ITC technical report, which we have included as part of this response.

As requested, we have also conducted analyses using SMDs and MDs as outcome measures. The full results are available in appendices to the ITC report, which we have included as part of this response; however we would like to draw the EAG's attention to the following:

- It was not possible to run the analysis using MD for the combined networks OPP and TPP as different scales have been used (B&B for TPP and NRS for OPP) and are not proportional so they cannot be compared - thus the use of Standardized Mean Difference (and OR)

- A few discrepancies in the results were observed when comparing OR vs SMD vs MD, however it is important to note that the new analyses were not adjusted for multiple testing, and the number of analyses has now substantially increased. The discrepancies shown below all involve the TPP endpoint which has limitations as discussed in question B11.
    - *TPP - Random effects - Empirical*
        - SMD Relugolix-CT vs placebo becomes unsignificant −0.55 (−1.4, 0.26)
        - Whereas it is significant for OR (0.37 (0.14, 0.96)) and MD (1.1 (0.22, 2.0))
    - *Combined OPP and TPP - using TPP from SPIRIT - Random effects - Empirical*
        - SMD Relugolix-CT vs placebo becomes unsignificant −0.60 (−1.4, 0.25)
        - Whereas it was significant using OR 0.34 (0.12, 0.96)
    - *Combined OPP and TPP - using OPP from SPIRIT - Random effects - Empirical*
        - SMD Relugolix-CT vs placebo becomes unsignificant −0.45 (−1.4, 0.45)
        - Whereas it was significant using OR 0.34 (0.12, 0.99)
    - *Combined OPP and TPP - using OPP from SPIRIT - Fixed effects*
        - SMD significant for dienogest vs Relugolix-CT 0.34 (0.087, 0.59) and leuprolide acetate vs Relugolix-CT 0.29 (0.025, 0.55)
        - Whereas it was not significant with OR: dienogest vs Relugolix-CT 1.5 (0.93, 2.5) and leuprolide acetate vs Relugolix-CT 1.1 (0.66, 1.9)

Regarding the assumptions related to the equation used to convert SMDs into ORs, $ln(OR) = \frac{\pi}{\sqrt{3}}SMD$, one can notice that standard deviations between the two interventions groups are of similar magnitude at a specific timepoint for all studies

included in the OPP (Table 9 of the ITC report) and TPP (Table 18 of the ITC report) analyses, as would be expected for randomized controlled trials.

Section B: Clarification on cost-effectiveness data

Literature review

B 1. Please clarify why the studies presented in Table 39 of the CS could not be used to inform some parts of the current economic analysis and model.

Company response: The studies presented in table 39 were not used to inform parts of the current economic analysis and model predominantly because they did not include the intervention technology (Relugolix CT) or the comparator of interest (GnRH agonist). This therefore limited greatly the number of generalisable inputs from these studies, i.e. we could not take any efficacy data, long-term treatment discontinuation data or incidence of treatment-related adverse events from any of these studies.

B 2. On page 136 of the CS, references 83-91 are cited for previous models that have evaluated treatment of endometriosis. Please clarify why this list of references is larger than those presented in Table 39.

Company response: We believe that the EAG is referring to the following statement on page 138 'The choice of a Markov model to evaluate the cost effectiveness of Relugolix CT is largely in line with previous models that have evaluated various interventions in the treatment of endometriosis (83-91)'. This list of references is larger than those presented in table 39 because the studies in table 39 are those that were specifically identified through the SLR detailed in Appendix G. The remaining studies were identified from alternative sources such as internet searches and were not identified systematically.

B 3. The EAG noticed that some clinical parameters such as the disutilities for headaches and hair loss are derived from relatively old studies. Please explain if that may be related to the EAG concerns on the appropriateness of the literature

searches (see above). Alternatively, please explain if there is a reason these studies were considered more appropriate over more recent ones.

Company response: The impact of disutility and costs of AEs on overall results is small. The difference in costs between Relugolix CT arm and the GnRH agonist arm is £7 favouring Relugolix while the disutility is 0.001 lower in the Relugolix arm compared to the GnRH agonist arm. Disutility for AEs was identified through a targeted literature review. As commented below, the probability of hair loss is set to 0% in both the Relugolix CT arm and the GnRH agonist arm so this does not impact the results. The difference in risk of headache between the Relugolix CT arm and the GnRH agonist arm is 0.04% favouring the Relugolix CT arm which again has a very low impact on overall results.

Population

B 4. Priority question: Please clarify whether the population demographics in the economic model (i.e., the trial population in SPIRIT 1 & 2) can be generalised to the patient population in England and Wales. Also, the decision problem described *"Adults with symptoms of endometriosis"* while on page 135 of the CS it is stated that the population considered are those with moderate to severe endometriosis-related pain. Please clarify how moderate and severe are defined in this context.

Company response: The generalisability of SPIRIT 1 and 2 to the patient population in England and Wales is discussed in section B.2.12 of the CS. As stated on page 132 of the CS, a clinical expert currently practicing in England who was consulted during the development of this submission stated that overall, the baseline demographics of patients in SPIRIT 1 and 2 were representative of the patients seen in clinical practice (81). However, the expert also noted that in some regions of the country, there may be a larger proportion of black patients than was included in the trials. Most patients enrolled in SPIRIT 1 and 2 were white (>90%), potentially reflecting under recognition or under diagnosis of endometriosis, or suboptimal clinical trial engagement among other races and ethnicities. This demographic makeup is consistent with the generally described epidemiology of endometriosis

although recent studies suggest there may be an ascertainment bias due to differences in the odds of endometriosis diagnosis by race and ethnicity (3-6). Gedeon Richter believe that the trial population in SPIRIT 1 and 2, and by extension the cost-effectiveness model, is broadly generalisable to the patient population in England and Wales.

We would like to apologise for the discrepancy in wording in B3 compared to the clinical sections of the CS. As stated in section B.1.2 (page 16 of the CS), [redacted] The wording on page 135 should have been updated in line with this and the correct model population is adult patients with symptoms of endometriosis who have a history of previous medical or surgical treatment for their endometriosis, as discussed in the response to clarification question A12.

Model Structure

B 5. Priority question: Regarding Figure 38 that presents the model structure:

a) Please explain the model structure in more detail, providing clear guidance on how patients move through the model from the beginning until the end.

b) Patients not responding to initial treatment are assumed to switch to best supportive care (BSC; treatment switch C in Figure 38). However, responders to BSC also seem to be assumed to go through surgery, via the waiting time before surgery health state, defined as treatment switch C. Please explain why responders to BSC are still assumed to go through surgery and provide further details on the exact path patients can follow following nonresponse to initial treatment.

c) The model assumes that patients undergoing conservative surgery and for whom pain recurs (treatment switch E in Figure 38; post conservative, PCS), may either receive BSC or undergo an additional surgery through the waiting time to surgery health state.

i. **However, for patients undergoing surgery as a subsequent treatment, there is also a waiting time of six months, in which time patients are assumed to again receive BSC. That would mean that all patients undergoing conservative surgery and experiencing pain recurrence, would receive BSC. Please explain what the discrepancy is between patients going to BSC directly or to BSC through the 'waiting time before surgery' health state. Is their response level different or is it the severity of pain that defines their path? Would it mean patients in the waiting time before surgery cannot respond to BSC anymore? Are these patients assumed to respond differently to BSC treatment?**

ii. **Please explain if there is a difference in the composition of the BSC in the two treatment pathways for PCS patients.**

iii. **If response levels are indeed different for the two BSC paths following conservative surgery as mentioned above (in question B5)i), please justify the validity of these assumptions.**

iv. **Patients experiencing PCS recurrence can transition back via section C to BSC. Please clarify if these patients are assumed to be identical to patients that did not have a surgery before. Please comment on the validity of this assumption. Also, please clarify why it is not needed in the model to "keep track" of any history of other surgeries.**

d) In treatment switch D in Figure 38, the last health state is called post-hysterectomy reoperation. Please clarify how often post-hysterectomy reoperation can occur and what happens after this to the patient (e.g., is it assumed that the patients will remain pain-free or something else).

Company response:

a) The model structure including the figure and the description of patients flow in the model have been updated in the newly submitted cost-effectiveness model.

b) All patients, independent of intervention start in Initial treatment (A in the figure). Treatment response is evaluated at six months and only patients who have obtained complete response remain on treatment. Patients without complete response move on to a subsequent treatment strategy. The options for subsequent treatment strategy include indeed BSC or surgery. Patients who switch to BSC (instead of surgery) are evaluated after three months and transit either into Response BSC or Non-response BSC. Patients with non-response move on to surgery, via the waiting time before surgery health state (Point D in model figure). Within the model horizon, the surgery would represent the third strategy of pain management after initial treatment (Relugolix CT or GnRH agonist) and BSC. Likewise, patients who respond to BSC may also eventually (in subsequent cycles) switch to surgery due to the loss of effect over time. This is captured by the discontinuation rate of BSC. But, *a priori* patients with response to BSC remain in that health state.

c) i) Patients who opt to undergo surgery are assumed to receive BSC to not be left untreated as BSC may still provide some pain relief albeit not sufficient to provide a pain-reduction equivalent to complete response. This captures the fact that many patients have already tested hormonal therapy prior to initial treatment with Relugolix CT or GnRH agonist. In case patients experience recurrence of pain following conservative surgery, they may switch to BSC and have a positive probability of obtaining complete response from BSC (point G in the model figure) and transit into BSC Response. The model assumes the same probability of response as to those who switched directly to BSC instead of surgery. This is likely a conservative approach.

ii) No, both pathways are assumed to receive the same type of BSC.

iii) Response levels are different. No complete response from BSC in waiting time before surgery is expected while patients following recurrence of pain post conservative surgery may obtain response from BSC. This is a conservative approach due to lack of evidence of level of response for patients who suffer recurrence of pain following conservative surgery. This approach increases the ICER of Relugolix CT.

d)The probability of re-operation is 10% and it is then assumed that patients will remain pain-free.

B 6. Priority question: Please justify why the current time horizon used by the company in the base-case analysis (i.e., 16-year time horizon) is not a lifetime time horizon. Please adjust the model to allow for a lifetime time horizon or, in other case, please discuss the expected impact on the cost-effectiveness results of using a lifetime time horizon. Please compare this approach with the approaches in similar submissions where a lifetime time horizon may have been considered (e.g., NICE TA832, also a relugolix TA).

Company response: The growth of the endometriotic tissue is estrogen-dependent and is associated with the time between menarche and menopause. The time horizon was therefore set to 16 years in the initial submission, age at which patients reaches menopause. The time horizon was therefore deemed long enough to reflect all important differences in costs or outcomes between the technologies being compared. The model has now been updated to be able to capture a lifetime approach and a post-menopause health state has been introduced. The user can now set the age of the cohort at the entry in the model (Input sheet!D47), the age of the cohort reaching menopause (Input sheet!D21), and the time horizon (Input sheet!D20). We have updated the base-case such that costs and outcomes are measured over a lifetime horizon.

Associated utility and cost for the post-menopause health state can be adjusted by the user. Currently, the health state cost is set to £0 while the utility is set to 1. The utility value of menopause as well as for all other health states is now adjusted by an age-factor as suggested by the EAG (question B20). The utility of post-menopause (1.000) is multiplied by the adjustment factor so that patients following menopause obtain the utility of the general population. The health state cost for the post-menopause state been set to £0 since no direct medical cost related to endometriosis is assumed to occur beyond menopause.

B 7. Priority question: Considering the patient population is adults with moderate to severe endometriosis-related pain who have a history of previous medical or surgical treatment, fertility, and hospital admissions would be expected to be clinically and economically relevant outcomes for the economic analysis. Similarly, overall pain, recurrence of endometriosis, or complications of treatment are missing from the outcomes list as mentioned in question A14. Please discuss the impact on the cost effectiveness results of not considering fertility, hospital admissions and the other outcomes in question A14 in the economic analysis.

Company response: We consider each of these outcomes in turn below:

Admission to hospital: the majority of hospital admissions are related to procedures, which are already captured in the model. In a recent Australian report, in 2021–22 there were 40,500 endometriosis-related hospitalisations compared with 3,600 endometriosis-related emergency department presentations (7, 8). The contribution of emergency admissions to cost effectiveness is therefore considered to be small.

Fertility: Disutility from infertility would only be expected to have an impact on the proportion of people actively trying to have a family. Both GnRH and relugolix are contraceptive, and the disutility of infertility related to this would already have been captured within the trial EQ-5D values given that the women participating in the trials would have been aware of this. A utility benefit for faster recovery of fertility following discontinuation of Relugolix CT was considered too uncertain a parameter to include and would likely have little impact in results, given that the difference in time to regain in fertility between the two treatments is likely to be months rather than years.

Recurrence of endometriosis: Gedeon Richter would like to clarify that Relugolix CT is not a disease-modifying drug; it relieves the symptoms of endometriosis rather than removing diseased endometrial tissue. Therefore, it is not possible for endometriosis to ‘recur’ after treatment with Relugolix CT.

Overall pain: The model uses utility values collected directly from the clinical studies, pain is by definition therefore captured in the model.

Complication of treatment: adverse events of treatment were captured in the economic analysis, as were the costs and disutilities of surgical interventions.

B 8. Priority question: Please confirm if the partial response health state is only active if response is assessed at 3 months and thus in the base case analysis there are 11 health states considered. In addition,

a) Please explain what the rationale is for including partial response as an additional health state in the model.

b) Please explain what it would mean for patients with endometriosis if that option was selected in the model.

c) Please provide a scenario analysis on the impact of using this option in the model.

Company response: We would like to apologise for the unclear and slightly inaccurate description of the partial response health state in the CS. Partial response is defined against the same definitions as complete response (change from baseline response; A NRS score reduction of either 2.8 for dysmenorrhea or 2.1 for NMPP and no increase of analgesic use) but relates to when patients have responded on either dysmenorrhea or NMPP after 3 months. At 6 months, patients can only be complete responders (i.e., response on both dysmenorrhea and NMPP). The partial response state is only active if treatment is evaluated at six months and thus it is considered in the base-case. We erroneously stated in Document B that partial response would only be considered if outcomes are assessed at month 3 and would like to clarify that this is not the case.

In response to question 8a, the rationale is that clinicians would see patients at three months and leave them on treatment if any sort of response was observed. The partial response therefore captures patients with response to *either* NMPP or dysmenorrhea. We also account for the proportion of complete responders (NMPP and dysmenorrhea) at 3 months.

In response to question 8b, if 3 months is selected as the timepoint at which response to assessment is assessed, then this would mean that patients who are

non-responders to treatment (Relugolix CT or GnRH agonist) would discontinue treatment earlier than they would if response is assessed at 6 months. Furthermore, there is no partial response health state included if 3 months is selected as the timepoint for response assessment, as any patient who fails to respond to both the NMPP and dysmenorrhea criteria would be characterised as a non-responder and would thus discontinue treatment.

In response to question 8c, the results of the scenario where response is assessed at 3 months, and there is therefore no partial responder health state, are reported below.

Table 3: Cost-effectiveness results, scenario where treatment response is assessed at 3 months

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)
Relugolix CT	£10,849	23.10	16.92	-	-		
GnRH agonist	£10,057.56	23.10	16.44	£791	0.01	0.48	£1,661

Comparator

B 9. Priority question: Please provide a clear definition of <u>all</u> comparators included in the model, as this is unclear throughout the CS.

a) On page 136 for example it is mentioned “The model compares treatment with Relugolix CT to alternative treatments that are currently available”, whereas on page 142 it is mentioned that “The modelled comparator is GnRH agonist”. Also on page 137, BSC is included in the list of initial treatment options like Relugolix CT and GnRH agonists, whilst it is not included in the comparator list.

b) The electronic model includes BSC, GnRH and surgery as relevant comparators. Please explain the role of BSC (including a clear definition of what constitutes BSC) and surgeries in the model, providing clear definitions. As also questioned in A13, the company may have missed

relevant comparators from the economic analysis. Please explain why BSC and surgery are listed as comparators in the electronic model but not in the CS documentation.

Company response: The cost-effectiveness model that has been submitted as part of this appraisal is an adaptation of a global cost-effectiveness model and as such, includes certain functionalities that are not applied in the NICE base-case or any of the scenarios included in the submission. This includes additional comparators such as BSC and surgery. However, the only comparator included in this NICE submission for Relugolix CT is GnRH agonists. As discussed in response to clarification question A13, Relugolix CT is indicated for symptomatic treatment of endometriosis in women with a history of previous medical or surgical treatment for their endometriosis (i.e. second-line). NSAIDs, neuromodulators, and surgical procedures would be used before Relugolix CT, and are therefore not considered relevant comparators. Relugolix CT is the only oral GnRH antagonist; there are no direct, licensed comparators. GnRH agonists are the closest comparator at second-line in the clinical pathway of care. Both BSC and surgery are subsequent treatment options for patients who discontinue either Relugolix CT or GnRH agonist and they should be considered only in this context. That is, the additionally functionality in the model which includes BSC and surgery as comparators is not relevant to this submission.

B 10. Priority question: For BSC, many parameters in the model are assumed to be informed from the placebo arm of the SPIRIT 1 & 2 trials (for example the response rates, the rates of AEs, and health state utilities for BSC). Moreover, on page 150 of the CS, for the AEs of GnRH agonist it is stated that "*the risk ratios for GnRH agonist alone vs. placebo (i.e., BSC) were applied to the placebo data from the SPIRIT 1 & 2 trials".*

a) Please clarify if the composition of the placebo arm of the SPIRIT 1 & 2 trials is the same as the composition of BSC in England and Wales. If that is not the case, please clarify on what evidence basis it could be assumed that

the various input parameters for BSC are equal to the values seen in the placebo arm of the SPIRIT trials.

b) In case the placebo arm of the trials is not equivalent to BSC in the UK, please use other appropriate methods (e.g., ITC or scientific literature) to inform all BSC parameters in the model that are currently informed from the placebo arm of the SPIRIT trials. Please run scenario analyses using these parameters.

c) The ITC presented in section B.2.9 of the CS includes dienogest in the network for total pelvic pain (TPP). However, it seems that these results were not used to inform the clinical effectiveness parameters of BSC. Please explain why this evidence was disregarded. Note that based on this ITC, placebo and dienogest do not seem to be equally effective (despite showing overlapping confidence intervals). Also, the forest plots presented as results of the ITC report odds ratios above 1, favoring thus dienogest over relugolix CT. Irrespective of significance level, please adjust the model to appropriately incorporate these ITC results for BSC versus Relugolix CT (i.e., the deterministic model should be based on the ITC point estimate and the PSA should include the confidence interval) and run scenario analyses using these results.

Company response:

a) In our original CS, we had described that BSC in the model was comprised purely of treatment with Dienogest, however this was a carry-over from the original global model developed by the company and is not reflective of BSC in England and Wales. BSC in England and Wales for this patient population is comprised of symptomatic treatment for pain management, such as analgesics, which is the same as the definition of BSC in SPIRIT 1 & 2. We have thus updated the model to reflect this.

b) As discussed above, the placebo arm of the SPIRIT 1 & 2 trials is the same as the composition of BSC in England and Wales.

c) Please see our response to clarification question **Error! Reference source not found.**for the description of the incorporation of ITC estimates in the model. However, as stated in response to B10a) the composition of BSC in England and

Wales is the same as the placebo arm of SPIRIT 1 & 2 trials, and the evidence from the placebo group was therefore used for efficacy inputs related to BSC.

Clinical effectiveness parameters

B 11. Priority question: Many effectiveness parameters in the model are assumed to be equal between GnRH agonist and Relugolix CT. Please clarify where the benefits of Relugolix CT over the comparator are expected to be observed and how these benefits were included in the model. Please explain whether ITC results were included in the model and how. In general, please do not assume equal effects in the model based on non-significant results. Instead, for all parameter where equal effects were assumed, please implement a deterministic model based on the ITC point estimate (obviously translating the odds ratio to a relative risk first) and probability sensitivity analysis (PSA) including the ITC-derived confidence intervals.

Company response: The cost-effectiveness model has been updated to incorporate the ITC point estimate (via the odds ratio) as well as confidence/credible intervals in the PSA. As detailed in the ITC report, several limitations exist in the network of total pelvic pain (TPP) and the analysis of the TPP network could be prone to major risk of biased results due to:

- the high risk of performance and detection bias for the Strowitzki et al. 2010 study, connecting leuprolide acetate to dienogest, due to the absence of blinding as showed in the risk of bias assessment (see Table 17 of the ITC report),
- the between-study differences with differences in baseline pain score, prior surgical treatments, and concomitant use of NSAIDs,
- as well as the difference in race with Lang et al. 2018 study, connecting dienogest to placebo, having been conducted on Chinese women which could lead to population bias.

We have updated the model with the point estimates from the ITC. The results from the analysis of the network for overall pelvic pain, with random effects and weakly informative priors, were used to derive the response rates of GnRH agonists using an odds ratio of 1.1 [CrI: 0.032; 41] comparing Relugolix CT to Leuprolide acetate as displayed in the forest plot Figure 4 of the ITC report.

B 12. Priority question: On page 149 of the CS, it is stated that "*the GnRH agonists that are licensed for the treatment of endometriosis in the UK are leuprolide acetate, goserelin, triptorelin, nafarelin and buserelin*". Nafarelin and buserelin are not used as based on clinical experts experiencing that patients who are currently receiving GnRH agonist treatment in the UK opt for the subcutaneous formulations.

a) Of the subcutaneous GnRH agonists, only leuprolide is included in the ITC presented in section B.2.9 of the CS. Please clarify why the other treatment options were not part of the ITC. Was it because they could not be directly or indirectly connected to the network, or for any other reason?

b) Treatment effectiveness of the GnRH agonists is assumed equal to the effectiveness of the Relugolix CT based on the non-significant outcomes of the ITC. However, all forest plots presented as results of the ITC report odds ratios above 1, favoring thus leuprolide acetate over relugolix CT. Please adjust the model to appropriately include the results of the ITC analyses for Relugolix CT versus GnRH agonists irrespective of the significance levels (i.e., the deterministic model should be based on the ITC point estimate and the PSA should include the confidence interval).

Company response:

a) Only studies with leuprolide acetate as GnRH agonist treatment were identified for the network on overall pelvic pain (OPP – Table 3 of the ITC report). For the network of total pelvic pain (TPP), studies with other GnRH agonist treatment option were identified but could not be connected to the network (Table 12 of the ITC report).

b) Please see company response to clarification question B 11.

B 13. Priority question: On page 143 of the CS, it is mentioned that *"The model allows for the selection of two possible definitions of complete response to treatment"*.

a) Please clarify (and explain why) which one can be assumed to be representative of clinical practice in England and Wales, and select this for the base-case.

Company response: the feedback from three advisory boards (one European, two UK) is that it is difficult to define a responder due to patient heterogeneity in what they perceive as relief of symptoms, and lack of a standardised clinical scale in clinical practice:

- "*In clinical practice, for any QoL measure, a response is based on whether the patient is satisfied with the treatment*";
- "*In clinical practice, QoL as defined by the patient would be used to determine response rather than the numerical rating scale*";
- "*The decision to continue treatment with Ryeqo® would be a result of a discussion between the healthcare provider (HCP) and the patient*";
- "*Response may be measured in different ways and scores, and improvement in any domain may be enough to continue treatment, especially in the beginning of treatment*";
- UK clinician: "*the decision to continue will be likely made between HCP and patient based on response and side effects*";

In the UK HTA advisory board (which included three clinicians), no response criterion was offered by the clinicians, therefore a suggestion was made by a health economist to remove the existing continuation rules altogether and/or include a flexible stopping rule based on a change from baseline in pain score. This option would have required restructuring the model and respecifying the utility regressions to use a continuous pain score, which unfortunately was not feasible within the timeframe of the submission.

In the European advisory board (which included 1 UK clinician), it was agreed that a threshold for response was hard to determine, as the effect is also dependent on e.g., the level of pain at baseline, and patients may respond to treatment well but not reach a certain absolute threshold. However, the consensus was that achieving a threshold below 4 in NRS scale (mild pain) in both NMPP and dysmenorrhea at 6 months may be used for complete response. To explore this further, further details of the trial primary endpoint was sent to the attending clinicians along with choice of 3 options:

- Both options would be considered as relevant as definition of response in clinical practice
- Only the NRS score reduction of 2.8 both for dysmenorrhea and 2.1 for NMPP would be considered as relevant as definition of response in clinical practice
- Only the threshold below 4 in NRS scale would be considered as relevant as definition of response in clinical practice

All clinicians subsequently selected the “Both” option.

Given the lack of sensitivity of the model to response criteria (a difference of <£100 in the ICERs) and that both could represent reasonable scenarios, the decision was made to use the trial responder criteria in the base case for consistency with the clinical data.

b) The option of the “threshold” response is stated to be defined using clinical expert opinion during a global advisory board HCP (not defined in the CS, presumably healthcare professional) meeting. Please provide a full reference for the advisory board meeting, including participant list and minutes of the meeting as this meeting is referred throughout the CS, but a reference seems to be missing (e.g., also on page 136, the CS states that “the model structure was also validated with clinical experts during the global advisory board.”) Please provide the documentation including the minutes of the meeting and participants of the global advisory board. Also, would this statement mean that the option of threshold response is more representative for England and Wales as it is informed using clinical expert opinion?

Company response: The advisory board reports have been provided as part of this response. As explained in part a), the UK clinicians could not reach a consensus regarding responder criteria. As also explained in part a), the two different criteria have little impact on model results and both may represent reasonable scenarios.

B 14. Priority question: In the base case analysis, patients on Relugolix CT will continue treatment until response, discontinuation or until the age of menopause. That also indicates that response to Relugolix CT treatment is assumed to be constant over time (and it does not appear to be an option in the model structure [figure 38] for people to move from complete response to partial or non-response). Please comment on the validity of this assumption. Please present scenario analyses exploring the effect of long-term treatment effect waning of Relugolix CT treatment.

Company response: The efficacy of Relugolix CT was demonstrated in the SPIRIT OLE study. At Week 52, 84.8% (95% CI: 80.06, 88.85) of patients met the dysmenorrhea responder definition and 73.6% (95% CI: 68.04, 78.74) of patients met the NMPP responder definition. These rates of response were sustained through Week 104/EOT: 84.8% (80.06, 88.85) for dysmenorrhea and 75.8% (70.33, 80.74) for NMPP. Waning of treatment effect is captured through the discontinuation rates applied in the model. At discontinuation, patients move from complete response to non-response.

B 15. Priority question: On page 148 of the CS, it is mentioned that "*The model base-case assumes that the treatment duration of GnRH agonists is capped at one year*".

a) Please clarify what happens in the economic model after one year.

Company response: Upon discontinuation of GnRH agonist treatment, patients' endometriosis symptoms are managed through subsequent treatment (BSC) or by surgery (conservative surgery or hysterectomy). The split of patients amongst BSC and the different types of surgery is detailed in section B.3.3, Treatment distributions of the CS.

b) Please explain why all patients in the responder group after 1 year move to the non-responder group for 1 cycle, rather than moving them to the 2nd line treatment immediately.

Company response: This accounts for the time spent by patients and their clinicians to assess which subsequent treatment patients will receive, i.e. whether they will switch treatment to BSC or undergo surgery to manage endometriosis.

c) Please explain how to interpret discontinuation rates for GnRH agonist after one year in Table 48.

Company response: In the model base-case, given that GnRH agonist treatment is restricted to one year, the discontinuation rate at months 12 and beyond (i.e. at months 15, 18 etc) will be 100%. However, if the duration of GnRH agonist treatment is increased to 2 years for example, as it is in a scenario analysis reported in section B.3.9, then the treatment discontinuation rates reported in table 48 will apply. That is, discontinuation at months 12, 15, 18 and 21 would be 12% and would be 100% at month 24, as treatment duration is capped at 24 months.

B 16. Priority question. On page 147 of the CS, it is mentioned that the discontinuation rates for Relugolix CT and BSC were based on post-hoc analysis of discontinuation data from the SPIRIT open-label extension (OLE) study. However, Table 87 of the CS indicates that data on discontinuation were derived from the SPIRIT 1 & 2 trials for Relugolix CT. Please clarify the discrepancy in these statements and explain why discontinuation data were informed from the OLE study and not from the SPIRIT 1 and 2 trials if that is the case.

Company response: We apologise for the discrepancy in the reporting, the correct reference for discontinuation is the SPIRIT OLE study. The duration of the SPIRIT 1 and 2 studies was 24 weeks whereas the duration of the SPIRIT OLE study was 104 weeks, thus providing a longer duration over which treatment discontinuation could be observed and hence why this was chosen as the source of discontinuation data in the model.

B 17. Priority question: On page 151, the company states that *"AEs are weighted according to the proportion of patients that receive add-back therapy"*. As

in the base case analysis 100% of patients in GnRH agonist arm is assumed to receive add-back therapy:

a) Please explain what the proportions of patients receiving add-back therapy are and how these weights are exactly used in the calculations (base case or scenario analysis).

Company response: All patients (100%) in the Relugolix CT arm and the GnRH agonist arm use add-back therapy. The probability of a particular AE is weighted between the probabilities of GnRH agonist monotherapy and GnRH agonist with add-back therapy using the share of patients using add-back therapy. In the base case analysis, only probabilities of AEs for GnRH agonist with add-back therapy would then be used in the calculations.

b) Also, Table 52 is presenting "total probability for AEs", whereas Table 53 is presenting "3-monthly probabilities for AEs". Please explain what "total probabilities for AEs" means.

Company response: Please see response to part c below.

c) Please explain how probabilities in Table 53 are calculated based on Table 52 and provide an example of the computations.

Company response: The probabilities of Relugolix CT and BSC in Table 52 represents those observed in the SPIRIT 1 & 2 trials. The AE profile for GnRH agonist was derived by applying risk ratios to the probabilities for AEs linked to BSC/placebo arm from SPIRIT 1 & 2. If an AE is assumed to take place at treatment initiation, then "acute" is selected as risk type and the number of observed adverse events is applied to the first model cycle and patients are thereafter assumed to not experience any further AEs. If risk type "constant" is selected, the number of observed events in SPIRIT trial is re-calculated to a 3-month probability taking into account that AEs in SPIRIT 1 & 2 trials were observed during the duration of the trial (24 weeks) until the safety follow-up visit approximately 30 days after the last dose of study drug so 24 weeks plus 30 days in total.

d) In Table 52, the total probability for AEs for patients on GnRH agonist monotherapy and GnRH agonist therapy in combination with add-back

therapy are presented. Please explain exactly how these parameters are used in the model computations to define AEs.

Company response: The probabilities are weighted by the share of patients using add-back therapy and are thus set to 100% in the base case analysis. Hence, the values in the column labelled 'GnRH agonist (monotherapy)' have no impact on the results in the current base case where all patients on GnRH agonists use add-back therapy.

e) According to Table 52, decreased libido and hair loss are assigned to a probability value of 0%, whereas costs related to treatment of these AEs are non-zero (as well as disutility for hair loss), please explain if these two AEs are eventually considered (or not) in the model calculations.

Company response: The probabilities of decreased libido and hair loss have been set to 0% and are therefore not considered in the model calculations. That is, the disutility and costs associated with these adverse events are not incurred in the model.

B 18. Priority question: Regarding the risk of experiencing cardiovascular (CV) events, major osteoporotic fractures, and death:

a) Please clarify if and to what extent Relugolix CT and the comparator are expected to increase or decrease the risk of experiencing CV events, major osteoporotic fractures, and death (through other ways besides these two events) at 1 year and beyond 1 year.

Company response: An increase in total cholesterol levels and a decrease in high-density cholesterol levels was observed in the Relugolix CT arm in SPIRIT 1 & 2. Both changes are associated with an increased risk of CV events as estimated by the Framingham prediction model. Patients on GnRH agonist were also assumed to experience the same changes in total cholesterol levels and high-density cholesterol levels as patients in the Relugolix CT arm. No change in BMD was observed in the Relugolix CT arm in SPIRIT 1 & 2 and therefore no excessive fracture risk was estimated in the model. Patients on GnRH agonist were also assumed to maintain their BMD levels given the use of add back therapy. The risk of CV events and fractures are only applied while on treatment. A population-based risk of mortality is

applied and modelled as a 3-month probability of death (age-dependent). Patients who undergo a surgery also face an additional risk of death. This risk is limited to the subsequent cycle following surgery.

b) Please also explain how mortality has been operationalized in the economic model.

Company response: Please see above in response to part a).

c) On page 154, it is mentioned that "in the base case analysis, all patients are assumed to be initiated on add-back therapy treatment and thus no decrease in BMD [bone mineral density] and no excess fracture risk is applied". Please explain if this sentence means that a potential change in both, BMD and risk of fracture, is eventually not included in the base case analysis neither for Relugolix CT nor for the comparator treatment.

Company response: That is correct, please refer to a) for further details.

d) Please run a scenario analysis including the impact of Relugolix CT treatment and comparator(s) on BMD and excess fracture risk.

Company response: The results for the requested scenario are reported below. The results for this scenario are extremely similar to the updated model base-case results, which is not unexpected. This is because the underlying risk of low-energy fractures is very low in the model population. Moreover, any difference between the two comparators only appears after one year, when those who were allocated to GnRH agonists have discontinued treatment. However, at that stage some patients will have also been taken off treatment with Ryeqo which further reduces the differences between the two arms.

Table 4: Cost-effectiveness results, scenario including the impact of Relugolix CT treatment and comparator(s) on BMD and excess fracture risk

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)
Relugolix CT	£11,486	23.11	17.17	-	-	-	-
GnRH agonist	£10,279	23.10	16.46	£1,207	0.01	0.70	£1,715

e) Please explain if any excess mortality due to CV events has or has not been incorporated in the model and how. In case excess mortality due to CV events is not included, please justify this choice and explain the potential impact.

Company response: No excess mortality due to CV events is included in the model. The number of CV events is low (0.003 in the Relugolix CT arm) which is a reflection of the absence of males and high age (two well-established risk-factors for CV events) in the patient population. The impact of omitting excess mortality due to CV events is therefore minimal.

f) In terms of long-term impact of Relugolix CT treatment, if the treatment is expected to be used until menopause while it can increase the risk of CV and fractures, these risks would also be expected to be higher and relevant for patients around or above the age of menopause. Please comment on this expectation. Note that if there are risks above the age of menopause, these are not currently captured in the model.

Company response: The increased risk of CVD (and fractures) is associated with the use of Relugolix CT treatment and therefore is assumed to disappear following menopause when all patients stop treatment. While a residual risk of CV events may remain even after menopause, this is assumed to have little impact on the ICER.

g) Please clarify how the equations of the CV and fracture have been incorporated into the model. Are the parameters informed from the trials and are they constant or do they change across the model cycles (time horizon)?

Company response: The risk calculation for CVD events uses the Framingham Risk Function. The coefficients included in the cost-effectiveness model are based on those obtained from the original publication and are constant over the model time horizon. The value of variables is also constant and are based on those observed at baseline in SPIRIT 1 & 2 with the exception of age which is updated each cycle.

h) Please provide detailed examples of the risk calculations for CV events and fractures for patients at 1, 5, 10 and 15 years in the treatment arms.

Company response: The requested calculations are provided in the supplementary excel spreadsheet submitted with our responses.

l) Please clarify why on sheet ‘CV and fx calculation’, cell D8, the value of systolic blood pressure (SBP) has been hard-coded?

Company response: The share of patients treated with statin was not available in the SPIRIT trials and therefore omitted from the calculation. The effect of systolic blood pressure is still captured through the variable SBP(untreated).

Health-related quality of life

B 19. Priority question: Please discuss the (face) validity of the EQ-5D values presented in Table 59 and the disutility values presented in Table 64 (e.g., compare the values presented in this submission with other sources of utilities for this or similar diseases – e.g., studies retrieved by the SLR, and with the utility values for the general population – also indicate if these were validated with clinical experts and how). The EAG is concerned for example that the baseline utility value (in Table 59) seems low and the disutility for headache (in Table 64) high. Please also explain how the utilities are adjusted to each model cycle (3 months).

Company response: The baseline utility value in Table 59 is within the range of those identified in the SLR of QoL data and reported in Table 121 of the company submission, in which baseline values ranged from 0.15 to 0.689 pre-surgery. In a study by Grundstorm et al. of biopsy-confirmed endometriosis in women with moderate to severe pelvic pain the baseline EQ-5D was 0.45 (9). The highest baseline value observed was 0.78 in a prospective observational study in France, however these women did not have moderate-severe symptoms of endometriosis and were being treated with high-dose progestin, thus can be considered a less severe cohort than that treated with Relugolix CT.

Disutilities were not validated with clinical experts, however the sensitivity of the model to these inputs should be considered; removal of the disutility of headache does not change the ICER and removal of them altogether increases the ICER by only £1.

The utility and disutility values are adjusted to the cycle length in the calculation-sheets (Result calculation comp#) where the sum of the discounted QALYs over the time horizon is divided by 4 (Columns HC to JC, row 106).

B 20. Priority question. Please include in the model utility decrements with age, based for example on Ara and Brazier 2010 (DOI: 10.1111/j.1524-4733.2010.00700.x).

Company response: The model has been updated to include age-related utility decrements, sourced from Szende et al., 2014 (10).

B 21. Priority question: The model does not account for the number of treatment failures as the utility of non-response following Relugolix CT treatment is used equally for both non-responders in the initial treatment, and non-responders in the subsequent BSC treatment, post-hysterectomy recurrence and post-conservative surgery recurrence. In that respect, the company states that 'to not decrease further the utility after failing another line of treatment is likely a conservative approach'. However, also for the response health states the utility estimated in the initial response health state of the Relugolix CT arm is assumed to be the same for all response health states, i.e. *"Response"*, *"Response BSC"*, *"Post-hysterectomy stable"*, and *"Post-conservative (PCS) response"*. Please comment on the validity of this assumption and the potential impact on the results. Would this assumption also be considered conservative as indicated for the non-response patients? Would this assumption be expected to favour patients in the Relugolix CT arm versus the comparator arm?

Company response: The model has been updated by introducing individual utility-values to each health state. However, due to lack of evidence, it is still assumed in the base case analysis that the utility of response is the same for all response health states, i.e. "Response", "Response BSC", "Post-hysterectomy stable", and "Post-conservative (PCS) response". Since previous treatment failures are likely to contribute negatively to a patient's quality of life, assuming the same utility for all response health states is a conservative approach which is not expected to favour patients in the Relugolix CT arm versus comparator arm. A scenario analysis was conducted where we set the health state utility values for subsequent treatment

(rows 322-333 in the 'Input sheet' worksheet) equal to the utility value of initial treatment health state (0.584). The results for this scenario are presented in the table below.

Table 5: Cost-effectiveness results, scenario assuming health state utility for subsequent lines of treatment are equal to initial treatment

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)
Relugolix CT	£11,486	23.11	15.49	-	-	-	-
GnRH agonist	£10,279	23.10	13.63	£1,207	0.01	1.859	£649

B 22. Priority question: A long-term disutility of 0.180 is applied to post-hysterectomy to account for people who would no longer be able to get pregnant. This disutility value is informed by a global burden of disease report published by the World Health Organization and represents the disutility linked to infertility. Please comment if this value is plausible as infertility linked to hysterectomy may not be a concern for all patients undergoing hysterectomy. Furthermore, patients are assumed to not experience pregnancy while being on Relugolix CT treatment. Would that also imply a disutility related to not becoming pregnant while on Relugolix CT treatment?

Company response: we agree with the EAG that this parameter is uncertain, but it is difficult to determine what proportion of patients with a hysterectomy would be wishing to have children, particularly as this would change as the cohort ages. The intention was to capture QoL losses additional to those of infertility, which may include feelings of a loss of femininity associated with the loss of the uterus.

With respect to disutility while receiving Relugolix CT and/or GnRH agonist, as explained in B7 the patients recruited to the study would have been aware that both treatments were contraceptive and would preclude pregnancy, therefore this would be expected to have been captured within the trial EQ-5D values.

Resource use and costs

B 23. Priority question: The model assumes that 100% of patients in the BSC arm use Dienogest as part of their hormonal treatment. Although this has been stated to be the case for the concomitant medication in the post-operative medical therapy (page 169), it is currently unclear why the costs in Table 69 of the other hormonal treatments including Estrogen-progestin oral contraceptive, Medroxyprogesterone acetate, and Levonorgestrel-releasing intrauterine system were incorporated in the intervention and comparators' costs section. Please explain if the other hormonal treatments presented in Table 69 were part of the BSC or not.

Company response: We would like to apologise for making this unclear. As discussed in response to clarification question B9, the only relevant comparator included in this submission is GnRH agonists and thus BSC drug costs should not have been included in table 69 as this is not included as a comparator. The acquisition costs of BSC should have only been included as part of the concomitant medication costs (table 71). As discussed in response to question B10, the submitted model erroneously allocated all BSC patients to subsequent treatment with Dienogest (a legacy assumption from the original global model), whereas BSC is only comprised of symptomatic treatment for pain management, i.e. analgesics (NSAIDs). Thus, only the costs of NSAIDs should be included in table 71 as Dienogest is not a concomitant medication in clinical practice. The costs of other hormonal treatments listed above are not actively applied in the model (comparison of Relugolix CT vs. GnRH agonist only) and should be disregarded.

B 24. Please clarify why drug wastage is not included in the cost calculations (as part of the BSC there are injections which might be affected by wastage) and whether this is expected to have some impact on the model results. If possible, please include this option in the model.

Company response: This option is available via the Input sheet. Given that BSC is assumed to only include NSAIDs, which are an oral tablet, in the base case analysis, drug wastage was not included.

B 25. Please provide information on the dosing schedule of patients using the add-back therapy in the GnRH agonist arm as it is missing from the add-back therapy section.

Company response: Apologies for omitting the complete details regarding the dosing schedule of add-back therapy for the GnRH agonist arm. The dosing schedule for tibolone is 2.5mg daily as an oral tablet, whilst for raloxifene it is 60mg as a daily oral tablet. The model assumes 100% dose intensity for add-back therapy, that is all GnRH agonist patients comply with daily add-back therapy. As described in section B.3.5, Add-back therapy, we assume an equal split of GnRH agonist patients between tibolone and raloxifene. With GnRH agonists, patients are initiated on add-back therapy typically at three months, thus these costs are only incurred after the first cycle for the GnRH agonist arm.

Validation

B 26. Priority question: Please provide details about what validation efforts were performed in Section B.3.14 of the company submission and the results of these validation efforts. This could be presented for example (but not necessarily) with the help of the validation tool AdViSHE (https://advishe.wordpress.com/author/advishe/). Please confirm whether black-box tests to detect modelling errors were conducted. If not, please include these steps as well.

Company response: Both external and internal validations of the model were conducted.

The external validation included the following considerations:

- Model structure, clinical/treatment pathway, and key assumptions were validated with clinical expert opinions based on a global advisory board
    - Alternative treatment response definition based on clinicians' inputs was for instance implemented with cost-effectiveness results within the same magnitude than with the response definition derived from SPIRIT's co-primary endpoints
    - Clinical experts also concluded that patients may undergo a maximum of 2 conservative surgeries during their lifetime, which was reflected in

the CE model with an average number of conservative surgeries of 0.5 for Relugolix CT, 0.7 for BSC 0.7, and 0.8 for GnRH agonists

- Clinical and quality of life inputs to the model were based on pivotal trial data where possible
- To translate clinical results to ultimate health outcomes, i.e. quality of life, transparent and standard statistical approaches were applied.
- The modelled population corresponds to the pivotal clinical trial's population.
- The model used the best available evidence from external sources to inform the input parameters and assumptions.

The internal validation followed a formal technical quality control protocol that was conducted at a late stage during the development of the model. Examples of the quality control and model review performed by a second modeller can be found in the next two pages. This protocol includes black-box tests and validation of the expected results such as:

- Turn off mortality ➔ Life-years equal in both comparators and equal to model horizon (undiscounted)
- Equal efficacy and AEs for all comparators ➔ QALYs equal in both comparators
- Increase treatment cost of Relugolix CT ➔ The total cost of Relugolix CT increases and the ICER of Relugolix CT increases
- Costs of treatments and health care resource use set to £0 ➔ Costs equal to £0 in all comparators
- Increase/decrease of model horizon ➔ Increase/decrease of life years in all comparators
- Increase/decrease of utility-values for all health states ➔ Increase/decrease of QALYs in all comparators

Standardized verifications

		Greta Bütepage		Gustaf Ortsäter	
Test	**Result**	**Pass/fail/NA**	**Comment**	**Addressed**	**Comment**
All unit costs = 0	Total cost = 0	Fail	For all comp: Costs are still incurred for health care vists; error in the Markov trace; Gynaecologist visit; column CG; value series instead of 0	Yes	Hardcoded as 0
All resource use = 0	Total costs = 0	Fail	No input values for quarterly frequencies for medical tests; when setting the frequencies to 0, no costs occur for "Medical tests and exams" but for healthcare visits, see comment above.	Yes	The link to the input-sheet was missing
All societal costs = 0	Costs in payer perspective = costs in societal perspective	Pass			
~~All indirect utilities = 0~~	~~QALYs in payer perspective = QALYs in societal perspective~~	~~N/A~~	No indirect utilities were applied, i.e., utilities/QALYs for the societal and payer perspective are identical.		
Turn off mortality	LYG equal in all treatment arms	Pass			
Equal efficacy and AEs for all treatments	LYG equal in all treatment arms	Pass			
Equal efficacy and AEs for all treatments	QALYs equal in all treatment arms	Fail	Slight differences occur despite setting efficacy and AE rates equal – haven't looked more into the differences yet. I repeated this test twice, but it could be that I did not catch all efficacy/AE inputs.	Yes	
Increase treatment costs of tx A	Total cost increases	Fail	Only true for Ryeqo; and costs for the other comparators (2 – 4) increase when increasing the drug cost for Ryeqo; Reference error in the Markov traces for comp 2 – 4 = referencing drug cost of comp 1	Yes	
Decrease treatment costs of tx A	Total cost decreases	Fail	See comment above.	Yes	
Increase time horizon	Increase in LYG	Pass	Maybe add some information on the max time horizon that can be selected.		
Increase time horizon	Increase in QALYs gained per tx arm	Pass			
Decrease time horizon	Decrease in LYG	Pass			
Decrease time horizon	Decrease in QALYs gained per tx arm	Pass			
Increase efficacy of tx A	Increased QALYs for tx A	Pass	Changed the response rate to up to 100% for all comparators; difference in QALYs was pretty small though		
Decrease efficacy of tx A	Decreased QALYs for tx A	Pass	See comment above		
Increase utility for all health states	Increased QALYs in all treatment arms	Pass			
Decrease utility for all health states	Decreased QALYs in all treatment arms	Pass			
Click all buttons	Each performs expected function	Fail	The drop down menu for time for evaluation (Input sheet D64) does not work (and should be removed?) & it is implemented twice in the model (Clinical D11; wrong values, i.e. yes and no and not 3 or 6 months) - hence the values are also wrong e.g., in the Markov traces. When selecting 3 months (after correcting the list) there is a value error in the Markov traces.	Yes	

Standardized validations

		Greta Bütepage		Gustaf Ortsäter	
Validation	**Pass/fail/NA**	**Comment**		**Addressed**	**Comment**
The model structure appears to reflect the economically and clinically significant aspects of the disease	Pass			No	
The patient population appears to be reflective of the medical indication for the new intervention	Pass			No	
The decision problem is clear and relevant for the disease area and intervention	Pass			No	
The comparator interventions inluded in the model appear to be meaningful alternatives to the innovative product	Pass			No	
Check that the technical report is comprehensive, free from errors, and reader-friendly	Pass			No	
The models results make intuitive sense given the differences between the treatment arms	Pass			No	
Model interface is intuitive and clear, free from ambiguity	Pass			No	
Check that units are displayed for all numeric inputs	Pass			No	
~~For model adaptations, verify that the model structure, inputs, and outputs correspond to the local HTA requirements~~	~~N/A~~			~~No~~	

B 27. Several inputs in the model have been informed from studies performed in other countries, such as the disutility of hot flush, decreased libido and depression which were derived from a Canadian and United States study respective. Please comment on the validity of these inputs for the England and Wales setting.

Company response: We would like to draw the EAG's attention to the fact that the model is not very sensitive to the disutilities. This was confirmed by setting all disutility values in the model to 0, the results for this scenario are presented below. The ICER in this scenario (£1,717) lies extremely close to the updated base-case ICER (£1,715). The utility decrements applied thus have very little impact upon the results. Furthermore, we do not believe that the fact that the country populations mentioned above are likely to impact the validity of the disutility inputs because they are utility values elicited from relevant populations of interest. For depression and decreased libido, the disutility values were sourced from patient populations with endometriosis, whilst for hot flush, it was taken from a very similar population, with uterine fibroids.

Table 6: Cost-effectiveness results, scenario where disutilities are removed

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus baseline (£/QALY)
Relugolix CT	£11,486	23.11	17.17	-	-	-	-
GnRH agonist	£10,279	23.10	16.46	£1,207	0.01	0.703	£1,717

Electronic model

B 28. Priority question: Please explain how the response rates and use of analgesics in "Input sheet" columns AH-AX, rows 67-75 were derived. Please clarify whether the results from SPIRIT 1 & 2 were simply pooled and provide a rationale for it. Please provide results based on meta-analysis and compare them with those obtained with a simple pooling, in case the latter was used to derive the results.

Company response: Response rates and use of analgesics results were indeed based on a simple pooling of SPIRIT 1 & 2 during post-hoc analyses of individual

patient trial data. Individual trial results are not available for the response definition and use of analgesics used in the cost-effectiveness model, however as both SPIRIT 1 & 2 were randomized controlled trials with similar patient population size (SPIRIT 1: Relugolix CT, n=212 vs. Placebo, n=212 & SPIRIT 2: Relugolix CT, n=206 vs. Placebo, n=204) and similar efficacy results on the co-primary endpoints, one can expect to have similar results obtained with a simple pooling or with a pairwise meta-analysis pooling via an inverse variance weighted average (fixed effect model would be assumed in case of no between-study heterogeneity).

This rationale can be confirmed when comparing the extremely similar results obtained with the two pooling methods on the co-primary endpoints from SPIRIT trials. Calculations are included in the Excel spreadsheet submitted alongside this response document.

B 29. Please confirm that the life years that are calculated in column AX of worksheets "Result calculation compXX" are incorrect, and should be calculated using the formula =sum(AI_;AU_) instead of =sum(AI_;AT_)

Company response: We would like to thank the EAG for identifying this error. This error has been rectified in the updated post-clarification model.

Section C: Textual clarification and additional points

C 1. Table 39 of the CS presents relevant cost-effectiveness studies identified through the SLR. Please clarify if the difference between the Bohn et al. 2020 and Bohn et al. 2021, is only the perspective of the analysis (i.e., societal vs payer). The column indicating the incremental cost-effectiveness ratios (ICERs) for the different strategies in the study of Bohn et al. 2021 seems to be incomplete as only the ICER for strategy 3 is presented. Please complete any potential missing information.

Company response: We can confirm that yes, the only difference between the two Bohn et al. papers is the perspective adopted for the analysis. We would like to apologise for the oversight regarding the omission of the complete ICERs from Bohn et al., 2021. The complete information is provided in the table below.

Table 7: Summary list of published cost-effectiveness studies

Study	Year	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Grand	2019	A cost-utility analysis that compares oral contraceptives vs no hormonal therapy. It uses a Markov sate transition model structure with five health states. Results are discounted at 3.5%. The model used a time horizon of 1 month with a cycle length of 1 month. The analysis was conducted from the perspective of the NHS England.	Hypothetical cohort: 1000, Starting age of cohort: 32	No hormonal treatment: 9.88 Oral contraceptives: 10.31 Mean difference: 0.43	£1707 for no hormonal treatment £1113 for oral contraceptives £594 is the mean difference:	NR
Bohn	2020	A cost-utility analysis that compares Strategy 4: Proceeding directly to surgery without attempting medical management first.vs Strategy 1: NSAIDs followed by surgery if there was no improvement, Strategy 2: NSAIDs, then a short-acting reversible contraceptive or a long-acting reversible contraceptive (LARC) followed by surgery if no improvement, and Strategy 3: NSAIDs, then a shortacting reversible contraceptive or LARC, then a LARC or a GnRH agonist or antagonist, followed by surgery if no improvement. It uses Decision Tree model structure. The model used a time horizon of 3 years. The analysis was conducted from the societal perspective.	Hypothetical cohort: 10,018,400, 18-45 years	Strategy 4: 1.96 Strategy 1: 2.18 Strategy 2: 2.28 Strategy 3: 2.34	* Strategy 4: 3,980 USD Strategy 1: 2,328 USD Strategy 2: 1,831 USD Strategy 3: 2,842 USD	Strategy 4: 2027.34 Strategy 1: 1067.91 Strategy 2: 803.27 Strategy 3: 1216.66

Bohn	2021	A cost-utility analysis that compares Strategy 4: Proceeding directly to surgery without attempting medical management first.vs Strategy 1: NSAIDs followed by surgery if there was no improvement, Strategy 2: NSAIDs, then a short-acting reversible contraceptive or a long-acting reversible contraceptive (LARC) followed by surgery if no improvement, and Strategy 3: NSAIDs, then a shortacting reversible contraceptive or LARC, then a LARC or a GnRH agonist or antagonist, followed by surgery if no improvement. It uses Decision Tree model structure. Results are discounted at 3%. The model used a time horizon of 3 years. The analysis was conducted from the payor perspective.	Hypothetical cohort: 4,817,894 18-45 years	Strategy 4: 9.7 million Strategy 1: 10.7 million Strategy 2: 11.2 million Strategy 3: 11.4 million	** Strategy 4: 42.1 billion Strategy 1: 22.6 billion Strategy 2: 12.9 billion Strategy 3: 13.2 billion	Strategy 2: -£20,204 Strategy 3: $1,352/QALY Strategy 4: -£17,269

References

1. Veth VB, van de Kar MMA, Duffy JMN, van Wely M, Mijatovic V, Maas JWM. Gonadotropin-releasing hormone analogues for endometriosis. Cochrane Database of Systematic Reviews. 2023(6).

2. Astellas. Scientific results summary: A study to assess the effectiveness and safety of different doses of ASP1707 compared to placebo for endometriosis associated pelvic pain.; 2016.

3. Giudice LC, As-Sanie S, Arjona Ferreira JC, Becker CM, Abrao MS, Lessey BA, et al. Once daily oral relugolix combination therapy versus placebo in patients with endometriosis-associated pain: two replicate phase 3, randomised, double-blind, studies (SPIRIT 1 and 2). Lancet (London, England). 2022;399(10343).

4. Bougie O, Yap MI, Sikora L, Flaxman T, Singh S. Influence of race/ethnicity on prevalence and presentation of endometriosis: a systematic review and meta-analysis. BJOG : an international journal of obstetrics and gynaecology. 2019;126(9).

5. Myovant Sciences GmbH. SPIRIT 1 (MVT-601-3101) Clinical Study Report. Data on file. 2021.

6. Myovant Sciences GmbH. SPIRIT 2 (MVT-601-3102) Clinical Study Report. Data on file. 2021.

7. Australian Institute of Health and Welfare. Endometriosis, Hospitalisations. 2023.

8. Australian Institute of Health and Welfare. Endometriosis, Emergency Department presentations. 2023.

9. Grundström H, Gerdle B, Alehagen S, Berterö C, Arendt-Nielsen L, Kjølhede P. Reduced pain thresholds and signs of sensitization in women with persistent pelvic pain and suspected endometriosis. Acta Obstetricia et Gynecologica Scandinavica. 2019;98(3):327-36.

10. In: Szende A, Janssen B, Cabases J, editors. Self-Reported Population Health: An International Perspective based on EQ-5D. Dordrecht (NL): Springer
Copyright 2014, The Editor(s) (if applicable) and the Author(s). 2014.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

[ID3982]

Additional clarification questions – company response

November 2023

File name	Version	Contains confidential information	Date
ID3982 Relugolix EAG clarification letter_additional company responses 16Nov23 [No CON]	Final	No	16/11/23

Updated response to clarification question A1

As indicated in our original response dated 9th November 2023, we have undertaken two supplementary search activities. These are additional to the original SLR as well as a pragmatic literature review which we had not explicitly reported in our submission (see the NMA feasibility assessment, included as data on file with this response). This pragmatic literature review was carried out by Gedeon Richter to supplement the results of the original SLR, which was deemed to report a low number of studies. Note that this original SLR had not been commissioned by Gedeon Richter but by Myovant, the developer of Relugolix CT.

The two supplementary search activities informing this response comprise:

- Comparing the included studies from the Cochrane review with those identified in the SLR and pragmatic literature review. The studies were analysed to see whether they captured Overall Pelvic Pain (OPP) and/or Total Pelvic Pain (TPP) and what pain scales were used.
- Updating the Cochrane search strategy to include May 2022 to present day. Note that the Cochrane search strategy included GnRH antagonist terms and thus would be expected to include any Relugolix CT studies.

The results of these two further searches are as follows:

<u>Studies identified from the Cochrane review</u>

Seven studies were identified from the Cochrane review that were not included in the original clinical SLR/pragmatic literature review and for which reasons for exclusion were not provided.

Records were not available for three of those studies. We are awaiting delivery of the following full texts:

1. Agarwal et al. Nafarelin vs. leuprolide acetate depot for endometriosis. Changes in bone mineral density and vasomotor symptoms. Nafarelin Study Group. Journal of Reproductive Medicine 1997;42(7):413-23.
2. Jelley & Magill. The effect of LHRH agonist therapy in the treatment of endometriosis (English experience). Progress in Clinical & Biological Research 1986;225:227-38.

3. Minaguch et al. Clinical study on finding optimal dose of a potent LHRH agonist (buserelin) for the treatment of endometriosis--multicenter trial in Japan. Progress in Clinical & Biological Research 1986;225:211-25.

The remaining four studies reported outcomes for either OPP or TPP:

1. Odukoya et al. Serum-soluble CD23 in patients with endometriosis and the effect of treatment with danazol and leuprolide acetate depot injection. Human Reproduction 1995;10(4):942-946.
2. Dmowski et al. Ovarian suppression induced with buserelin or danazol in the management of endometriosis: a randomized, comparative study. Fertility & Sterility 1989;51(3):395-400.
3. Crosignani et al. Leuprolide in a 3-monthly versus a monthly depot formulation for the treatment of symptomatic endometriosis: a pilot study. Human Reproduction 1996;11(12):2732-5.
4. Tummon et al. A randomized, prospective comparison of endocrine changes induced with intranasal leuprolide or danazol for treatment of endometriosis. Fertility & Sterility 1989;51(3):390-4.

Consideration of potential inclusion of the above studies in the ITC is ongoing.

<u>Updated searches</u>

Three databases were searched (from May 2022 to present) to identify any new studies:

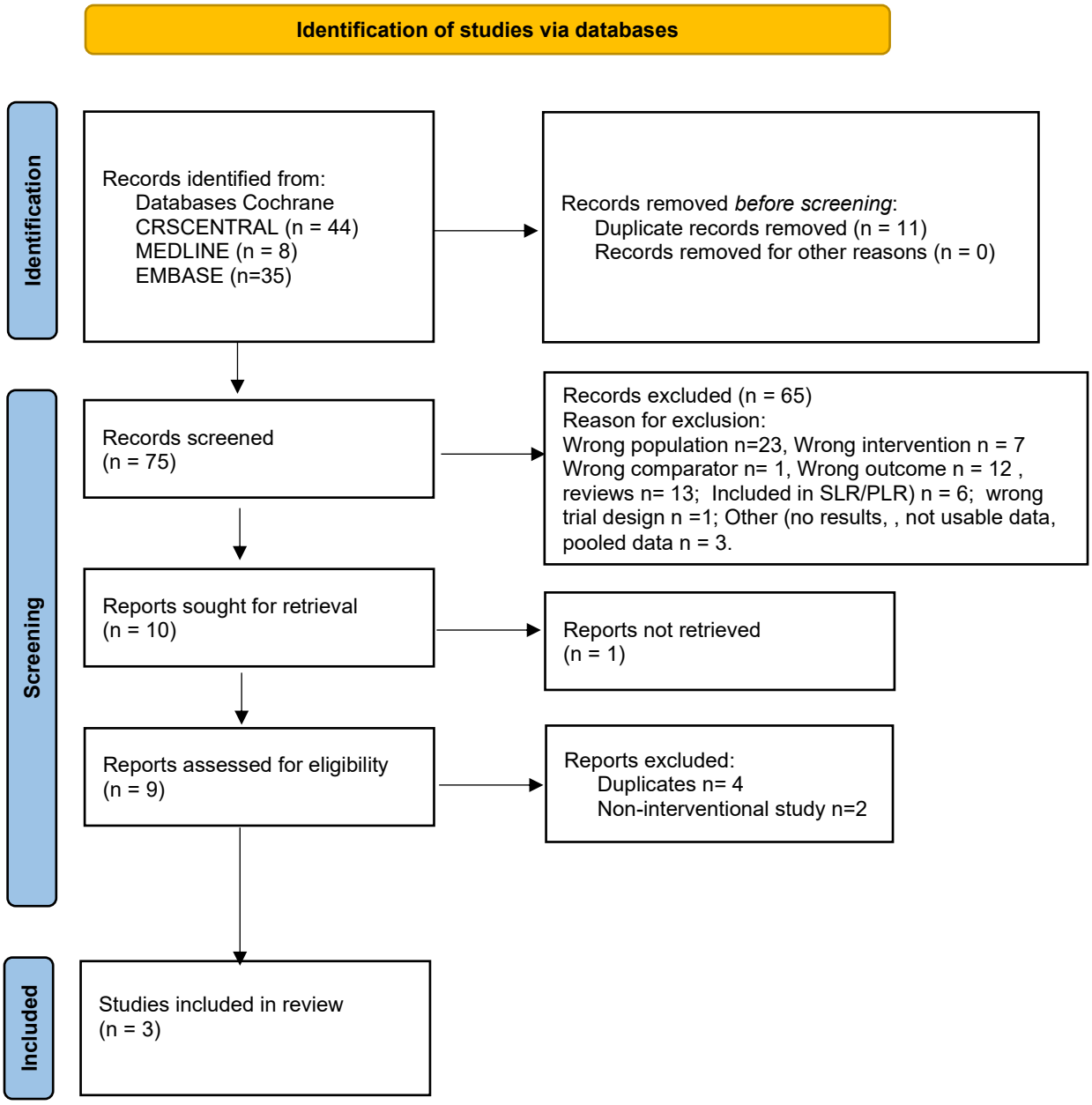
- CENTRAL via the Cochrane Register of Studies Online (CRSO); searched from May 2022--10th November 2023
- MEDLINE (Ovid platform); searched from 2022-13th November 2023
- Embase (Ovid platform); searched from 2022-15th November 2023

The search strategy used was identical to the Cochrane review.
https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD014788.pub2/pdf/full

The PRISMA diagram below shows the process for screening and selection of relevant studies. Nine records were assessed for eligibility and six were excluded. The remaining three studies included outcomes for either OPP or TPP:

1. As-Sanie et al. Relugolix combination therapy in North American women with endometriosis-associated pain: SPIRIT 1 and 2 trials. Fertility and Sterility 2022; 118(4 suppl):E223.
2. Li et al. Assessment of two formulations of triptorelin in Chinese patients with endometriosis: a phase 3, randomized controlled trial. Adv Ther 2022;39(10):4663-77.
3. Tang et al. Comparison of the efficacy of dienogest and GnRH-a after endometriosis surgery. BMC Women's Health 2023;23(1):85.

Figure 1: PRISMA diagram for updated searches



A spreadsheet showing a list of included studies and pdfs of the additional articles identified are provided with this response.

Responses to additional clarification questions on parameter uncertainty

B 30 Priority question: Please fix the PSA in the model. The company's PSA results are nearly identical to the deterministic ones, and we think this is caused by assuming a fixed 10% variation from the mean for all parameters in the PSA. We think this is not correct. The uncertainty around each input parameter should be implemented according to the parameter's source, and this uncertainty is unlikely to be the same for all parameters. Please would the company implement this in the model, if possible. At a minimum, the fixed standard error for all parameters should be removed from the model and let all parameters to have their own SE. Special attention is required to those parameters for which a non-symmetric confidence interval is to be expected (for example, a hazard ratio).

Company's response: The SE has been replaced by the uncertainty of the parameter's source, either directly or calculated from reported confidence intervals where available. For parameters which source did not report a variation or measure of uncertainty, the model assumes a SE equals to 10% (this value can be changed by the user in *Variable bank* – cell D6) of the input's mean estimate which was also the case in the previously submitted model. This allows to account then for the respective distribution associated to each input and for instance to have asymmetric confidence interval for relevant parameters like odds ratios.

Graphics are available in the model (in Excel sheet *Results - Sensitivity*) to show the convergence of the probabilistic results, average of the ICER over the n^{th} first PSA iterations, to the deterministic ones (horizontal line). One can notice that the average of ICERs of REL-CT vs GnRH agonists over 1,000 iterations is however below the deterministic ICER due to a number of iterations where REL-CT has lower mean total costs than GnRH agonists and thus inducing negative ICER.

B 31 Priority question: Many parameters are excluded from the PSA for no obvious reason to us. Please clarify why this has been done, or if this has been overlooked, please include them in the PSA.

Company's response: This has now been corrected.

B 32 Priority question: Some parameters associated to regression equations were included in the PSA but are sampled independently instead of using the variance/covariance matrices. If possible, please also fix this.

Company's response: The covariance matrix from the Framingham Risk Function for Cardiovascular Events was not available and could unfortunately not be implemented in the PSA. However, SEs from the estimated beta-coefficients have now been included.

B 33 Priority question: When the fixed SE is increased to 20%, the model returns an error. This is because the starting age can occasionally drop below 18 years and, in that case, the lookup function for mortality (on Mortality sheet) produces an error. Please fix this too, and make sure that this would not happen for other input parameters.

Company's response: We have adjusted the model so that the age never goes below 18 years. Age has also been included in the PSA, as well as the model time horizon and age at menopause. The model horizon has also been adjusted so that it never runs beyond 100 years (e.g., it can never be longer than [100 – age]. Furthermore, the model horizon has also been conditioned to be at least 1 year long. Finally, a restriction on the sampled value has been implemented to ensure that age at menopause is at least one year more than age at baseline. This is to ensure that the model runs for at least one year.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Relugolix–estradiol–norethisterone acetate for treating symptoms of endometriosis [ID3982]

Patient Organisation Submission

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type. [Please note that declarations of interests relevant to this topic are compulsory].

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you

1.Your name	[redacted]
2. Name of organisation	Endometriosis UK
3. Job title or position	[redacted]
4a. Brief description of the organisation (including who funds it). How many members does it have?	We are the leading UK charity for all those affected by endometriosis, determined to see that everyone gets prompt diagnosis and the best treatment and support. We provide support and information for those living with the disease, their partners, and families; raise awareness; support research and campaign for change. We are the only UK organisation providing endometriosis support through a Helpline, web chat, online moderated community (Health Unlocked) and 54 local support groups. Over 200 volunteers with personal lived experience of the disease are trained to provide support through our services, directly reaching over 70,000 impacted by endometriosis last year. We lead collaboratively across our community, acting to inform, empower and advocate for all those affected by the disease. The endometriosis community is at the heart of all we do, and we strive to support and represent all experiences. Endometriosis UK delivers a range of projects including supporting employers around Women's Menstrual Wellbeing, and an Endometriosis Friendly Employer Scheme. Funding is received from a range of sources, the majority through voluntary income donations and community fundraising. Grants and donations are also received from a range of funders including statutory, corporates and charitable trusts and foundations.
4b. Has the organisation received any funding from the company bringing the treatment to NICE for evaluation or any of the comparator treatment companies in the last 12 months? [Relevant	No

companies are listed in the appraisal stakeholder list.] **If so, please state the name of the company, amount, and purpose of funding.**	
4c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No
5. How did you gather information about the experiences of patients and carers to include in your submission?	We regularly gather information on experience of patients about the diagnosis and treatment of endometrioses. In the last 3 years this has included 4 surveys, receiving between 2,000 and 10,000 responses. In addition to surveys, we regularly consult the endometriosis community, analyse information gathered through our support services such as Helpline enquiries, and undertake qualitative and focus group research. In addition to refine our submission, we held semi-structured interviews with 10 patients with endometriosis to whom use of this drug would be relevant, gathered their experiences and feedback.

Living with the condition

6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	Endometriosis affects 1 in 10 women and those assigned female at birth from puberty to menopause, although the impact may be felt for life. Symptoms and impact vary for each individual depending on the location and extend of their disease, with the most common symptom being chronic pain. Many patients describe endometriosis as debilitating, with daily pain and an overall lower quality of life. Those we interviewed for the submission said they often feel like their life is on hold, and struggle with daily tasks especially when they have a high pain day. In the Endometriosis APPG Report[1] which was published in 2020,10,000 patients were asked about their quality of life living with endometriosis and 95% of respondents said endometriosis symptoms have had a negative or very negative impact on their wellbeing. Respondents that we interviewed for this submission also described living with endometriosis as extremely challenging and the detrimental impact that the condition's symptoms can have upon their day-to-day lives both logistically and in terms of quality of life and general physical and mental wellbeing requires exhaustive resilience.

[1] Endometriosis APPG Report Oct 2020.pdf (endometriosis-uk.org)

Current treatment of the condition in the NHS

7. What do patients or carers think of current treatments and care available on the NHS?	Those we interviewed said they found the overall treatment on the NHS currently is inadequate, reinforcing the messages that we have identified through surveys over the last 3 years expressed that the found the process of getting a diagnosis and also receiving treatments and aftercare with follow up appointments a very long-winded struggle. Despite the lengthy processes currently in place, patients said that they often don't find they get a lot of out of their appointments unless they really advocate for themselves and "fight" in appointments in order not to be dismissed by medical professionals. Patients expressed that the options available as treatments at the moment can all have considerable side effects and expressed how they were often being prescribed a type of contraceptive for their pain without any real investigation as to what the pain might be or were not given enough information about the potential side effects of hormonal contraceptives. All medical treatment options are also currently all hormonal which are not suitable for those wishing to conceive.
8. Is there an unmet need for patients with this condition?	In the Endometriosis APPG Report 2020, survey respondents were asked how helpful they found their GP in supporting them with endometriosis/symptoms and 46% of respondents said that they GP had been unhelpful or very unhelpful. The majority of responders, 63% stated that if they wanted to speak to a medical professional about their endometriosis symptoms they would go to their GP first, identifying how difficult many people find the system in accessing useful support and help. Those who we interviewed on this occasion, similarly stated that there is an unmet need for patients with endometriosis. Interviewees described a lack of basic understanding about endometriosis as a condition, and a need for better treatment pathways or referrals once diagnosed. Patients also described their condition not being seen as a priority by healthcare practitioners with appointments being repeatedly cancelled and long waiting times for those appointments and surgeries. This was noted in RCOG's report, 'Left too Long'[2], where it was reported that Gynaecology waiting lists across the UK have now reached a combined figure of over 570,000 women across the UK – just over a 60% increase on pre-pandemic levels', and that 'Gynaecology waiting lists in England have grown the most in percentage terms of all elective specialties'. There is currently a need to treat and manage endometriosis like the long-term condition it is; at present if patients have recurrence of symptoms some time post-surgery, they feel that they often have to practically go through the diagnosis process again to access the care they need. There needs to be better training for healthcare practitioners, so they recognise the symptoms and understand the care pathways.

Advantages of the technology

9. What do patients or carers think are the advantages of the technology?	Those interviewed viewed having an all-in-one treatment that would be Relugolix–estradiol–norethisterone acetate as a positive step, as it has the treatment and HRT combined together, with one respondent noting that not having to remember to take HRT in tablet form could be a huge step forward. Respondents also said it was positive that it can be used for a much longer period of time compared to other, similar treatments currently available, which can only be used for a few months. Those asked also felt as the drug can be used as a contraceptive up until natural menopause that this would be extremely useful and one less thing to worry about. Those interviewed also said having a tablet which can be taken daily was very positive, adding if there were concerns about side effects, then the patient would be able to stop taking the drug, rather than as at present, being on similar drugs which may require a 3 month injection. Patients also praised about new treatments coming out for endometriosis, believing it to be a positive step for the future of endometriosis care.

[2] Gynae waiting lists (shorthandstories.com)

Disadvantages of the technology

10. What do patients or carers think are the disadvantages of the technology?	Overall, the responses we received from those we interviewed were positive. Some concerns were raised about the length of time the drug would be in a person's system, with reference to one drug not fitting all. However, the significant benefit of a daily dose over a 1 – 3 month injection was clearly noted. Some interviewees spoke about their experiences on HRT in particular and how they often have had to change the type of HRT they were prescribed as it did not agree with them. There were queries with reference to this drug, as to whether the HRT aspect would not suit some and it would mean that they would have to stop taking the drug. Interviewees wanted to understand more about the drug including potential side effects, and how these would be properly assessed and looked into. Effective patient information will be key. Some also spoke of their concern about the perceived lack of information about the long-term side effects of this drug, especially for those who may have already have had several years of taking prescribed GnRH Analogues/Agonists. Concerns include around including potential conditions such as cardiovascular disease and osteoporosis. Patients also spoke of their want of treatments or drugs that are not hormonal and with less potential side effects.

Patient population

11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	Some respondents said that they were not able to take GnRH Analogue treatments due to the side effects of that type of drug, therefore this would not be suitable for some people's treatment or management of endometriosis. People who want to conceive will also not be able to take this drug as a form of treatment. Bone density was also raised as a concern for people who will be undergoing this type of treatment. Those interviewed felt it would be important to have bone density scans to ensure that bone deterioration, if it does happen, will be swiftly picked up, treated and recognised.

Equality

12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	At the moment, what treatments and drugs are offered are based on the knowledge of individual medical professionals as well as what is currently available in certain areas and different NHS Trusts. This means the choices of treatments can be a postcode lottery.

Other issues

13. Are there any other issues that you would like the committee to consider?	Whilst we are keen to see this drug made available for those for whom it is appropriate, we would like to add that the treatment and management options for those with endometriosis are limited due to historic lack of research. We would urge NICE to look to identify further possible ways to identify and provide new technologies to endometriosis patients. Some patients are concerned about hormonal drugs and their long-term effects. There should be clear patient information provided due to potential concerns or misunderstanding around the impact of taking any HRT with endometriosis.

Key messages

14. In up to 5 bullet points, please summarise the key messages of your submission.	• There was positive feedback in regard to having an all-in-one treatment where the patient does not have to remember to additionally take HRT as a separate tablet. • Respondents were also positive at the prospect of being able to have this treatment for a longer period of time than current available treatments. • A benefit of this drug over other treatments is that it is taken daily, so if side effects are deemed unmanageable it can be stopped quickly (comparison with 3 month injection). • HRT is included, which should help mitigate the negative effects of menopause such as bone density. •

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

Please select YES if you would like to receive information about other NICE topics - YES or NO

For more information about how we process your personal data please see our privacy notice.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Relugolix–estradiol–norethisterone acetate for treating symptoms of endometriosis [ID3982]

Clinical expert statement

Information on completing this form

In part 1 we are asking for your views on this technology. The text boxes will expand as you type.

In part 2 we are asking you to provide 5 summary sentences on the main points contained in this document.

Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted as 'confidential [CON]' in turquoise, and all information submitted as 'depersonalised data [DPD]' in pink. If confidential information is submitted, please also

send a second version of your comments with that information redacted. See Health technology evaluations: interim methods and process guide for the proportionate approach to technology appraisals (section 3.2) for more information.

The deadline for your response is **5pm** on **<insert deadline>.** Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Part 1: Treating endometriosis and current treatment options

Table 1 About you, aim of treatment, place and use of technology, sources of evidence and equality

1. Your name	Elizabeth Bruen
2. Name of organisation	University Hospital of Wales Cardiff
3. Job title or position	Endometriosis CNS; Surgical assistant; Nurse Hysteroscopist; Nurse Sonographer
4. Are you (please tick all that apply)	☐ An employee or representative of a healthcare professional organisation that represents clinicians? ☒ A specialist in the treatment of people with endometriosis? ☐ A specialist in the clinical evidence base for endometriosis or technology? ☐ Other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☒ Yes, I agree with it ☐ No, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ Other (they did not submit one, I do not know if they submitted one etc.)
6. If you wrote the organisation submission and/or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission)	☐ Yes
7. Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	none
8. What is the main aim of treatment for endometriosis?	To improve symptom management and maximise fertility

(For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability)	
9. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount)	Reduction in pain score; increased quality of life sand patient satisfaction; delay to further surgical intervention; less appointment in clinic
10. In your view, is there an unmet need for patients and healthcare professionals in endometriosis?	Absolutely delayed doaiagnosos;1 in 10 women minimal access to service a and this is supported by NICE 2017; Task and Finish Group 2018; APGG 2020; Scottish report 2021; Womans Health Plan Wales 2022 a;l identify the need to improve.

11. How is endometriosis currently treated in the NHS?
• What is first-, second-, third-line treatment etc?
• Are any clinical guidelines used in the treatment of the condition, and if so, which?
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)
• What impact would the technology have on the current pathway of care?
  ○ where in the treatment pathway would you expect this treatment to be used? – i.e first-, second-, third-line, etc
  ○ what treatments might it displace? (i.e. what are the comparator treatments)
• Regarding best supportive care:
  ○ How would you define best supportive care?
  ○ Does it include analgesics? If so, which ones?
	First line basic analgesics then hormonal if willing and appropriate. If no improvement refer to core gynaecology.
	Full history to cover menstrual health, sexual wellbeing, bowel, bladder, thoracic function; fertility wishes; psychological well being; assessment of pelvic pain and chronic pain; assessment of previous management and responses;
	Investigations can be TVS
	Bimanual examination; assessment of vaginal vault if no i
	mprovement diagnostic laparoscopy.
	If endometriosis present consider excision rather than ablation when able. Diagnostic laparoscopy shuold include a 360 degree assessment of pelvis and elevation of Pouch of Douglas to assess for rectovagianal endometriosis
	NICE clinical guidelines 2017
	Access to service is problematic. There continue to be significant delays in diagnosis
	This treatment would be considered if symptoms unmanageable or wish to avoid surgical intervention
	In Wales each Health Board has an Endometriosis CNS but this needs to be extended to more nurses to bridge between primary and secondary care. Access to Consultants; Pelvic Physiotherapists; Nutritionist; Psychosexual and counselling are vital but sporadically accessible or available

○ Is it appropriate to use best supportive care after not responding to treatments (i.e. relugolix CT, GNRH agonists, or surgery subsequent to either relugolix CT or GNRH agonists). And to use at a further timepoint in the pathway if BSC had previously 'failed'?	For more severe endometriosis MDT approach is essential including Colorectal; Pain specialists; Urologist; Cardo thoracic surgeon; Radiologist; Analgesics range from paracetomol; Codeine; NSAIDS; TENS; lidocaine patches; opiates slow release and breakthrough; europathic medication. What is essential is pain assessment and body maping with contructive assessmeny ==t and evaluation of impact. Hormonal manipulation can also improve symptoms and reduce pain. Best supportive care is fundamentally not funded and patients tend to access through self funding.
12. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice? • How does healthcare resource use differ between the technology and current care? • In what clinical setting should the technology be used? (for example, primary or secondary care, specialist clinic) • Would the technology be considered for all women or only those women who do not seek to have children? • What investment is needed to introduce the technology? (for example, for facilities, equipment, or training)	To be used as alternative to prostap Secondary tertiary commencement then continue in primary care Unlikely to conceive but advised to use contraceptive but can be given as treatment prior to time wish to conceive Training for staff and information for patients
13. Do you expect the technology to provide clinically meaningful benefits compared with current care? • Do you expect the technology to increase length of life more than current care? • Do you expect the technology to increase health-related quality of life more than current care?	Yes as will not need to attend for injections every month or require HRT management Direct saving on service provision

14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	May be impacted by comorbidity such as liver failure
15. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use? (For example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed)	Easier less attendance to clinic or |GP for injection. Woman avoids injection and severe menopause symptoms and reduced risk of osteoporosis.
16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Local intention woud be to evaluate at 4 months for impact
17. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation? • Do the instruments that measure quality of life fully capture all the benefits of the technology or have some been missed? For example, the treatment regimen may be more easily administered (such as an oral tablet or home treatment) than current standard of care	Yes better symptom management and improved quality of life
18. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	Yes We hope will be significant development in management of symptoms and impact on quality of life

• Is the technology a 'step-change' in the management of the condition? • Does the use of the technology address any particular unmet need of the patient population?	
19. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	
20. Do the clinical trials on the technology reflect current UK clinical practice? • If not, how could the results be extrapolated to the UK setting? • What, in your view, are the most important outcomes, and were they measured in the trials? • If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes? • Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	I would expect so?
21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No I am not
22. How do data on real-world experience compare with the trial data?	Don't have sufficient information
23. NICE considers whether there are any equalities issues at each stage of an evaluation. Are there any potential equality issues that should be taken into account when considering this condition and this treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged.	Not that I am aware of this is exciting development

Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics. Please state if you think this evaluation could • exclude any people for which this treatment is or will be licensed but who are protected by the equality legislation • lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population • lead to recommendations that have an adverse impact on disabled people. Please consider whether these issues are different from issues with current care and why. More information on how NICE deals with equalities issues can be found in the NICE equality scheme. Find more general information about the Equality Act and equalities issues here.	

Part 2: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

Improved access to care

Improved management of symptoms

Increase options for management available to clinician for individualisation of care

Highlight the ongoing issues with access to service and delayed diagnosis

Provide options with fewer potential long term side effects

Thank you for your time.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☒ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Relugolix–estradiol–norethisterone acetate for treating symptoms of endometriosis [ID3982]

Clinical expert statement

Information on completing this form

In part 1 we are asking for your views on this technology. The text boxes will expand as you type.

In part 2 we are asking you to provide 5 summary sentences on the main points contained in this document.

Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted as 'confidential [CON]' in turquoise, and all information submitted as 'depersonalised data [DPD]' in pink. If confidential information is submitted, please also

send a second version of your comments with that information redacted. See Health technology evaluations: interim methods and process guide for the proportionate approach to technology appraisals (section 3.2) for more information.

The deadline for your response is **5pm** on **<insert deadline>.** Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Part 1: Treating endometriosis and current treatment options

Table 1 About you, aim of treatment, place and use of technology, sources of evidence and equality

1. Your name	Karolina Afors
2. Name of organisation	University College London Hospital
3. Job title or position	Consultant Gynaecologist
4. Are you (please tick all that apply)	☒ An employee or representative of a healthcare professional organisation that represents clinicians? ☒ A specialist in the treatment of people with endometriosis? ☒ A specialist in the clinical evidence base for endometriosis or technology? ☐ Other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☒ Yes, I agree with it ☐ No, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ Other (they did not submit one, I do not know if they submitted one etc.)
6. If you wrote the organisation submission and/or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission)	☐ Yes
7. Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	Nothing to disclose
8. What is the main aim of treatment for endometriosis? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability)	There is no cure for endometriosis. The main aim of treatment for endometriosis is to reduce the severity of symptoms and to improve the quality of life for individuals affected by the condition.

9. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount)	A significant reduction in symptoms and improvement in quality of life can be considered a significant treatment response. This may include a reduction in pain, improvement in overall quality of live, including physical and emotional well-being as well as ability to engage in daily activities. A clinically significant response may involve a decrease in size or number of endometriotic lesions or for patients trying to conceive a successful pregnancy outcome.
10. In your view, is there an unmet need for patients and healthcare professionals in endometriosis?	There is a huge unmet need with regards to endometriosis. The disease can negatively effect a patient's physical health, impacting their quality of life and productivity or ability to work. Despite its high prevalence, there is also a lack of disease awareness among patients, health care providers, and the public. It is often misdiagnosed and takes on average 7.5 years to be diagnosed, placing significant economic and social burden on patients, their families, and society as a whole. Despite its high prevalence, there is limited understanding of the disease process and generally endometriosis remains underfunded and under-researched. There is a need for innovation in both diagnostic and treatment options.
11. How is endometriosis currently treated in the NHS? • What is first-, second-, third-line treatment etc? • Are any clinical guidelines used in the treatment of the condition, and if so, which? • Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.) • What impact would the technology have on the current pathway of care?	For women with endometriosis related pain, benefits and risk of analgesics can be discussed. Taking into account any comorbidities and the woman's preferences. Consider a short trial (for example, 3 months) of paracetamol or NSAID, alone or in combination with hormonal treatment. It is recommended to offer women hormonal treatment to reduce endometriosis associated pain (such as combined oral contraceptive, progestogen, levonorgestrel releasing intrauterine system or an etonogestrel-releasing subdermal implant. GnRH agonists are prescribed as second-line (e.g. if hormonal contraceptives or progestogens have been ineffective) due to their side effect profile. Clinicians should consider prescribing hormonal add back therapy alongside GnRh agonist therapy to reduce risk of hypo oestrogenic symptoms and osteoporosis.

○ where in the treatment pathway would you expect this treatment to be used? – i.e first-, second-, third-line, etc ○ what treatments might it displace? (i.e. what are the comparator treatments) • Regarding best supportive care: ○ How would you define best supportive care? ○ Does it include analgesics? If so, which ones? ○ Is it appropriate to use best supportive care after not responding to treatments (i.e. relugolix CT, GNRH agonists, or surgery subsequent to either relugolix CT or GNRH agonists). And to use at a further timepoint in the pathway if BSC had previously 'failed'?	Clinical guidelines used in management include NICE endometriosis guideline and ESHRE Guideline If initial hormonal treatment for endometriosis is not effective, not tolerated, or is contraindicated (ie woman wishes to conceive), referral to a specialist endometriosis service is recommended. It is recommended to offer surgery as one of the options to reduce endometriosis associated pain. I would expect GnRh antagonists to be prescribed as second line (for example if hormonal contraceptive or progestogens have been ineffective). Similarly to when GnRh agonists are currently prescribed to reduce endometriosis associated pain. GnRh antagonists can also be used as combination treatment for those patients awaiting surgery for symptom relief or where ovarian suppression is considered of benefit pre or post surgical excision of endometriosis. The most appropriate comparator is GnRh agonist with add back treatment. Other alternative medical treatment is dienogest (4th generation progestogen) which is licensed for treatment of endometriosis. Continuous or chronic pain can lead to central sensitisation which results in lowering of pain threshold. Chronic pain can lead to posture changes and muscles cramps etc Complementary measures can provide best supportive care notably. Physiotherapy including pelvic floor physio Psychological support Acupuncture/osteopathy. Nutrition/dietary changes Other treatment options also available in the NHS would be pain medication i.e neuromodulators in those experiencing recurrence of pain symptoms with neuropathic type pain component. Theses can be used in conjunction with hormonal treatment.

	Due to the chronic, non-curative nature of endometriosis and limited evidence with regards to treatment strategies I think it is important to use best supportive care alongside conventional treatment and to offer it at further time points if previously failed as may improve quality of life and living with the disease more manageable.
12. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice? • How does healthcare resource use differ between the technology and current care? • In what clinical setting should the technology be used? (for example, primary or secondary care, specialist clinic) • Would the technology be considered for all women or only those women who do not seek to have children? • What investment is needed to introduce the technology? (for example, for facilities, equipment, or training)	Relugolix–estradiol–norethisterone acetate is recommended, within its marketing authorisation, as an option for treating moderate to severe symptoms of uterine fibroids in adults of reproductive age and as symptomatic treatment of endometriosis in women with a previous history of medical or surgical treatment for their endometriosis. It is currently under review by NICE for treatment moderate to severe symptoms of fibroids the results of which are due to be published
13. Do you expect the technology to provide clinically meaningful benefits compared with current care? • Do you expect the technology to increase length of life more than current care? • Do you expect the technology to increase health-related quality of life more than current care?	Relugolix–Estradiol–Norethisterone Acetate represents a novel advancement in women's healthcare, offering a promising alternative to the current standard of care **Improved Patient Experience:** - **Reduced Treatment Burden:** Relugolix–Estradiol–Norethisterone Acetate eliminates the need for frequent clinic visits associated with injectable GnRH agonists. Patients no longer require regular injections every 3 months, thereby reducing inconvenience and discomfort. - **Enhanced Convenience:** With oral administration, patients have the flexibility to take their medication at home, eliminating the need for clinic appointments. This contributes to greater autonomy and adherence to treatment regimens.

	Reduced Healthcare Resource Utilisation: - **Reduction in clinic visits:** By eliminating the requirement for frequent injections, Relugolix–Estradiol–Norethisterone Acetate reduces the strain on healthcare facilities and resources associated with administering injectable medications. This may alleviate clinic overcrowding and improve overall healthcare efficiency. - **Cost Savings:** The streamlined patient journey facilitated by oral therapy may lead to cost savings for both healthcare systems and patients, including reduced transportation expenses and fewer missed workdays due to clinic visits.
14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Relugolix–estradiol–norethisterone acetate is less likely to be effective and in postmenopausal women. Equally it would not be considered appropriate in those patients wishing to conceive. In patients with a history of a low trauma fracture or other risk factors for osteoporosis or bone loss, including those taking medications that may affect BMD, should be individually assessed and may not be considered appropriate for initiating treatment.
15. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use? (For example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed)	The most appropriate comparator to Relugolix–estradiol–norethisterone acetate is GnRh agonist. The licensed use of GnRh agonists is limited to 6 months due to the risk of osteoporosis. To ensure no harm is done whilst in treatment, a base line bone mineral density scan is recommended before treatment started and repeated after 12-18 months. Prospective observational study on Relugolix–estradiol–norethisterone acetate demonstrated non-clinically relevant decrease in bone mineral density (BMD), which stabilized after 12-24 weeks of treatment and thereafter remained stable. However, decreases of > 3% were seen in 21% of the patients. Therefore, a DEXA scan is recommended after the first 52 weeks of treatment and as considered appropriate thereafter. The technology will be similar in terms of ease of use for patients and health care professionals alike. Relugolix–estradiol–norethisterone acetate as compared to GnRh agonists appears to demonstrate lower risk of decrease in bone mineral density loss.

16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	The benefits and risks of Relugolix–estradiol–norethisterone acetate in patients with a history of a low trauma fracture or other risk factors for osteoporosis or bone loss, including those taking medications that may affect BMD, should be considered prior to initiating treatment. Relugolix–estradiol–norethisterone acetate should not be initiated if the risk associated with BMD loss exceeds the potential benefit of the treatment. As a result a DEXA scan is recommended after 1 year of treatment. Any hormonal contraception should be stopped prior to commencement of treatment. Pregnancy should be ruled out prior to initiating treatment. Discontinuation should be considered when patient enters menopause as endometriosis related pain symptoms are known to regress during menopause. Relugolix–estradiol–norethisterone acetate is contraindicated in the presence or history of severe hepatic disease where liver function values have not returned to normal. Treatment should be discontinued if jaundice develops.
17. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation? • Do the instruments that measure quality of life fully capture all the benefits of the technology or have some been missed? For example, the treatment regimen may be more easily administered (such as an oral tablet or home treatment) than current standard of care	I consider that Relugolix–Estradiol–Norethisterone Acetate medical therapy largely provides health related benefits that are likely to be included in the QALY calculation. The only quality of life measure that may have been omitted is that this represents a second line treatment alternative in the form of an oral treatment which is more easily administered then the current comparable GnRh agonists which are administered as an injection.
18. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	I consider that Relugolix–Estradiol–Norethisterone Acetate address an important unmet clinical need. It provides an extra choice to tackle a significant gap in medical care for endometriosis by offering effective, safe, and well-tolerated treatment options that can be utilized over an extended period. This helps decrease reliance on opioids and enhances quality of life for individuals dealing with the condition.

• Is the technology a 'step-change' in the management of the condition? • Does the use of the technology address any particular unmet need of the patient population?	I believe this is a step change in the management of the condition providing an oral alternative, patients have the flexibility to take their medication at home, eliminating the need for clinic appointments. This contributes to greater autonomy and adherence to treatment regimens.
19. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	If side effects are severe or intolerable, they may lead to discontinuation of treatment, affecting the efficacy of managing endometriosis associated pelvic pain endometriosis. Hypo-oestrogenic side effects such as hot flashes could affect compliance. Additionally, side effects like nausea, headaches, or changes in mood can reduce the patient's overall well-being and quality of life, impacting their daily activities and emotional state. Furthermore, certain adverse effects, such as changes in bone density, may require additional monitoring or interventions to mitigate potential long-term health risks, adding complexity to the management of the condition.
20. Do the clinical trials on the technology reflect current UK clinical practice? • If not, how could the results be extrapolated to the UK setting? • What, in your view, are the most important outcomes, and were they measured in the trials? • If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes? • Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	The clinical trials in my opinion do reflect current clinical practice in the UK. The most important outcome is improvement in endometriosis associated pelvic pain which was measured in SPIRIT 1 and 2 TRIAL where the co-primary endpoints were responder rates for dysmenorrhoea and non-menstrual pelvic pain. No adverse effects that I am aware of that have come to light subsequently. However, non-clinically relevant decrease in bone mineral density (BMD), which stabilized after 12-24 weeks of treatment and thereafter remained stable was observed. Decreases of > 3% were seen in 21% of the patients. Therefore, a DEXA scan is recommended after the first 52 weeks of treatment and as considered appropriate thereafter. Recently published data of 80-week long-term extension study demonstrated Relugolix–Estradiol–Norethisterone Acetate effective, safe, and well-tolerated medical treatment for endometriosis that can be used longer-term, reducing the need for opioids and improving quality of life.
21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	I am not aware of any relevant evidence that might not be found by systematic review of trial evidence

22. How do data on real-world experience compare with the trial data?	There is limited real world experience compared with trial data as only recently been approved by FDA and EMA for treatment of endometriosis associated pelvic pain symptoms
23. NICE considers whether there are any equalities issues at each stage of an evaluation. Are there any potential equality issues that should be taken into account when considering this condition and this treatment? Please explain if you think any groups of people with this condition are particularly disadvantaged. Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics. Please state if you think this evaluation could • exclude any people for which this treatment is or will be licensed but who are protected by the equality legislation • lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population • lead to recommendations that have an adverse impact on disabled people. Please consider whether these issues are different from issues with current care and why. More information on how NICE deals with equalities issues can be found in the NICE equality scheme.	When considering equality issues in patients diagnosed with endometriosis I think it is important to consider: 1.Access to Healthcare: People with limited access to healthcare facilities may face challenges in obtaining treatment which can lead to disparities in health outcomes. Given this is an oral treatment, oral medications are generally more convenient for patients compared to injections, as they can be taken at home without the need for frequent visits to healthcare facilities. This convenience may particularly benefit individuals with transportation barriers, mobility issues, making it easier for them to adhere to the treatment regimen. The most appropriate comparator to Relugolix–Estradiol–Norethisterone Acetate is GnRh agoinsts which are generally administered as an injection. 2. Health literacy and education background: Disparities in health literacy and education can impact understanding, decision-making, and adherence to treatment regimens. Ensuring that information is presented in a clear, culturally competent manner can help mitigate these disparities

Find more general information about the Equality Act and equalities issues here.	

Part 2: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

Click or tap here to enter text.

Relugolix–Estradiol–Norethisterone Acetate demonstrated efficacy in alleviating symptoms such as pelvic pain and menstrual irregularities, improving quality of life for individuals with endometriosis.

Relugolix–Estradiol–Norethisterone Acetate represents a promising alternative to the current standard of care and eliminates the need for frequent clinic visits associated with injectable GnRH agonists

Non-clinically relevant decrease in bone mineral density has been observed, therefore a DEXA scan is recommended after the first 52 weeks of treatment and as considered appropriate thereafter.

Click or tap here to enter text.

Thank you for your time.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

Relugolix–estradiol–norethisterone acetate for treating symptoms of endometriosis [ID3982]

Produced by	Kleijnen Systematic Reviews (KSR) Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University
Authors	Robert Wolff, Managing Director, KSR Ltd, United Kingdom (UK) Isaac Corro Ramos, Health Economics Researcher, Institute for Medical Technology Assessment (iMTA), EUR, the Netherlands (NL) Eline Krijkamp, Health Economist, Erasmus School of Health Policy & Management (ESHPM), EUR, NL Venetia Qendri, Health Economist, ESHPM, EUR, NL Jiongyu Chen, Systematic Reviewer/Health Economist, KSR Ltd, UK Pawel Posadzki, Reviews Manager, KSR Ltd, UK Caro Noake, Senior Information Specialist, KSR Ltd, UK Nigel Armstrong, Health Economics Manager, KSR Ltd, UK Isaac Corro Ramos, Health Economist, ESHPM, EUR, NL Maiwenn Al, Health Economics Researcher, iMTA, EUR, NL
Correspondence to	Kleijnen Systematic Reviews Ltd Robert Wolff Unit 6, Escrick Business Park Riccall Road, Escrick York, YO19 6FD United Kingdom
Date completed	18/12/2023

Source of funding: This report was commissioned by the National Institute for Health Research (NIHR) Evidence Synthesis Programme as project number STA 13/61/26.

Declared competing interests of the authors None.

Acknowledgements

We gratefully acknowledge the expert clinical advice input from Lucky Saraswat MBBS, MRCOG, PhD (Consultant Gynaecologist, Aberdeen Royal Infirmary; Honorary Senior Lecturer, University of Aberdeen) and Professor T Justin Clark MB ChB, MD (Hons), FRCOG (Consultant Gynaecologist, Birmingham Women's and Children Hospital; Honorary Professor of Gynaecology, University of Birmingham; Immediate Past President, British Society for Gynaecological Endoscopy).

Commercial in confidence (CiC) data are highlighted in blue throughout the report.

Confidential comparator prices are highlighted in green throughout the report.

Any de-personalised data are highlighted in pink throughout the report.

Copyright belongs to Kleijnen Systematic Reviews Ltd.

Rider on responsibility for report

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

This report should be referenced as follows:

Wolff R, Corro Ramos I, Krijkamp E, Qendri V, Chen J, Posadzki P, Noake C, Armstrong N, Al M. Relugolix–estradiol–norethisterone acetate for treating symptoms of endometriosis [ID3982]: a Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2023.

Contributions of authors

Robert Wolff acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Isaac Corro Ramos acted as health economic project lead, critiqued the company's economic evaluation and contributed to the writing of the report. Eline Krijkamp, Venetia Qendri and Nigel Armstrong acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Jiongyu Chen and Pawel Posadzki acted as systematic reviewers, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Caro Noake critiqued the search methods in the submission and contributed to the writing of the report. Maiwenn Al acted as health economist on this assessment, critiqued the company's economic evaluation, contributed to the writing of the report and provided general guidance.

Abbreviations

AdViSHE	Assessment of the Validation Status of Heath-Economic decision models tool
AE	Adverse event
AUD	Australian Dollar
B&B	Biberoglu-Behrman
BMD	Bone mineral density
BNF	British National Formulary
BSC	Best supportive care
CADTH	Canadian Agency for Drugs and Technologies in Health
CDSR	Cochrane Database of Systematic Reviews
CE	Cost effectiveness
CEA	Cost effectiveness analysis
CEAC	Cost effectiveness acceptability curve
CI	Confidence interval
CiC	Commercial in confidence
CR	Complete response
CRD	Centre for Reviews and Dissemination
CrI	Credible interval
CS	Company submission
CT	In combination with oestradiol and norethisterone acetate
DBC	Disease Burden Calculator
DSA	Deterministic sensitivity analysis
DSU	Decision Support Unit
EAG	Evidence Assessment Group
ED	Emergency department
EED	Economic Evaluation Database
EHP-30	Endometriosis health profile-30
EQ-5D	European Quality of Life-5 dimensions
EQ-5D-3L	European Quality of Life-5 dimensions 3 levels
EQ-5D-5L	European Quality of Life-5 dimensions 5 levels
ESHPM	Erasmus School of Health Policy & Management
EUCTR	European Union Clinical Trials Register
EUR	Erasmus University Rotterdam
GBP	Great British Pound
GnRH	Gonadotropin-releasing hormone
GP	General practitioner
HCRU	Healthcare resource utilisation
HR	Hazard ratio
HRQoL	Health-related quality of life
HRT	Hormone replacement therapy
HSUV	Health state utility value
HTA	Health Technology Assessment
ICER	Incremental cost-effectiveness ratio
ICTRP	International Clinical Trials Registry Platform
iMTA	Institute for Medical Technology Assessment
KOL	Key opinion leader
KSR	Kleijnen Systematic Reviews Ltd
LNG-IUS	Levonorgestrel intrauterine system
LYs	Life years
LYG	Life years gained
MD	Mean difference
Mg	milligram
mITT	modified intention-to-treat
MRI	Magnetic resonance imaging

N/A	Not applicable
NETA	Norethisterone acetate
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NL	Netherlands
NMA	Network meta-analysis
NMPP	Non-menstrual pelvic pain
NRS	Numerical rating scale
NSAID	Non-steroidal anti-inflammatory drug
OLS	Ordinary least square
OPP	Overall pelvic pain
OR	Odds ratio
OWSA	One-way sensitivity analysis
PCS	Post-conservative surgery
PGA	Patient Global Assessment
PICO	Population, intervention, comparator(s), and outcome measures
PR	Partial response
PRESS	Peer Review of Electronic Search Strategies
PRISMA	Preferred reporting items for systematic reviews and meta-analyses
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
QALY	Quality-adjusted life year
QoL	Quality of life
RCT	Randomised controlled trial
RR	Risk ratio
SAE	Serious adverse events
SE	Standard error
SERM	Selective oestrogen receptor modulator
SLR	Systematic literature review
SMD	Standardised mean difference
SoC	Standard of care
TECH-VER	Technical verification
TPP	Total pelvic pain
TSD	Technical Support Document
UK	United Kingdom
UKCTOCS	United Kingdom Collaborative Trial of Ovarian Cancer Screening
US	United States
USD	United States Dollar
VAS	Visual analogue scale
WHO	World Health Organization

Table of Contents

Table of Tables

Table of Figures

1. EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the Evidence Assessment Group (EAG) as being potentially important for decision making. If possible, it also includes the EAG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues. Section 1.2 presents the key model outcomes. Section 1.3 discusses the decision problem, Section 1.4 issues relate to the clinical effectiveness, and Section 1.5 issues relate to the cost effectiveness (CE). Other key issues are discussed in Section 1.6 while a summary is presented in Section 1.7.

Further details and information on key as well as non-key issues are in the main EAG report, see Sections 3 (decision problem), 4 (clinical effectiveness) and 5 (CE) for more details.

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

1.1 Overview of the EAG's key issues

Table 1.1: Summary of key issues

ID3982	Summary of issue	Report Sections
1	Lack of clarity in the decision problem population.	2.1, all
2	Not all relevant comparators were included.	2.3, all
3	SLR had serious methodological limitations that might have resulted in the omission of important evidence, including that to inform the ITCs.	3.1.1
4	Lack of justification for choice of outcomes, transformation, and source of data for the ITCs.	3.4
5	The link between clinical effectiveness evidence and health economic analyses should be stronger	4.2.9 5.2.1
6	The operationalisation of (in)fertility in the model needs to be carefully considered	4.2.2 4.2.3 4.2.8 5.1
7	The definition and role of BSC in the model should be clarified	4.2.4 4.2.6
8	The number of relevant scenario analyses to test key modelling assumptions is insufficient	4.2.6 4.2.8
9	Model validation efforts are needed to improve model transparency and credibility of results	5.3.3 5.3.4
BSC = best supportive care; ITC = indirect treatment comparison; SLR = systematic literature review		

The EAG was unable to define a new base-case. Assessing most of the uncertainties identified in the cost effectiveness analyses (CEAs) would require major changes to the economic model, which cannot be conducted with the current evidence.

1.2 Overview of key model outcomes

National Institute for Health and Care Excellence (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 new technology is modelled to affect QALYs by:

- Increasing the number of QALYs in "response" health states.
- Reducing the number of QALYs post-hysterectomy.
- In all other health states, the difference in QALYs is not substantial.

Overall, the technology is modelled to affect costs by:

- Its higher unit price compared to current treatments.
- Decreasing costs associated to surgery and health care visits.

The modelling assumptions that have the greatest effect on the ICER are those related to the operationalisation of long-term utility decrements due to loss of fertility.

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

There is lack of clarity of the population addressed in the decision problem and the company submission (CS) (Table 1.2). It is unclear whether all relevant comparators were included (Table 1.3).

Table 1.2: Key issue 1: Lack of clarity in the decision problem population

Report Section	**2.1, all**
Description of issue and why the EAG has identified it as important	Initially, the CS positioned relugolix CT at 2nd line after what was described as conventional hormonal [contraceptive] therapy. In the clarification phase, surgical treatment was added as well. Clinical expert opinion obtained by the EAG indicated that relugolix CT might be suitable for 2nd or 3rd line treatment. It is therefore unclear what the precise nature of the population is in terms of treatment history, which could affect what are the appropriate comparators as well as the nature of subsequent treatment.
What alternative approach has the EAG suggested?	The population should be clearly defined in terms of the place in the care pathway.
What is the expected effect on the CE estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	The population should be clearly defined in terms of the place in the care pathway.
CE = cost effectiveness; CS = company submission; CT = in combination with oestradiol and norethisterone acetate; EAG = Evidence Assessment Group	

Table 1.3: Key issue 2: Not all relevant comparators were included

Report Section	**2.3, all**
Description of issue and why the EAG has identified it as important	The CS focussed on GnRH agonists as the relevant comparator for relugolix CT. Clinical expert opinion obtained by the EAG confirmed GnRH agonists to be appropriate at 2nd line, which is where the company have positioned relugolix CT. In addition, clinical expert opinion obtained by the EAG indicated that linzagolix (Theramex) is another oral peptide GnRH antagonist that should be considered once approved.

Report Section	2.3, all
	As outlined in key issue 1, the precise place in the care pathway is unclear in terms of which medical or surgical treatments have already been experienced. This means that appropriate comparators might have been omitted. Furthermore, serious problems with the SLR (see key issue 3) might mean that evidence to inform the clinical effectiveness of relugolix CT in relation to all relevant comparators via an ITC might have been omitted. This also implies that the results of the economic analyses presented by the company are likely to be invalid.
What alternative approach has the EAG suggested?	The EAG requested that all analyses (ITC and economic) with the appropriate comparators, including analgesia or hormonal treatment if first line and surgery if 2nd line be conducted.
What is the expected effect on the CE estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	Linzagolix (Theramex) should be considered as a comparator once approved. Depending on the line of therapy (treatment history) established in response to key issue 1, further comparators might need to be considered.
CE = cost effectiveness; CS = company submission; CT = in combination with oestradiol and norethisterone acetate; EAG = Evidence Assessment Group; GnRH = gonadotropin-releasing hormone; ITC = indirect treatment comparison; SLR = systematic literature review	

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

The EAG identified a number of serious shortcomings in the systematic literature review (SLR) supporting the CS which might have resulted in the omission of important evidence, including that to inform the indirect treatment comparisons (ITCs; Table 1.4).

As a well conducted SLR is the cornerstone of a methodologically robust appraisal of the available evidence, it is likely that a new SLR, following the relevant guidance, would result in a different evidence base for this CS and could therefore affect the deliberations of the appraisal committee. Therefore, the EAG has not reviewed the evidence presented in this CS and instead highlighted issues which should be addressed when conducting a methodologically robust SLR to underpin a revised CS.

Furthermore, the EAG identified a number of issues with the ITC, especially a lack of justification for choice of outcomes, transformation, and source of data (Table 1.5).

Table 1.4: Key issue 3: SLR had serious methodological limitations that might have resulted in the omission of important evidence, including that to inform the ITCs

Report Section	3.1.1
Description of issue and why the EAG has identified it as important	The EAG identified a number of serious errors and limitations in the SLR that would have affected the overall recall of results and in the EAG's opinion, render the review not fit for purpose. Key areas for concern were raised in the request for clarification and the EAG asked that the searches be rerun and expanded with the identified issues in mind. Whilst acknowledging the failings of the SLR, the company declined to rerun the searches due to time constraints and instead offered to use a Cochrane review by Veth et al. 2023 to supplement the original CS. Disparities in the scope of the Cochrane review and lack of transparency in the running of the update searches leaves the

Report Section	3.1.1
	EAG concerned about both the robustness and appropriateness of the CS.
What alternative approach has the EAG suggested?	As the submission relies on a methodologically robust SLR, the EAG recommend that these concerns be best addressed by a new full SLR designed to cover the final scope as outlined by NICE.
What is the expected effect on the CE estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	As the submission relies on a methodologically robust SLR, the EAG recommend that these concerns be best addressed by a new full SLR designed to cover the final scope as outlined by NICE.
CE = cost effectiveness; CS = company submission; EAG = Evidence Assessment Group; ITC = indirect treatment comparison; NICE = National Institute for Health and Care Excellence; SLR = systematic literature review	

Table 1.5: Key issue 4: Lack of justification for choice of outcomes, transformation, and source of data for the ITCs

Report Section	3.4
Description of issue and why the EAG has identified it as important	The company chose TPP and OPP as outcomes, despite dysmenorrhoea and NMPP being co-primary endpoints in the relugolix CT trials. While TPP and OPP might be valid outcomes, the EAG would argue that the justification for choosing them to inform the economic model is insufficient as all outcomes necessary to inform a comparison of effectiveness should be included, not all of which might be required in the economic model. It Is also unclear why OR was used given that in the economic model comparator response was not estimated (as claimed by the company), but in fact assumed to be equal to the relugolix CT value. Also, response was not informed by either TPP or OPP, but instead by the co-primary endpoints. Given that OR had to be estimated from the original data either on a continuous or ordinal scale, it was not clear to the EAG why the ITC could not have used these data unconverted. It is also the case that the OR calculated in this way is challenging to interpret. It is also unclear that it is superior to using SMD or that the MD might not be used if the measures of pain are considered to be sufficiently similar e.g. VAS and NRS. It is unclear to the EAG where the NRS values for the SPIRIT trials came from, given that these were not reported in the clinical effectiveness Section. NRS was used for several outcomes, including OPP. However, TPP was not explicitly included as an outcome in the SPIRIT trials and the EAG also could not locate TTP in the main clinical effectiveness Section of the CS.
What alternative approach has the EAG suggested?	The EAG requested the ITCs be conducted using SMDs and the original MDs, which the company did.
What is the expected effect on the CE estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	ITCs should be conducted with all outcomes relevant to an assessment of clinical effectiveness, including dysmenorrhoea and NMPP. The scale for estimation should be the original if possible and any transformation adequately justified with a clear presentation

Report Section	3.4
	of the original data, its source and method of transformation. The comments highlighted in key issues 1 to 3 should be considered.
CE = cost effectiveness; CT = in combination with oestradiol and norethisterone acetate; EAG = Evidence Assessment Group; ITC = indirect treatment comparison; MD = mean difference; NMPP = non-menstrual pelvic pain; NRS = numerical rating scale; OPP = overall pelvic pain; OR = odds ratio; SMD = standardised mean difference; TPP = total pelvic pain; VAS = visual analogue scale	

1.5 The CE evidence: summary of the EAG's key issues

A full summary of the cost effectiveness (CE) evidence review conclusions can be found in Section 6.4 of this report. The company's CE results are presented in Section 5, the EAG's summary and detailed critique are in Section 4, and the absence of EAG's exploratory analyses is discussed in Section 6.

The key issues in the CE evidence are discussed in Tables 1.6 to 1.10.

Table 1.6: Key issue 5: The link between the clinical effectiveness evidence and health economic analyses should be stronger

Report Sections	4.2.9 and 5.2.1
Description of issue and why the EAG has identified it as important	The ITC conducted by the company informs only one parameter of the model, and its impact on the model results is negligible. It is unclear whether the studies identified in the cost/resource use search have been used in the model or not. Also, relevant outcomes identified in the final NICE scope might be included in the model (key issue 4).
What alternative approach has the EAG suggested?	Inform CE parameters using clinical effectiveness results as much as possible.
What is the expected effect on the CE estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	Update evidence synthesis and economic model to include missing outcomes from the final NICE scope.
CE = cost effectiveness; EAG = Evidence Assessment Group; ITC = indirect treatment comparison; NICE = National Institute for Health and Care Excellence	

Table 1.7: Key issue 6: The operationalisation of (in)fertility in the model needs to be carefully considered

Report Sections	4.2.2, 4.2.3, 4.2.8, and 5.1
Description of issue and why the EAG has identified it as important	The operationalisation of (in)fertility in the model needs to be reassessed since the current approach (applying a utility decrement to all women post-hysterectomy) seems too simplistic and it has a major impact on the model results.
What alternative approach has the EAG suggested?	Linking women (in)fertility concerns to the treatment pathway. When a utility decrement due to infertility is applied, a more recent estimate should be considered, which should be applied only to the proportion of women actively seeking to have a family (instead to all

Report Sections	4.2.2, 4.2.3, 4.2.8, and 5.1
	of them) and make it age dependent. Estimate model input parameters for this group of patients where possible.
What is the expected effect on the CE estimates?	The ICER is likely to increase.
What additional evidence or analyses might help to resolve this key issue?	Update the model structure. A targeted search for more recent estimates may be needed. As a key driver of the model results, carefully address the uncertainty around implementation in the model (scenario analyses).
CE = cost effectiveness; EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratio	

Table 1.8: Key issue 7: The definition and role of BSC in the model should be clarified

Report Sections	4.2.4, and 4.2.6
Description of issue and why the EAG has identified it as important	The definition and role of BSC in the model is unclear. It is defined as analgesics by the company, but other (?) types of analgesics seem to be included in the model. It is also assumed to be equal to placebo in the SPIRIT trials, but in the trials, placebo do not seem to be defined as analgesics. Also, it is unclear if patients can/should receive BSC after failing to respond (in the current model this seems possible).
What alternative approach has the EAG suggested?	Provide a clear definition of BSC, placebo and analgesics, and how these are used in the model.
What is the expected effect on the CE estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	Update evidence synthesis (if available) and economic model to include effectiveness estimates of BSC in the correct population (BSC after treatment discontinuation).
BSC = best supportive care; CE = cost effectiveness; EAG = Evidence Assessment Group	

Table 1.9: Key issue 8: The number of relevant scenario analyses to test key modelling assumptions is insufficient

Report Sections	4.2.6, and 4.2.8
Description of issue and why the EAG has identified it as important	The number of relevant scenario analyses to test key modelling assumptions is insufficient.
What alternative approach has the EAG suggested?	Explore impact of potentially relevant assumptions such as treatment effect waning, infertility (after update in operationalisation as mentioned in previous issue) or a multiplicative approach when implementing disutilities.
What is the expected effect on the CE estimates?	None.
What additional evidence or analyses might help to resolve this key issue?	Conduct additional scenario analyses.

Report Sections	4.2.6, and 4.2.8
CE = cost effectiveness; EAG = Evidence Assessment Group	

Table 1.10: Key issue 9: Model validation efforts are needed to improve model transparency and credibility of results

Report Section	5.3.3, 5.3.4
Description of issue and why the EAG has identified it as important	The model contains "legacy" parameters and assumptions from previous submissions, which makes external assessment inefficient and difficult. The model seems to produce counterintuitive results (e.g., relugolix CT seems to be more cost effective with lower response rates).
What alternative approach has the EAG suggested?	Remove all "legacy" parameters and assumptions from the model. Explore (and explain) counterintuitive model results.
What is the expected effect on the CE estimates?	Unknown (but likely none).
What additional evidence or analyses might help to resolve this key issue?	An updated economic model.
CE = cost effectiveness; CT = In combination with oestradiol and norethisterone acetate; EAG = Evidence Assessment Group	

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

No other key issues were identified by the EAG.

1.7 Summary of the EAG's view

Considering the key issues above, the EAG was unable to define a new base-case. Assessing most of the uncertainties identified in the CEAs would require major changes to the economic model, which cannot be conducted with the current evidence. The EAG would suggest the company to resolve the key issues presented above as much as possible. It should be emphasised though that the current submission deals with a comparison between a long-term intervention (relugolix in combination with oestradiol and norethisterone acetate (CT) – given for 16 years) against a short-term intervention (GnRH agonists – given for at most 12 months). The focus of the clinical effectiveness Sections of this submission lies on estimating the relative treatment effects of relugolix CT against GnRH agonists for at most 12 months since treatment initiation. When these relative treatment effects are parameterised in the company's economic model, which has a lifetime time horizon, the comparison between relugolix CT versus GnRH agonists is largely irrelevant, since the CE results are then driven by the relative treatment effects of relugolix CT compared to the subsequent treatments after GnRH agonists discontinuation, namely BSC and surgeries. With the currently modelled clinical pathway, BSC and surgeries can be considered as the comparator options for 15 years in the economic model (all treatments are assumed to stop when women become 50 years old). The clinical effectiveness Sections of this submission did not assess how to estimate the relative treatment effects of relugolix CT versus BSC or surgery (after treatment discontinuation) through evidence synthesis. Instead, these relative treatment effects were directly sourced from the SPIRIT trials. Therefore, any changes to the clinical effectiveness Sections, including for example an updated SLR and ITC, are expected to have a minimal impact on the CE results unless the missing comparators can be applied in the long-term. With the current model structure, the EAG anticipates that the only change that can have a major impact on the

model results is the operationalisation of loss of fertility (because it carries a long-term effect in the model).

2. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 2.1: Statement of the decision problem (as presented by the company)

	Final scope issued by NICE	Decision problem addressed in the CS	Rationale if different from the final NICE scope	EAG comment
Population	Adults with symptoms of endometriosis.	Same as scope. The ITC and economic analysis presented in this submission focus on the subgroup of patients who remain symptomatic following treatment with conventional hormonal therapy, including combined hormonal contraception and oral and intra-uterine progestogens.	The company acknowledged that in line with the current UK clinical practice, relugolix CT will be used as a second-line therapy i.e., the population will be narrower than the final scope issued by NICE.	Relugolix CT will be indicated only for women with a history of previous medical or surgical treatment for their endometriosis.
Intervention	Relugolix in combination with oestradiol and norethisterone acetate herein referred to as relugolix CT.	Same as scope	N/A – in line with the NICE final scope.	The intervention is in line with the NICE scope; however, it is unclear whether it can be administered to patients with postmenopausal endometriosis, see Section 2.2.
Comparator(s)	Established clinical management without relugolix in combination with oestradiol and norethisterone, including: • analgesics or NSAID alone or in combination with each other • neuromodulators • hormonal treatment such as combined hormonal	The submission will focus on GnRH agonists as the relevant comparator for relugolix CT.	The company claims that there are no direct, licensed comparators for relugolix CT; and GnRH agonists are the closest comparator in the clinical pathway of care although licensed for use for up to 6 months.	Some potentially relevant comparators have not been included, see Section 2.3.

	Final scope issued by NICE	Decision problem addressed in the CS	Rationale if different from the final NICE scope	EAG comment
	contraception, oral progestogens, GnRH agonists.			
Outcomes	The outcome measures to be considered include: • overall pain • opioid use • analgesic use • recurrence of endometriosis • admission to hospital • subsequent surgical treatment • fertility • adverse effects of treatment • complications of treatment • HRQoL	The outcome measures in the clinical effectiveness Section included: • dysmenorrhoea • non-menstrual pelvic pain • dyspareunia • EHP-30 pain domain • opioid use • analgesic use • EQ-5D-5L • adverse effects The outcome measures in the ITC included: • OPP • TPP The outcome measures in the CE model included: • Dysmenorrhoea • non-menstrual pelvic pain • recurrence of pain • analgesic use • subsequent surgical treatment • subsequent medical treatment • complications related to surgery • HRQoL	Outcomes in the clinical effectiveness, CE Section and the ITC differed.	The outcomes reported are not fully in line with the NICE scope and there is a mismatch between outcomes in different Sections of the CS, see Section 2.4 for details. Of note, in response to the request for clarification, two outcomes were added namely, 1) overall pelvic pain (clinical effectiveness Section) and 2) adverse events (outcomes measures in the CE model)

	Final scope issued by NICE	Decision problem addressed in the CS	Rationale if different from the final NICE scope	EAG comment
Special considerations including issues related to equity or equality		There is evidence to suggest that women from some minority ethnic groups may be underdiagnosed and/or present later for help with endometriosis and thus have more severe symptoms.		The issue of underdiagnosis of endometriosis should be considered.

Based on Table 1 and pages 10 to 14 of the CS[1]

CE = cost effectiveness; CS = company submission; CT = in combination with oestradiol and norethisterone acetate; EAG = Evidence Assessment Group; EHP-30 = Endometriosis health profile-30; EQ-5D-5L = European Quality of Life 5 dimensions 5 levels; GnRH = gonadotropin hormone-releasing hormone; ITC = indirect treatment comparison; N/A = not applicable; NICE = National Institute of Health and Care Excellence; NSAID = non-steroidal anti-inflammatory drug; OPP = overall pelvic pain; TPP = total pelvic pain; UK = United Kingdom

2.1 Population

The population defined in the scope issued by the National Institute for Health and Care Excellence (NICE) is *"adults with symptoms of endometriosis"*.[2] The population in the company submission (CS) is limited to *"the subgroup of patients who remain symptomatic following treatment with conventional hormonal therapy, including combined hormonal contraception and oral and intra-uterine progestogens"*.[1]

In response to the question A12 of the request for clarification, the company acknowledged that in line with the current United Kingdom (UK) clinical practice, relugolix in combination with oestradiol and norethisterone acetate (CT) is only *"indicated for symptomatic treatment of endometriosis in women with a history of previous medical or surgical treatment for their endometriosis"*.[3] Therefore, relugolix CT will be indicated as at least the 2nd treatment of endometriosis i.e., the population should be narrower.

EAG comment: Clinical expert opinion received by the Evidence Assessment Group (EAG) stated that relugolix CT should be used as a 2nd to 3rd line treatment for moderate to severe symptoms of fibroids as well as pain associated with endometriosis. It was noted that the treatment should be carefully considered because of its unfavourable side-effect profile.

As noted by the company, the population in the CS is narrower than in the NICE final scope. However, given the lack of clarity on precise line and previous treatment(s), there is uncertainty as to the population eligible for relugolix CT, which could have profound implications including for relevant comparators and subsequent treatments. Therefore, this constitutes a key issue which should be considered by the committee.[1]

2.2 Intervention

The intervention (relugolix CT) is in line with the NICE final scope.

One tablet of relugolix (40 mg) in combination with oestradiol (as hemihydrate; 1 mg) and norethisterone acetate (0.5 mg) is administered orally, once daily, with or without food (taken with liquid as needed), see Table 2 of the CS.[1]

According to Tables 41 and 87 of the CS, *"patients may continue treatment until discontinuation due to other reasons (e.g., adverse events)"* or the treatment duration is until menopause i.e., average of 16 years in the UK.[1, 4]

According to the company, a dual X-ray absorptiometry scan is recommended 1-year after starting relugolix CT treatment; and in patients with risk of osteoporosis or bone loss, the scan is recommended prior to starting the treatment, see Appendix C of the CS.[5]

EAG comment: It is unclear whether relugolix CT can be administered to women with postmenopausal endometriosis with an estimated prevalence of the condition of 2–5%.[6]

2.3 Comparators

The NICE final scope described the comparators as *"established clinical management without relugolix in combination with oestradiol and norethisterone, including: analgesics or non-steroidal anti-inflammatory drug (NSAID) alone or in combination with each other neuromodulators, hormonal treatment such as combined hormonal contraception, oral progestogens, gonadotropin-releasing hormone (GnRH) agonists"*.[2]

The CS focuses on gonadotropin-releasing hormone (GnRH) agonists as the relevant comparator for relugolix CT; and claims that there are no direct, licensed comparators for relugolix CT.[1] In response to question A13 of the request for clarification, the company acknowledged that *"relugolix CT is indicated for symptomatic treatment of endometriosis in women with a history of previous medical or surgical treatment for their endometriosis"*.[3] It is unclear which surgical treatments the company is referring to, as numerous are available including hysterectomy (removal of the uterus) following which relugolix CT would not be recommended.[7]

EAG comment: Clinical expert opinion obtained by the EAG confirmed that GnRH agonists are appropriate comparators, but that linzagolix (Theramex) is another oral peptide GnRH antagonist that could be used in clinical practice (once approved).

As stated in Section 2.1, the precise place in the care pathway is unclear in terms of which medical or surgical treatments have already been experienced. This means that appropriate comparators might have been omitted, which would have implications for relative clinical effectiveness and cost effectiveness (CE).

In addition, the EAG noted several issues with the searches conducted for the systematic literature review (SLR), including the omission of potentially relevant comparators, see Section 3.1.1 for details. Uncertainty as to comparators is therefore a key issue.

2.4 Outcomes

The NICE final scope lists the following outcome measures:

- overall pain
- opioid use
- analgesic use
- recurrence of endometriosis
- admission to hospital
- subsequent surgical treatment
- fertility
- adverse effects of treatment
- complications of treatment
- health-related quality of life (HRQoL).

EAG comment: The following outcomes were not collected in the SPIRIT trials: hospital admission, fertility, recurrence of endometriosis, or complications of treatment. Those same outcomes are also missing in the CE and the indirect treatment comparison (ITC).

2.5 Other relevant factors

The issue of underdiagnosis of endometriosis in people from minority ethnic groups should be considered.

3. CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

3.1.1 Searches

The following paragraphs contain summaries and critiques of the searches related to clinical effectiveness presented in the CS.[1, 5] The Canadian Agency for Drugs and Technologies in Health (CADTH) evidence-based checklist for the Peer Review of Electronic Search Strategies (PRESS), was used to inform this critique.[8, 9] The EAG has presented only the major limitations of each search strategy in the report.

A summary of the sources searched is provided in Table 3.1.

Table 3.1: Data sources for Appendix D: Identification, selection and synthesis of clinical evidence (as reported in CS)

Resource	Host/Source	Date Span	Date searched/Updated
Electronic databases			
Embase	Ovid (update via Embase.com)	Original: unclear Update: 2022/4/1-2022/11/1	8.4.22 Update 1.11.22
MEDLINE (unclear if this includes Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations and Daily)	Ovid (Update via PubMed)	Original: unclear Update: 2022/4/1-2022/12/1	8.4.22 Update 1.12.22
Trials registries			
ClinicalTrials.gov		Original: unclear Update: 2022/4/1-2022/12/1	8.4.22 Update 12.12.22
EUCTR		Original: unclear Update: 2022/4/1-2022/12/1	8.4.22 Update 12.12.22
WHO ICTRP		Original: unclear Update: 2022/4/1-2022/12/1	8.4.22 Update 12.12.22
Health Canada's Clinical Trials Database		Original: unclear Update: 2022/4/1-2022/12/1	8.4.22 Update 12.12.22

Resource	Host/Source	Date Span	Date searched/Updated
Additional searches			
Google and Google Scholar			
Hand searching reference lists of the included trials			
Based on the CS and Appendix D of the CS[1, 5] CS = company submission; EUCTR = European Union Clinical Trials Register; ICTRP = International Clinical Trials Registry Platform; WHO = World Health Organization			

EAG comment:

- Searches were undertaken in April 2022 to identify efficacy and safety outcomes of treatments for endometriosis and were updated in December 2022. Searches were carried out on MEDLINE and Embase with additional searches of four trials registers and a search of Google and Google Scholar. Where appropriate, strategies utilised a randomised controlled trial (RCT) study design filter. The EAG queried a number of issues with the reporting of the searches, which for the most part were addressed at clarification.[3, 10] However, some areas (such as the date span for the Embase searches and odd syntax in the PubMed search) remained unclear, which would affect any future attempts to reproduce the searches. The EAG would remind the company that best practice recommends that *"… bibliographic database search strategies should be copied and pasted into an appendix exactly as run and in full, together with the search set numbers and the total number of records retrieved by each search strategy. The search strategies should not be re-typed, because this can introduce errors"*.[11]
- The company confirmed that no separate searches were undertaken to retrieve information regarding adverse events (AEs) for safety outcomes for relugolix. Given that the main clinical effectiveness searches were restricted to RCTs, guidance by the Centre for Reviews and Dissemination (CRD) recommends that if searches have been limited by a study design filter, additional searches should be undertaken to ensure that adverse events that are long-term, rare or unanticipated are not missed.[12]
- Whilst searches were reported for MEDLINE and Embase, with additional pragmatic grey literature searches, the EAG queried the absence of any searches of the Cochrane Library (either Cochrane Database of Systematic Reviews (CDSR) or CENTRAL). The company responded that *"Cochrane Reviews and Editorials are indexed in PubMed, and searched were run in PubMed. Therefore, no separate search of the Cochrane Library was carried out"*.[3] While NICE no longer has a minimum requirement of databases to search, best practice for systematic reviews suggests searching a range of databases to maximise recall and minimize bias, with regard to CENTRAL in particular as *"CENTRAL, however, also includes citations to reports of randomized trials that are not included in MEDLINE, Embase or other bibliographic databases; citations published in many languages; and citations that are available only in conference proceedings or other sources that are difficult to access. It also includes records from trials registers and trials results registers beyond ClinicalTrials.gov and the WHO portal"*.[11]

A key issue for the EAG was the number of limitations and errors identified within the main search strategies, that may have accounted for the low recall of results (Embase = 509 hits). Many of the noted weaknesses were repeated throughout the clinical and economic searches. Taking the primary 2022 Embase search strategy reported for the clinical effectiveness searches in Appendix D[5] as an example, here are some of the key areas of concern:

- Conditions facet
  - This facet only contains subject headings, no free text terms.
  - The main subject heading for endometriosis was not exploded (although this did appear in the update searches)
  - Records containing subject headings for adenomyosis / uterus myoma / and ovary cancer/ were excluded from the search results using the Boolean operator NOT. The use of NOT is generally not recommended. If included in a strategy it should always be used with extreme caution, as it can easily remove relevant records containing both terms. In this example, the EAG requested that the strategy be amended to remove the use of the NOT operator in this facet.
- Interventions facet
  - Missing subject heading and synonyms for relugolix, although the use of the '.mp.' field tag may have negated some loss of recall.
  - Failure to explode subject headings for 'oral contraceptive agent/' (misses trade names of individual products) and 'contraceptive agent/' (misses subheading 'hormonal contraceptive agent/' and other subject headings below it in the EMTREE hierarchy)
  - Further missing relevant free text and subject headings for named comparators.
- Pain facet
  - Given the low number of results for Embase (n=509) the inclusion of a pain facet appeared both unnecessary and overly restrictive. The EAG recommend that this facet be removed.

3.1.1.2 Addressing identified issues related to the searches

Given the lack of relevant papers found, the EAG requested that the main Embase and MEDLINE clinical searches be rerun and expanded with the above points in mind and the resulting new papers screened for potentially relevant references.[10]

The company declined to rerun the searches citing the significant amount of time and delay to the appraisal that this would require.[3] However, partly acknowledging the aforementioned issues, the company proposed utilising a recently published Cochrane review by Veth et al. 2023 that assessed the efficacy and safety of GnRH agonists for the treatment of painful symptoms associated with endometriosis.[13] This would involve carrying out and reporting on a feasibility assessment of any studies in this review that were not identified in the submitted clinical SLR.[3] The company also proposed running update searches based on the Cochrane review search strategies (original Cochrane searches were carried out in May 2022).

EAG comment: As detailed below, the EAG raised concerns regarding the appropriateness of this course of action, not least as it is important that submission to NICE utilise robust methods i.e., include a SLR to support each submission, rather than adapting a previous systematic review which can create a range of issues, as discussed below.

3.1.1.2.1 Cochrane review

- The searches reported for the Cochrane review were clear and reproducible and contained a broad facet for endometriosis.[13] However, the interventions facet did not contain specific terms for relugolix and the company stated that *"the Cochrane search strategy included GnRH antagonist terms and thus would be expected to include any Relugolix CT studies"*.[14] The EAG would recommend including these as a separate search. The interventions facet only contained

terms for GnRH as used in the treatment of symptomatic endometriosis as per the scope of the Cochrane review, alternative treatments such as non-steroidal anti-inflammatory drugs (NSAIDs) and hormonal contraceptives, named in the NICE final scope for the CS, were not actively searched for.[2, 13]

- In their response to the request for clarification, the company compared the final scope issued by NICE with that of the Cochrane review in terms of eligible population, intervention, comparator(s) and outcome measures (PICO).[3] There are noticeable differences between the NICE final scope and the Cochrane review which makes PICO incomparable between the two.[2, 13]
    - For instance, the Cochrane review included analgesics, calcium-regulating agents, hormonal treatment (gestrinone, progesterone, danazol), placebo, no treatment as relevant comparators; and omitted e.g., neuromodulators or NSAIDs listed in the final scope by NICE.[13] The Cochrane review excluded surgical therapies, the combined oral contraceptive pill, progesterone receptor modulators or selective oestrogen receptor modulators (SERMs) or GnRH antagonists as relevant interventions/comparators, and all potentially eligible for the ITC.
    - In terms of outcomes, only overall pain associated with endometriosis and adverse effects were measured in the Cochrane review; and opioid use, analgesic use, recurrence of endometriosis, admission to hospital, subsequent surgical treatment, fertility, complications of treatment and HRQoL were not, however these were listed in final scope by NICE.[2, 13]

  Hence, the EAG does not agree with the company's statement that *"the only significant difference between the two is that the Cochrane review did not include relugolix CT"*.[3] The EAG does not fully understand exclusion of neuromodulators, which in company's view *"in general comprise medical devices such as vagus nerve stimulators"* (page 5 of the response to the request for clarification).[3] Neuromodulators, also known as neuropathic analgesics (i.e., drugs not devices), are used in the management of chronic, persistent pain; and primarily affect the central nervous system's modulation of pain, rather than peripheral meditators of inflammation.[15]
- Update searches for the Cochrane review were performed for the period from May 2022 to November 2023, despite reporting that *"the search strategy used was identical to the Cochrane review",*[14] the EAG was concerned that results for the MEDLINE search in particular appeared lower than expected. The EAG reran the Cochrane MEDLINE strategy using different combinations of the MEDLINE segments (i.e., Epub ahead of print, In-process etc.) with different date limits, and all yielded higher results than reported, however without full search strategies for the updates the EAG is unable to assess if strategies were correctly entered with appropriate date limits. Again, the EAG refers to the importance of providing full details of all searches in order to ensure that the methods used are both clear and reproducible as recommended by the Cochrane Handbook.[11]

3.1.2.1.2 Additional searches

- In the additional clarification documents received on the 16 November 2023, the company reported that as well as the original SLR a *"a pragmatic literature review which we had not explicitly reported in our submission"* was also carried out.[14] However, the network meta-analysis (NMA) feasibility assessment which was provided as evidence did not contain any search strategies. The methods Section reported that *"the methods and results of the SLRs are described in a separate report (data on file). The pragmatic literature search was done by*

searching the web using key words related to GnRH agonist therapies used to treat moderate-to-severe pain associated with endometriosis".[16] With no further information provided, the EAG is unable to assess the possible contribution that this may have made to the overall recall of results.

3.1.2.1.3 Summary of EAG view on attempts to address issues related to the searches

Despite the additional work carried out by the company at clarification, the EAG remains very concerned that the lack of clarity and appropriateness of search methods may have severely affected the robustness of the CS.

Best practice states the importance of well conducted and reported search methods i.e.,

- Guidance by CRD highlight that *"conducting a thorough search to identify relevant studies is a key factor in minimizing bias in the review process. The search process should be as transparent as possible and documented in a way that enables it to be evaluated and reproduced"*,[12]
- Authors of the ROBIS tool for assessing risk of bias in systematic reviews state that *"systematic flaws or limitations in the design or conduct of a review have the potential to bias results"* and that *"a sensitive search to retrieve as many eligible studies as possible is a key component of any systematic review"*,[17] while
- The Cochrane Handbook states that *"systematic reviews require a thorough, objective and reproducible search of a range of sources to identify as many eligible studies as possible (within resource limits). This is a major factor distinguishing systematic reviews from traditional narrative reviews, which helps to minimize bias and achieve more reliable estimates of effects and uncertainties"*.[11]

Therefore, the EAG would strongly recommend that these concerns be best addressed by a new full SLR designed to cover the final scope as outlined by NICE.

3.1.2 Inclusion criteria

This and the following Sections summarise and critique the further clinical effectiveness SLR methods. An initial SLR was conducted on 8 April 2022 with an update on 1 December 2022 to identify relevant clinical evidence. Full details of the SLR search strategy, study selection process and results were reported in Appendix D.[5] The eligibility criteria for the SLR are presented in Table 3.2 below. Of note, the EAG critique of the conducted searches, a key issue, is presented in Section 3.1.1.

Table 3.2: Eligibility criteria used in the SLR

Parameter	Original SLR		Updated SLR	
	Inclusion Criteria	**Exclusion criteria**	**Inclusion Criteria**	**Exclusion criteria**
Population	• Premenopausal women with a clinically confirmed diagnosis of EM • Women with EM-associated pain (no restriction applied on pain severity)	• Asymptomatic women with EM • Studies without specific results for the EM population	• Premenopausal women with a clinically confirmed diagnosis of EM • Women with EM-associated pain (no restriction applied on pain severity)	• Asymptomatic women with EM • Studies without specific results for the EM population
Intervention	Relugolix CT 40 mg with add-back therapy	-	Relugolix CT 40 mg with add-back therapy	Studies reporting data related to other GnRH antagonists (elagolix, linzagolix, oligolix etc.)

Comparators[a]	Trials including any pharmacological treatments/ surgeries used to manage EM-associated pain:
	Hormonal:
	• Combined and progestin-only contraceptives
	• LNG-IUS
	• GnRH agonists and antagonists with or without add-back therapies
	Surgery:
	• Conservative procedures:
	  • Surgical ablation/excision
	  • Ovarian cystectomy
	  • Laparoscopy
	• Definitive procedures:
	  • Removal of endometrioma
		• Dosing studies in which no control or comparator is used
		• Studies comparing different operative methods (i.e., microsphere sizes, types of sutures)
		• Studies comparing complimentary therapies (i.e., yoga, physical activity, diet, mindfulness, pelvic floor physiotherapy)
			Trials including the following treatments used to manage EM-associated pain:
			Hormonal:
			• Combined and progestin-only contraceptives
			• LNG-IUS
			• GnRH agonists with or without add-back therapies
			Surgery:
			• Laparotomy
			• Endometrial ablation techniques
			Analgesics
			Placebo
				• Dosing studies in which no control or comparator is used,
				• Studies comparing different operative methods (i.e., microsphere sizes, types of sutures)
				• Studies comparing complimentary therapies (i.e., yoga, physical activity, diet, mindfulness, pelvic floor physiotherapy)
				• Studies reporting data related to any other class of drug (For e.g.: anaesthetics,

Parameter	Original SLR		Updated SLR	
	Inclusion Criteria	**Exclusion criteria**	**Inclusion Criteria**	**Exclusion criteria**
	• Abdominal or vaginal hysterectomy • Salpingectomy/tuboplasty • Oophorectomy Control groups: • Placebo			hemorheological agents etc.)
Outcomes	Studies aiming to assess the efficacy/safety of the included treatments and reporting at least one of the efficacy outcomes of interest or safety outcomes Efficacy outcomes: • Dysmenorrhea • Dyspareunia • Non-menstrual pelvic pain • Overall pelvic pain • EM health profile, overall and by domain subscale • HRQoL using the EQ-5D form Safety outcomes: • Prevalence of AEs • Prevalence of SAEs • Prevalence of fatal AEs • Bone mineral density loss • Prevalence of hot flashes • Change in low-density lipoprotein cholesterol from baseline Tolerability outcomes: • Discontinuation (all-cause)	• Trials without any outcome of interest • Trials only reporting surgical-related outcomes (i.e., surgical pain, post-surgery pain)	Studies aiming to assess the efficacy/safety of the included treatments and reporting at least one of the efficacy outcomes of interest or safety outcomes Efficacy outcomes: • Dysmenorrhea • Dyspareunia • Non-menstrual pelvic pain • Overall pelvic pain • EM health profile, overall and by domain subscale • HRQoL using the generic measures (like SF-36, EQ-5D, NHP, WHOQOL-BREF, The Duke Health Profile, 15D, QLI) or endometriosis specific measures (like EHP-30, EHP-5) or self-developed specific scales like Colwell scale, Bodner scale, etc. Safety outcomes: • Prevalence of AEs • Prevalence of SAEs	• Trials without any outcome of interest • Trials only reporting surgical-related outcomes (i.e., surgical pain, post-surgery pain)

Parameter	Original SLR		Updated SLR	
	Inclusion Criteria	**Exclusion criteria**	**Inclusion Criteria**	**Exclusion criteria**
	• Discontinuation due to AEs		• Prevalence of fatal AEs • Bone mineral density loss • Prevalence of hot flashes • Change in low-density lipoprotein cholesterol from baseline Tolerability outcomes: • Discontinuation (all-cause) • Discontinuation due to AEs	
Study design	• RCTs that are comparative and prospective • Accepted control groups will be placebo or active control • No restriction on type of statistical hypothesis (e.g., non-inferiority trials)	• Cross-over trials (due to the potential reversibility of EM treatments) • Non-comparative trials • Preclinical/non-human studies • Clinical trials without publicly available or published results • Study protocol • Observational studies • Reviews (SLR, meta-analysis, NMA, narrative reviews, scoping reviews, etc.) • Descriptive studies (case report, case series) • Expert opinions, editorials, letters to editors	• RCTs that are comparative and prospective • Accepted control groups will be placebo or active control • No restriction on type of statistical hypothesis (e.g., non-inferiority trials)	• Cross-over trials (due to the potential reversibility of EM treatments) • Non-comparative trials • Preclinical/non-human studies • Clinical trials without publicly available or published results • Study protocols • Observational studies • Reviews (SLR, meta-analysis, NMA, narrative reviews, scoping reviews, etc.) • Descriptive studies (case report, case series) • Expert opinions, editorials, letters to editors

Parameter	Original SLR		Updated SLR	
	Inclusion Criteria	**Exclusion criteria**	**Inclusion Criteria**	**Exclusion criteria**
Other reasons	• Study type: Original publications only • Language: No restriction (except for searches conducted in Google and Google Scholar that were limited to French and English) • Time restriction: No restriction • Geographical scope: No restriction	• Language: N/A • Time restriction: N/A • Geographical scope: N/A	• Study type: Original publications only • Language: No restriction (except for searches conducted in Google and Google Scholar that were limited to English) • Time restriction: No restriction • Geographical scope: No restriction	• Language: N/A • Time restriction: N/A • Geographical scope: N/A

Based on Table 91 of Appendix D[5]

[a] The review will not be limited to studies only involving the intervention of interest (relugolix-CT). Any trial including a comparison involving relugolix-CT or a comparator of interest will be considered.

AEs = adverse events; CT = combined therapy; EM = endometriosis; EQ-5D = European Quality of Life 5 Dimension; GnRH = gonadotropin-releasing hormone; HRQoL = health-related quality of life; LNG-IUS = levonorgestrel intrauterine system; NMA = network meta-analysis; N/A = not applicable, RCTs = randomised controlled trials; SAEs = serious adverse events; SLR = systematic literature review

EAG comment:

- The EAG noted that Appendix D indicates no restriction by line of therapy.[5] However, analgesia is not listed as comparator. Various types of surgery are listed as comparators, but no study of surgery was included. The company was asked to conduct a systematic review that is consistent with the population and comparators of the decision problem (see Table 2.1) as requested in questions A12 and A13 and where the studies included are consistent with the eligibility criteria (clarification question A16).[10] In response, the company stated that *"the list of interventions and comparators was agreed with our Gedeon Richter global colleagues: only key terms related to surgical treatment actually used in the electronic search strategy were reported in the PICOS (i.e., Laparotomy and Endometrial Ablation Techniques)"*.[3] As the eligibility criteria should be consistent with the population and comparators specified in the decision problem the EAG does not consider this to be a reasonable explanation.
- In clarification question A17, the EAG asked the company to clarify on differences/discrepancies between the original and the updated SLR in terms of interventions, comparators, and outcomes e.g., an exclusion of GnRH antagonists in the updated SLR.[10] In response, the company stated that *"GnRH antagonists were removed from the updated SLR as those treatments are not available in Europe/UK (note that the original SLR was conducted with a US scope)"*.[3] According to the clinical expert opinion obtained by the EAG, linzagolix (Theramex), another oral peptide GnRH antagonists, as well as other off licence nasal or parenteral GnRH analogues and addback hormone replacement therapy (HRT) could be used in clinical practice in England and Wales. This means that potentially relevant comparators have not been included in the CS which has been identified as a key issue.

3.1.3 Critique of study selection and data extraction

Two reviewers, working independently, reviewed all title/abstract and full-text screening. Only the reasons for exclusion at full-text screening were documented and reported. Any discrepancy in study selection between reviewers were resolved by consensus or with the help of a third reviewer.

No details of the data extraction process or plan were provided in the CS[1] or the associated appendices[5].

EAG comment: This is of concern to the EAG considering the fundamental importance of conducting a robust SLR and identifying the relevant evidence. While it is not suggested that data extraction may be impacted by bias or error, the lack of reporting around this area means that an elevated risk must be assumed. The EAG emphasises the importance of robust methodologies and clear and descriptive reporting.

3.1.4 Quality assessment

According to the CS (Appendix D)[5], risk of bias for all studies that were included in the final networks was assessed using the ROB tool.[5, 18] Quality assessment of the Relugolix CT clinical trials (SPIRIT 1, SPIRIT 2 and SPIRIT OLE) using the York Centre for Reviews and Dissemination guidance for undertaking reviews in healthcare.

The CS did not include details of the quality assessment process (i.e., the number of reviewers involved and the approach for resolving disagreements).

EAG comment: The company was asked to clarify who adapted the quality assessment tool for studies included in the SLR and how the adaptation was done as well as to indicate whether one or two independent reviewers performed the quality appraisals (clarification question A18).[10] The company

responded by stating that *"we believe the EAG is referring to Appendix D1.3 of the company submission, which describes the quality assessment of the studies included in the ITC. This was done using the template in Section 2.5 of the user guide to the company evidence submission template as provided by NICE upon invitation to participate. The assessments were carried out by one independent reviewer and checked by a second".*[3] It is optimal to conduct such tasks in duplicate, independently, and where disagreements exist, to resolve by consensus or by intervention of a third reviewer. The EAG is satisfied with the tools used by the company, but remains concern about the process used for of risk of bias assessment. The potential for bias in the assessment process cannot be discounted.

3.1.5 Evidence synthesis

No meta-analysis was performed as stated in Section B.2.8 of the CS *"Not applicable"*.[1] According to Appendix D of the CS, NMAs were performed using the gemtc package in R and JAGS software.[5] The NMAs were based on published information from clinical trials identified by a systematic literature search and a pragmatic literature search. An ITC was conducted to compare the efficacy of relugolix CT with comparator therapies for the treatment of endometriosis-associated pain. Further details are provided in Section 3.3 of this report.

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

The population considered in the CS is also narrower than the NICE scope i.e., the two phase 3 trials (SPIRIT 1 and 2), and an open-label phase 3 extension (SPIRIT OLE) only included pre-menopausal people aged 18 to 50 years with endometriosis (CS, Table 3, page 31).[1] However, post-menopausal people with endometriosis were not considered eligible in the SPIRIT trials; and they also suffer from the condition.[19]

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

The company stated that since direct head-to-head RCT data are not available, an ITC was conducted to compare the efficacy of relugolix CT with comparator therapies for the treatment of endometriosis-associated pain.[1]

From the 58 studies found in the SLR, to be included in the ITC, studies had to fulfil the following criteria:

- Directly connect a comparator of interest to the intervention relugolix CT, or
- Indirectly connect relugolix CT with a comparator of interest (e.g., through placebo).

3.3.1 Overall pelvic pain

Three placebo-controlled trials were included in the network, SPIRIT 1 and 2 for relugolix CT and D'Hooghe et al. 2019 to compare with ASP1707 (opigolix; 3 mg, 5 mg, 10 mg, 15 mg) and leuprolide acetate (3.75 mg).[1]

3.3.2 Total pelvic pain

Four RCTs were included in the network.[1] Three were placebo-controlled trials, SPIRIT 1 and 2 for relugolix CT and Lang et al 2018 to compare with dienogest 2 mg. Strowitzki et al. 2010 was also included via its connection to dienogest 2 mg in order to compare with leuprolide acetate (3.75 mg).

EAG comment: The EAG refers to the serious problems with the SLR discussed in Section 3.1.1, which means that there is little confidence in the composition of these networks and therefore any estimate of effectiveness of relugolix CT versus an appropriate comparator (see also Section 2.3). This further supports the need to identify the SLR problems as a key issue.

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

3.4.1 Indirect treatment comparison methods

The outcomes that were estimated were overall pelvic pain (OPP) and total pelvic pain (TPP). The company explained that dysmenorrhea could not be estimated due to *"…inconsistencies in the this was measured across the trials. However, dysmenorrhea was captured in the ITC as an element of the TPP endpoint. In addition, analgesic and opioid use were not included in the ITC owing to the amount of heterogeneity between studies in terms of permitted use and reporting of use"* (page 88 of the CS).[1]

The company also stated that *"the outcomes of interest, TPP and OPP, were selected to inform an economic model of REL-CT and other GnRH agonists for endometriosis-related pain"* (page 70 of the CS appendices).[5] The company acknowledged that most trials reported endometriosis related pain as a continuous measure, summarised either as a mean or mean change from baseline at a particular follow-up time. However, the treatment effect based on these was converted to an odds ratio (OR) because of how treatment effect was estimated in the model: *"The probability of response for a comparator in the cost-effectiveness model for REL-CT is estimated using the relative treatment effect of REL-CT vs the comparator and the probability of response estimated from the SPIRIT 1&2 trials for REL-CT"* (page 74 of the CS appendices).[5]

It was stated that OR was calculated using the formula:[5]

$$ln(OR) = \frac{\pi}{\sqrt{3}} SMD$$

where standardised mean difference (SMD) is SMD in pain.

The company stated that this was recommended in the Cochrane Handbook.[11] The company also stated that there might be conversion from a non-continuous outcome: *"If OPP reported using the NRS scale was presented on an ordinal outcome scale using counts for each item, a responder threshold of <4 was used to convert the results into responder vs non-responder which was used to calculate the odds ratio. The rationale for the above responder definition was discussed and agreed on during an advisory board meeting involving clinical experts within endometriosis, where the clinical experts agreed that the responder definition should be corresponding to achieving absence or mild pain (66)"* (page 74 of the CS appendices).[5]

SPIRIT 1 and 2 trial data were pooled, although the method of pooling was not reported.[1, 5]

A random effects model with what the company described as a *"weakly informative prior"* (page 72 of the CS appendices) for the between trial variance was used in the base-case.[5] As a sensitivity analysis, an empirical, i.e., informative prior was used, assuming a log-normal distribution for σ with a mean of -3.23 and standard deviation of 1.88, *"…corresponding to the pharmacological vs pharmacological intervention comparison type and the semi-objective outcome type derived by Turner et al., 2012 (160)."* (page 72 of the CS appendices)[5] A fixed effect model was used in a sensitivity analysis.

3.4.1.1 OPP

Table 104 in Appendix D shows the trial data used to estimate OR based on numerical rating scale (NRS) at baseline, 12 Weeks and 24 Weeks, as well as change from baseline, although there were no 24-Week data for D'Hooghe et al. 2019.[5] The ITC was therefore conducted only at 12 Weeks. All D'Hooghe et al. 2019 data means were graphically estimated.

3.4.1.2 TPP

Table 106 in Appendix D shows the trial data used to estimate OR, which is at baseline, 12 Weeks and 24 Weeks, as well as change from baseline, although there were no 12-Week data for Lang et al 2018 or Strowitzki et al 2010.[5] The ITC was therefore conducted only at 24 Weeks. The nature of the data varied with mean 'score' used for the SPIRIT 1 and 2 studies and Lang et al 2018, as opposed to the Biberoglu-Behrman (B&B) score (percentages in each of the five categories) for Strowitzki et al 2010. Data for Strowitzki et al 2010 were graphically estimated.

EAG comment: Despite dysmenorrhoea being a co-primary endpoint, it was not included in a network and no evidence for inconsistency in measurement as purported was provided in the CS or the appendices. In fact, the other co-primary endpoint, non-menstrual pelvic pain (NMPP), was also not included without providing a justification. While TPP and OPP might be valid outcomes, the EAG would argue that the company's justification for choosing them to inform the economic model is insufficient as all outcomes necessary to inform a comparison of effectiveness should be included, not all of which might be require in the economic model. It is also unclear why OR was used given that in the economic model comparator response was not estimated as claimed by the company, but in fact assumed to be equal to the relugolix CT value. Also, response was not informed by either TPP or OPP, but instead by the co-primary endpoints, which the company had ruled out of any network, as reported in Section 3.3 of the CS.[1] Because they are used in the economic model, the definitions of treatment response are critiqued in Section 4.2.6, but it is important to note that, although derived at least partly from the co-primary endpoints, they were not measures of effectiveness as reported in the SPIRIT trials. Given OR had to be estimated from the original data either on a continuous or ordinal scale, it was not clear to the EAG why the ITC could not have used these data unconverted. It is also the case that the OR calculated in this way is challenging to interpret, as described in the paper cited by the company.[11, 20] It is also unclear that it is superior to using the SMD or that the mean difference (MD) might not be used if the measures of pain are considered to be sufficiently similar e.g., visual analogue scale (VAS) and NRS.

It is unclear to the EAG where the NRS values for the SPIRIT trials came from given that they were not reported in the clinical effectiveness Section, perhaps because they were pooled.[1] A NRS was used for several outcomes, but those did include OPP. However, TPP was not explicitly included as an outcome in the SPIRIT trials despite a statement in Appendix D that *"SPIRIT 1&2 were the only studies that reported both OPP and TPP"* (page 72).[5] The EAG also could not locate TTP in the main clinical effectiveness Section of the CS.

The EAG therefore requested this based both on the MD and the SMD, which the company provided in their response to clarification.[3] The company also stated that the ITC using ORs had been updated with actual values as opposed to those graphically estimated from D'Hooghe et al. 2019. These updated values and those using the MD and SMD were supplied in supplementary documents.[16, 21, 22]

The EAG note that the statistical model chosen in the base-case and the sensitivity analyses are acceptable and consistent with the NICE Technical Support Document (TSD) 3.[23] However, the lack of justification for choice of outcomes, transformation and source of data for the ITCs constitute a key

issue while other key issues e.g., placement in the care pathway, choice of relevant comparators, and conduct of SLR are relevant as well.

3.4.2 Indirect treatment comparison results

3.4.2.1 OPP

The OR point estimate of relugolix CT versus placebo favoured relugolix CT, but was against relugolix CT versus leuprolide acetate.[1] This was the case in the base-case analysis (random effects, weakly informative priors) and the sensitivity analyses (empirical priors or fixed effects). However, the 95% credible intervals (CrI) were very wide and overlapped the point of no difference in all analyses except fixed effect versus placebo.

3.4.2.2 TPP

As for OPP, the OR point estimate of relugolix CT versus placebo favoured relugolix CT, but was against relugolix CT versus leuprolide acetate.[1] This was the case in the base-case analysis (random effects, weakly informative priors) and the sensitivity analyses (empirical priors or fixed effects). The 95% CrIs were very also very wide and overlapped the point of no difference in all analyses, although the exceptions were versus placebo with the empirical priors and versus dienogest and leuprolide acetate as well as placebo with the fixed effect model.

3.4.2.3 Combined OPP and TPP

The pattern of results was very similar to those for TPP, regardless of whether OPP or TPP data from SPIRIT 1 and 2 were used in the network, the only exceptions being overlap of the 95% CrI versus all but placebo for the fixed effect model.

EAG comment: The results indicate that there appears to be little difference between relugolix CT and the two comparators, dienogest and leuprolide acetate and this was generally the case in the updated analyses, which included those based on the MD and SMD, supplied in response to clarification.[3, 16, 21, 22]

Nevertheless, the EAG have not reproduced the results of the ITCs largely because of the problems with the SLR, as well as the substantive issues with the ITC methodology, as already mentioned.

3.5 Additional work on clinical effectiveness undertaken by the EAG

No additional work on clinical effectiveness was undertaken.

3.6 Conclusions of the clinical effectiveness Section

In the original CS the company presented details of an SLR undertaken in April 2022 and updated in December 2022.[1, 5] The EAG noted a number of serious errors and limitations in the SLR that would have affected the overall recall of results and in the EAG's opinion, render the review not fit for purpose. Key areas for concern were raised in the request for clarification and the EAG asked that the searches be rerun and expanded with the aforementioned points in mind. Whilst acknowledging the failings of the SLR the company declined to rerun the searches due to time constraints and instead offered to use a previously identified Cochrane review by Veth et al. 2023.[13] to supplement the original CS.[3] However, despite this additional work, disparities in the scope of the Cochrane review and lack of transparency in the running of the update searches leave the EAG concerned about both the robustness and appropriateness of the company's submission and they believe that a properly conducted full systematic review would be the best way forward. Further to this the EAG would note that both the original SLR and subsequent responses to clarification contained a number of errors and omissions in the reporting

of the search methods that hampered the EAG's ability to fully critique the search methods. Best practice states the importance of well conducted and reported search methods i.e., *"conducting a thorough search to identify relevant studies is a key factor in minimizing bias in the review process. The search process should be as transparent as possible and documented in a way that enables it to be evaluated and reproduced"*.[12]

As a well conducted SLR is the cornerstone of a methodologically robust appraisal of the available evidence, it is likely that a new SLR, following the relevant guidance, would result in a different evidence base for this CS and could therefore affect the deliberations of the appraisal committee. Therefore, the EAG has not reviewed the evidence presented in this CS and instead highlighted issues which should be addressed when conducting a methodologically robust SLR to underpin a revised CS.

While the shortcomings in the SLR should be considered to be the main key issue, the EAG identified other areas of concerns, namely regarding the population of interest and the care pathway (Section 2.1), the inclusion of all relevant comparators (Section 2.3), and concerning the ITC (Section 3.4).

4. COST EFFECTIVENESS

4.1 EAG comment on company's review of CE evidence

This Section pertains mainly to the review of cost effectiveness analysis (CEA) studies. However, the search Section (4.1.1) also contains summaries and critiques of other searches related to CE presented in the CS. Therefore, the following Section includes searches for the CEA review, measurement and evaluation of health effects as well as for cost and healthcare resource identification, measurement and valuation.

4.1.1 Searches performed for CE Section

The following paragraphs contain summaries and critiques of all searches related to CE presented in the CS.[1, 5] The CADTH evidence-based checklist for the PRESS, was used to inform this critique.[8, 9] The EAG has presented only the major limitations of each search strategy in the report.

The company provided a single search combining facets designed to retrieve economic evaluations, costs and resource utilisation outcomes, and HRQoL data for patients with endometriosis-associated pain. These searches were performed on 5 December 2022.

A summary of the sources searched is provided in Table 4.1.

Table 4.1: Data sources searched for Appendix G: Published CE studies (as reported in CS)

Resource	Host/Source	Date Ranges	Date searched
Electronic databases			
Embase & MEDLINE	EMBASE.com	Last 5 years	05/12/22
MEDLINE® In-Process	Pubmed.com	Last 5 years	05/12/22
NHS EED	CRD	Inception-2015	05/12/22
HTA Database	CRD	Inception-2015	05/12/22
Conferences			
International Health Economics Association		2020-2022	15/12/22
World Congress on Health Economics, Health Policy, and Healthcare Management		2020-2022	15/12/22
European Health Economics Association		2020-2022	15/12/22
American Society of Health Economists		2020-2022	15/12/22
The Professional Society for Health Economics and Outcomes Research		2020-2022	15/12/22
Additional searches			
Hand searching the reference lists of the included SLRs with objective similar to the current SLR			

Based on the CS and Appendices G, H and I of the CS[1, 5]
CRD = Centre for Reviews and Dissemination; CS = company submission; EED = NHS Economic Evaluation Database; HTA = Health Technology Assessment, NHS = National Health Service; SLR = systematic literature review

EAG comment:

- Appendix G reported on a single set of searches undertaken to retrieve papers containing relevant economic evaluations, costs and resource utilisation outcomes, and HRQoL data for patients with endometriosis-associated pain.[5] Economic evaluation studies and cost and resource use studies were both searched for the last 5 years, whilst utility studies were searched for the last 10 years.
- Searches were undertaken across a good range of resources. However, the EAG noted that searches reported for both MEDLINE and Embase, carried the same limitations as the clinical effectiveness searches (see Section 3.1.1) with regard to the conditions facet and the use of the overly restrictive pain facet, therefore the EAG has concerns that relevant studies may have been missed.
- Appendix G of the CS reported searches of EconLit, HTA database (CRD) and the National Health Service (NHS) Economic Evaluation Database (EED; CRD), however, no search strategies are provided.[5] The EAG queried this at clarification and the company responded that *"owing to time constraints, EconLit was not searched, so its inclusion in Appendix G is an error for which we apologise. We can confirm that CRD was searched; 32 hits were retrieved but none were deemed relevant at screening. The 32 hits were combined for economic evaluation, health care resource use and utility studies"*.[3] As no full search strategy was provided, the EAG is unable to comment on its suitability.

4.1.2 Inclusion/exclusion criteria

The in- and exclusion- predefined criteria used by the company to conduct the SLR for the CE studies are presented in Table 115 of Appendix G (search date December 2022) for the CE studies, in Table 120 of Appendix H (search date December 2022) for HRQoL studies, and in Table 123 of Appendix I for the studies on health care resource use and costs.[5] The EAG considers the in- and exclusion criteria suitable to capture all relevant evidence.

4.1.3 Findings of the CE review

The preferred reporting items for systematic reviews and meta-analyses (PRISMA) flow diagrams for the CE studies can be found in Figure 50 (search date December 2022) of the Appendix G, for the quality of life (QoL) studies in Figure 51 (search date December 2022) of the Appendix H and for the costs and health care resource use in Figure 52 (search date December 2022) of Appendix I.[5] A total of six CE studies (five from the structured databases search in December 2022 and one from the pragmatic search of the grey literature), of which three were linked to one of the other included records leading to the identification of three unique CE studies (including Bohn 2020, as cited in the CS but no details provided).[24, 25] Furthermore, 12 HRQoL studies (all from the December 2022 search) were included and 14 studies for the costs and health care resource use (12 from the structured databases search in December 2022 and two from the pragmatic search of the grey literature), which then led to 11 original studies. As there were no CE studies evaluating relugolix CT treatment identified in the SLR, a de novo CE model was constructed. Furthermore, the SLR did not identify any previous NICE technology appraisals for endometriosis treatment. Of the 12 HRQoL studies, five studies used the European Quality of Life-5 Dimensions (EQ-5D) for measuring HRQoL. Of the 11 studies on costs and resource use, five reported costs only data, four reported only resource use data and the remaining two studies reported both costs and resource use data.[5]

4.1.4 Conclusions of the CE review

Searches were conducted on 5 December 2022 to retrieve published economic models, available economic evidence including economic evaluations, costs, and resource use, as well as relevant utility data for patients with endometriosis-associated pain.[1, 5] The EAG noted that searches reported for both

MEDLINE and Embase, carried the same limitations as the clinical effectiveness searches, with regard to the condition facet and use of the overly restrictive pain facet, therefore the EAG concludes that it is likely that relevant studies may have been missed. Since no CE models to address the impact of relugolix CT treatment were identified by the company, a de novo model was built, which is discussed in the remainder of this Section. Finally, the SLR did not identify any previous NICE technology appraisals for endometriosis treatment.

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

4.2.1 NICE reference case checklist

Table 4.2: NICE reference case checklist

Element of HTA	Reference case	EAG comment on CS
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	As per the reference case
Perspective on costs	NHS and PSS	As per the reference case
Type of economic evaluation	Cost utility analysis with fully incremental analysis	As per the reference case
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	As per the reference case
Synthesis of evidence on health effects	Based on systematic review	As per the reference case
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults	Health effects expressed in QALYs. HRQoL measured using the EQ-5D-5L (mapped to EQ-5D-3L)
Source of data for measurement of HRQoL	Reported directly by patients and/or carers	As per the reference case
Source of preference data for valuation of changes in HRQoL	Representative sample of the UK population	For the health state utilities, quality-of-life was prospectively measured in the SPIRIT 1 and 2 trials using the EQ-5D measure as per the NICE reference case. For disutility values other sources were used with some uncertainty of whether they are representative of the UK population
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the	As per the reference case.

Element of HTA	Reference case	EAG comment on CS
	individuals receiving the health benefit	QALY weighting due to disease severity is not applicable.
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	As per the reference case
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	As per the reference case

CS = company submission; EAG = Evidence Assessment Group; EQ-5D = EuroQoL-5 Dimensions; EQ-5D-3L = EuroQoL-5 Dimensions, 3 levels; EQ-5D-5L = EuroQoL-5 Dimensions, 5 levels; HTA = Health Technology Assessment; HRQoL = health-related quality of life; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; PSS = Personal Social Services; QALY = quality-adjusted life year; QoL = quality of life; UK = United Kingdom

4.2.2 Model structure

The company developed a de novo semi-Markov model in Microsoft Excel® to assess the CE of relugolix CT compared to GnRH agonists for UK adult patients with symptoms of endometriosis who have a history of previous medical or surgical treatment for their endometriosis. The model consists of 13 mutually exclusive health states which are defined based on patients' response to medical therapies and surgical interventions. The company justified their choice for the model type based on previous models that have been used to evaluate interventions in the treatment of endometriosis.[24-32] Furthermore, the company explained that the model structure was based on reviews of the study designs for the Phase 3 SPIRIT trials,[33, 34] prior economic models of treatments for symptomatic endometriosis identified from a targeted literature review, and clinical practice guidelines for endometriosis.[19, 24-32] The company indicated that the model structure was validated by clinical experts during a global advisory board meeting. A schematic representation of the model structure is shown in Figure 4.1.

All patients start in the "Initial treatment" health state (A in Figure 4.1). Initial response to treatment can be selected at either 3 months or 6 months. Six months was deemed to be in line with clinical practice (based on feedback collected from a global advisory board) and it also aligned with the evaluation time for the two co-primary endpoints in the SPIRIT trials (24 Weeks). Therefore, 6 months was selected for the base-case. Patients may then transition to either complete response or non-response. For the relugolix CT arm, these transitions are based on observed rates at the study visit at 12 Weeks in the SPIRIT trials (B in Figure 4.1). At 6 months, treatment response is evaluated for all patients and a decision of subsequent management is taken. Patients with complete response move to or remain in the complete response health state and continue treatment until the end of the model horizon if response is maintained. Only patients in the complete response health state receive active treatment. Patients in the non-response health state stop treatment, and after one cycle (reflecting the time that clinicians need to assess which subsequent treatment is more appropriate), they switch treatment to best supportive care (BSC) (C in Figure 4.1) or undergo surgery (D in Figure 4.1). Best support care consists of NSAIDs, and the surgical options modelled are conservative surgery (laparoscopy) and hysterectomy. For some patients one or both ovaries as part of the hysterectomy (e.g., an oophorectomy) may be removed. The company mentioned that the proportion of patients undergoing hysterectomy and oophorectomy can be changed in the model. In the base-case, a waiting time of 6 months is assumed before surgery (D in

Figure 4.1), but this can be changed up to 1 year. In the waiting time health state, patients are assumed to receive BSC. After patients undergo hysterectomy, they transition to the post-hysterectomy stable health state (E in Figure 4.1), and they are assumed to remain in that health state unless pain recurs. If that happens, patients either opt for BSC and remain in the post-hysterectomy recurrence health state or undergo an additional surgery (which in all cases is an oophorectomy) and move into the post-hysterectomy reoperation health state. If patients opt for conservative surgery instead of hysterectomy (F in Figure 4.1), they transition to the post-conservative surgery response health state where they are supposed to remain until pain recurs. If pain recurs (G in Figure 4.1), patients may either receive BSC (C in Figure 4.1) or undergo additional surgery. In addition, patients may transition to the dead health state from all other modelled health states. Note finally that if at the beginning of the model, treatment response is evaluated at 3 months instead of 6 months, patients would transition from initial treatment to complete response or non-response after the first model cycle and the partial response health state is not used in this case.

The starting age of the patient cohort is 34 years, reflecting the average age at baseline of the population enrolled in the SPIRIT 1 and 2 trials. The model considered a lifetime time horizon, but shorter time-horizons may also be considered (the scenario until menopause – assumed to happen at 50 years, i.e., time horizon of 16 years – is considered relevant by the company and it was originally considered for the base-case). The model has a cycle length of 3 months. Costs and utilities are applied to each health state to calculate total costs and quality-adjusted life years (QALYs) per model cycle. The input values of the model, and their underlying assumptions, are further elaborated in the remainder of Section 4 of the EAG report.

EAG comment: The model description was unclear before and after clarification. There is text in Document B of the CS that does not match the model implementation. The company explained that these correspond to "legacy assumptions" from an original global model that includes certain functionalities that are not applicable to this NICE submission such as additional comparators such as BSC and surgery or the definition of BSC. However, after clarification these this has not been completely corrected. Some assumptions/functionalities mentioned in the CS could not be validated by the EAG. After pain recurs after hysterectomy, patients transit into post-hysterectomy recurrence where they are treated with BSC. Some patients subsequently undergo an oophorectomy and transit into post-hysterectomy reoperation. However, in the model description BSC can only be reached from conservative surgery, not from hysterectomy. In the model sheet "Variable bank" there is one parameter which is supposed to change the proportion of patients treated with surgery who switch to BSC following discontinuation of treatment. This value is set to 0%, but it seems that changing this input does not change results at all. There is another parameter which is supposed to change the proportion of patients treated with conservative surgery who switch to BSC following recurrence of pain 80%. This parameter does have an impact on the results.

The role and definition of BSC is not clear. The CS defines BSC as a treatment option that includes hormonal therapy with or without analgesics, but this is incorrect.[1] After clarification, the company confirmed that BSC consists of NSAIDs only.[3] Moreover, it seems that patients are allowed to receive BSC multiple times after being a non-responder (see position of BSC at C, and then again after G in Figure 4.1). It is thus unclear why patients would be re-treated with BSC when it has already failed. More details about BSC are discussed in Section 4.2.4.

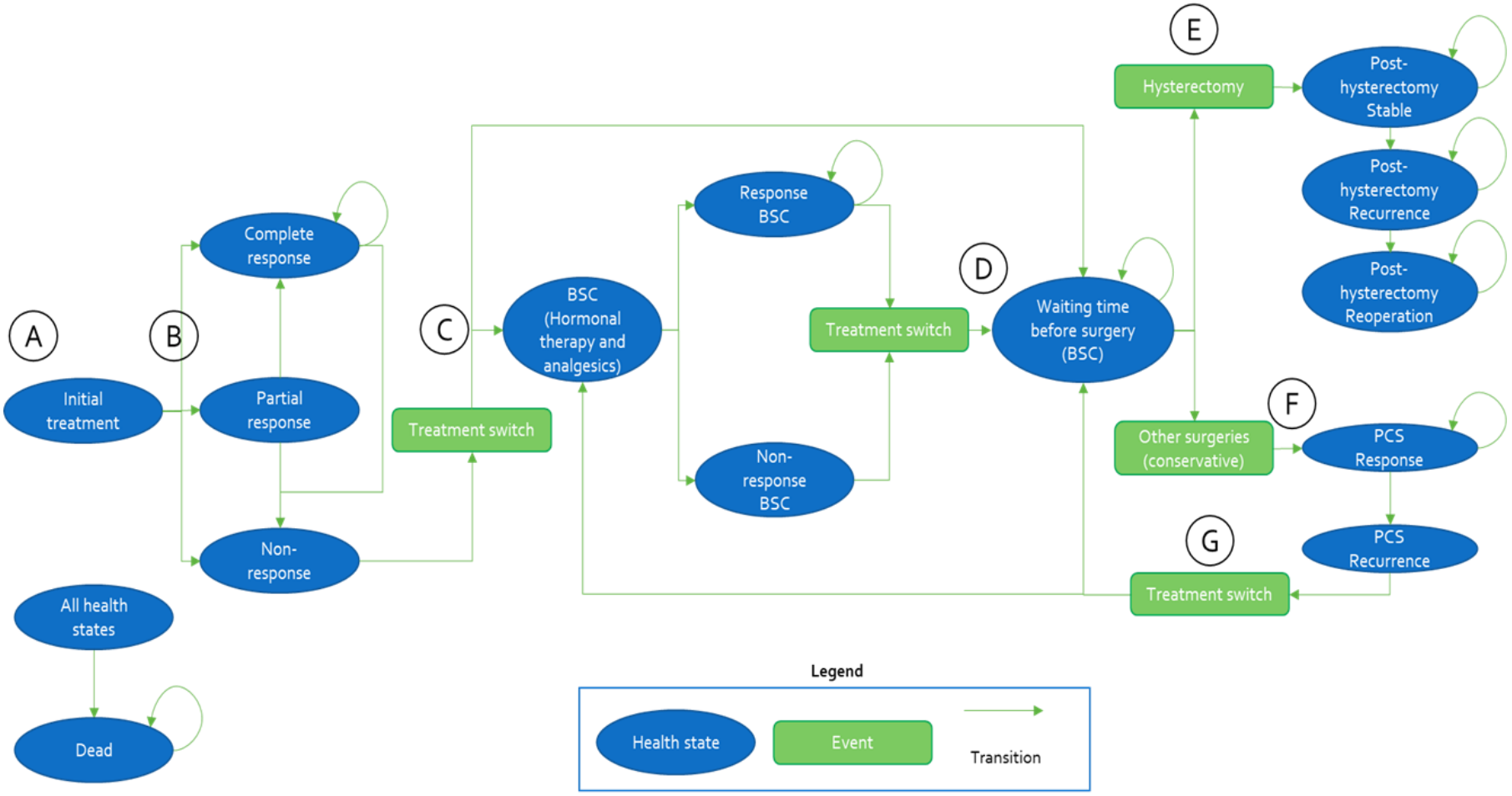
It is also unclear whether the health state "Waiting time before surgery" is needed at all. At point C in Figure 4.1, patients can either move to BSC or "Waiting time before surgery"; however, waiting time before surgery means in practice that patients are treated with BSC. Therefore, it is unclear why this

distinction is needed. The EAG wonders whether the same pathway could have been obtained by removing the “Waiting time before surgery” health state and allow patients to transition from BSC to surgery directly. That would simplify the model structure by having one health state less. Furthermore, it is not clear if there is an upper limit on the number of times patients can "loop" over the model, for example, if there is a limit for the number of surgeries a patient can undergo.

In clarification question B7, the EAG pointed out that potentially relevant health economic outcomes such as fertility, hospital admissions, overall pain, recurrence of endometriosis, or complications of treatment were not included in the economic model, and asked the company to discuss the impact on the CE results.[3] For each outcome, the company indicated the following:

- Hospital admissions: most hospital admissions are related to procedures, and these are captured in the model. The company referred to a recent Australian report where 40,500 endometriosis-related hospitalisations compared with 3,600 endometriosis-related emergency department presentations were reported.[35, 36] The company concluded from this study that the contribution of emergency admissions to the CE model is considered to be minor. The EAG wonders whether these proportions are representative for the UK. If that would be the case, this might imply that approximately 10% of the surgery-related hospitalisations would end up in emergency admissions (which is presumably more costly and involves a higher risk to patients).
- Fertility: a utility decrement due to infertility should only be applied to the proportion of women actively seeking to have a family. The EAG considers it uncertain whether this proportion can be estimated. Since both GnRH agonists and relugolix CT are contraceptive, the company considered that the utility decrement related to infertility would have been captured within the EQ-5D values collected in the trials, given that women participating in the trials were aware of this. The EAG considers that even if this utility decrement would have been captured, it should be emphasised that GnRH agonists are only administered for a maximum of 1 year, whereas in the model women can remain on relugolix CT treatment for 16 years. Therefore, the impact of the utility decrement due to infertility would be much longer for women on relugolix CT compared to those on GnRH agonists. The company also explained that a utility benefit for recovery of fertility following discontinuation of relugolix CT was deemed too uncertain to be parameterised, but it was expected to have little impact on the CE results, given that the difference in time to regain in fertility between the two treatments is likely to be months rather than years. The EAG is uncertain about the company’s expectation given that relugolix CT and GnRH agonists may have a different timeframe. For example, if a woman stops treatment after 1 year or after 10 years (only possible with relugolix CT) it is likely that the time to regain fertility would not be the same. Also, as suggested by the NICE algorithm for diagnosing and managing endometriosis (see Figures 2 and 3 of the CS),[1] fertility concerns for patients seem to guide the clinical pathway to some extent. This does not seem to be captured by the current model structure.
- Recurrence of endometriosis: the company clarified that relugolix CT is not a disease-modifying drug; it relieves the symptoms of endometriosis rather than removing diseased endometrial tissue. Therefore, it is not possible for endometriosis to “recur” after treatment with relugolix CT.
- Overall pain: the company indicated that since the model uses utility values collected directly from the clinical studies, the impact of pain is captured in the model.
- Complication of treatment: the company explained that treatment complications are modelled through AEs, and by including costs and disutilities of surgical interventions.

Figure 4.1: Schematic representation of the model structure*



Based on electronic model submitted after clarification.[37]
* The definition of BSC in the figure is incorrect: it should be analgesics only – without hormonal therapy.
BSC = best supportive care; CS = company submission; PCS = post-conservative surgery

4.2.3 Population

The population included in the economic analyses was defined by the company as adult patients with symptoms of endometriosis who have a history of previous medical or surgical treatment for their endometriosis.[3] The baseline patient characteristics used in the model are summarised in Table 4.3 and were sourced from the SPIRIT 1 and 2 clinical trials.[38]

Table 4.3: Baseline patient characteristics used in the economic model

Patient characteristic	Value in model	Source
Age (years)	33.88	SPIRIT 1 and 2[38]
Weight (kg)	70.4	
Body surface area (m^2)	1.71	
Total cholesterol (mg/dL)	182.36	
High-density lipoprotein (mg/dL)	29.72	
Systolic blood pressure (mmHg)	115.72	
Smoker (%)	17.1	
Diabetes (%)	7.1	

Based on Table 40 of the CS[1]
CS = company submission; dL = decilitre; kg = kilogram; m = meter; mg = milligram; mm; Hg = millimetres of mercury

EAG comment: The main EAG concerns regarding the population in the CS were discussed in Section 2.1 of this report. These concerns are also relevant to CE.

In addition, as explained in Section 4.2.6.3, the company indicated that women who discontinued treatment due to pregnancy (or wish to get pregnant) were excluded from the estimation of the discontinuation rates since treatment with BSC or surgery are deemed as not feasible options for these patients. Based on this, the EAG is unclear whether these patients have been properly included in the model.

Finally, no measure of uncertainty was provided in Table 4.3. This is needed for the probabilistic sensitivity analysis (PSA) and should be based on the trial data. As discussed in Section 5.2.1, the EAG considers that assuming 10% variation for all parameters is incorrect.

4.2.4 Interventions and comparators

The intervention considered in the economic analysis was relugolix CT, defined as relugolix 40 mg in combination with oestradiol 1 mg and norethisterone acetate (NETA) 0.5 mg, administered orally once daily, until menopause.

The comparator considered in the economic analysis was GnRH agonists. The company indicated that defining GnRH agonists as per established clinical practice in the UK was challenging, given that there are multiple GnRH agonists that are currently licensed for the treatment of endometriosis, i.e., leuprorelin acetate, goserelin, triptorelin, nafarelin and buserelin. Leuprorelin acetate, goserelin and triptorelin are administered as subcutaneous (SC) injection, either in short-acting (monthly) or long-acting (3-monthly) formulation. Nafarelin and buserelin are administered as daily intranasal treatments. The company conducted an email survey to UK-based healthcare professionals, who gave guidance as to which GnRH agonists are most commonly used for the treatment of endometriosis associated pain.

Their answers were heterogenous, reaching no consensus. However, all the consulted experts stated that all patients opt for the SC formulations. Therefore, no patients in the model are assumed to receive nafarelin or buserelin. Furthermore, the company assumed that the efficacy of different GnRH agonists was equal, and in that case, the choice of specific GnRH agonists would only impact total costs. The company assumed then a 50/50 split amongst the cheapest short-acting GnRH agonist and the cheapest long-acting GnRH agonist. The company considered this appropriate on the basis that the more costly GnRH agonists would be dominated by the less costly GnRH agonists in an incremental analysis.

Based on the feedback of clinical experts during a global advisory board, treatment with GnRH agonists should be complemented with add-back therapy which is typically initiated after three months on GnRH agonists (which is assumed in the base-case). GnRH agonists in combination with add-back therapy are assumed to be administered for a maximum of 12 months. According to the company, this is in line with current recommendations for the treatment with GnRH agonists.[39, 40] For add-back therapy, the company included in the model two treatments only (tibolone and raloxifene) and assumed a 50/50 share.

EAG comment: The main EAG concern is related to the exclusion of relevant comparators as discussed in Section 2.3 of this report. The poor quality of the searches resulted in the omission of relevant comparators such as oral peptide GnRH antagonists, off licence nasal or parenteral GnRH analogues and addback HRT, which could all be used in the UK clinical practice as per clinical expert opinion. Therefore, since the economic analyses presented by the company in this submission are lacking relevant comparators, their results are likely to be invalid.

In addition, the EAG is concerned about the definition and role of BSC in the model:

- BSC is part of the modelled treatment pathway (after treatment discontinuation with relugolix CT or GnRH agonists). As previously mentioned, a clear definition of them, specially BSC, is missing from the CS. The company defined BSC as hormonal therapy with or without analgesics in the CS.[1] However, this was incorrect according to clarification question B23,[3] where it was clarified that BSC should have been defined as symptomatic treatment for pain management (NSAIDs, i.e., analgesics only). The company explicitly mentioned that other hormonal treatments should be disregarded.
- For some input parameters (e.g., BSC response and AE rates, and BSC health state utilities), the company assumed that the modelled BSC (NSAIDs*, after failing treatment with GnRH agonists or relugolix CT*) is equivalent to placebo in the SPIRIT trials. In response to question B10 a), the company explained that BSC in England and Wales for the patient population in this appraisal is symptomatic treatment for pain management, such as analgesics, which is the same as the definition of BSC in the SPIRIT trials.[3] However, the EAG is uncertain about the validity of this assumption. Following Giudice et al. 2022,[38] it seems that placebo is not defined as NSAIDs since the use of analgesics (opioids and non-opioids) in the trials is explicitly mentioned separately and in fact on of the aims of the trials was to reduce the use of them. However, it is also unclear what the impact on the CE results might be.

4.2.5 Perspective, time horizon and discounting

The economic analysis is performed from the NHS and Personal Social Services (PSS) perspective. The model had initially a time horizon of 16 years, which was not in line with the NICE reference case, since that is not considered as a lifetime horizon, given that the average age of patients at the start of treatment is 34 years. In response to clarification question B6, the company modified the CE model by introducing a post-menopause health state which allowed the model to be run for a lifetime time horizon.

The model cycle length is three months. Costs and QALYs were discounted at 3.5% as per the NICE reference case.

4.2.6 Treatment effectiveness and extrapolation

The main sources of evidence on relugolix CT treatment effectiveness are the Phase 3 SPIRIT 1 and 2 trials.[38] Details about the clinical effectiveness reported in these trials can be found in Section 3.3 of this report. Clinical parameters used in economic model include event probabilities, such as treatment discontinuation, choice of surgical interventions, re-surgeries and treatment schedules. These are discussed separately in the remaining of this Section.

4.2.6.1 Response to initial treatment with relugolix CT and GnRH agonists

As mentioned above, the model assumes complete response, partial response, and non-response to treatment. Two possible definitions of complete response to treatment have been included in the model:

- Change from baseline response: NRS score reduction from baseline of 2.8 for dysmenorrhea and 2.1 for NMPP, and no increase of analgesic use. This definition of complete response was informed by the co-primary endpoints of the SPIRIT 1 and 2 trials and was selected by the company for their base-case. The values 2.8 for dysmenorrhea and 2.1 for NMPP were established from an anchor-based approach using the Patient Global Assessment (PGA) measure as an anchor to correlate with changes in NRS. Patient Global Assessment was collected at baseline and thereafter every 4 Weeks at study visits during the SPIRIT 1 and 2 trials.[38]
- Threshold response: achieving or maintaining a threshold below 4 in NRS scale (mild pain) for both NMPP and dysmenorrhea and no increase of analgesic use. This definition of complete response was suggested to the company by clinical experts during an advisory board. Clinicians indicated that since treatments aim to minimise the level of pain in patients, measuring response by achieving a certain threshold value may be more feasible in practice.

Partial response is defined in the same way as complete response, but response is evaluated at 3 months. At 6 months, patients can no longer be classified as partial responders; there are either complete responders or non-responders.

Table 4.4 shows the response probabilities for both relugolix CT and GnRH agonists. Note that response rates for GnRH agonists are assumed to be equal to the response rates of relugolix CT following the indirect treatment comparison presented in Section 3.4 of this report.

Table 4.4: Probability of achieving response to treatment

Response type	Relugolix CT and GnRH agonists	Source
Change from baseline response		
CR: 3 months	40.4%	[33, 34]
CR: 6 months	58.9%	
PR: 3 months	30.6%	
Threshold response		
CR: 3 months	47.4%	[33, 34]
CR: 6 months	63.4%	
PR: 3 months	25.8%	
Based on Table 42 and Table 43 of the CS.[1]		

Response type	Relugolix CT and GnRH agonists	Source
CR = complete response; CS = company submission; CT = in combination with oestradiol and norethisterone acetate; GnRH = gonadotropin releasing hormone; PR = partial response		

EAG comment:

The main concerns of the EAG regarding response to treatment are the following:

- The model is insensitive to the choices between the two definitions of complete response, and the choice between complete and partial response. This is expected given that complete response is evaluated at 6 months, partial response at 3 months, and the comparator (GnRH agonists) is only applied for at most 1 year. Because the time horizon of the model is lifetime, its results are mainly driven by the difference in treatment effectiveness between relugolix CT and the subsequent treatment options after discontinuation from GnRH agonists (i.e., BSC and surgery – conservative or hysterectomy/oophorectomy), rather than between relugolix CT and GnRH agonists. Based on this, the EAG wonders whether the model could have been simplified by excluding one of the definitions of complete response and partial response.
- Equal effectiveness between relugolix CT and GnRH agonists was assumed based on no statistical significance. This was changed after clarification (clarification question B11) and now the results of the ITC, discussed in Section 3.4 of this report, has been included in the economic model.[3] In particular, the ITC of the network for OPP, with random effects and weakly informative priors, was used to derive the response rates of GnRH agonists using an OR of 1.1 with (0.032, 41) as 95% CrI.
- As mentioned in Section 2.4, The following outcomes were not collected in the SPIRIT trials: hospital admission, fertility, recurrence of endometriosis, or complications of treatment. These are therefore not included in the CEAs. It is unknown whether and to what extent these outcomes might change the CE results should they be considered in the analyses.
- The EAG considers that the model produces counterintuitive results when decreasing relugolix CT response rates. In the extreme scenario where the proportion of patients achieving complete response in the relugolix CT arm is assumed to be only 1%, relugolix CT still produces 0.011 incremental QALYs compared to GnRH agonists. However, because the 99% of patients discontinue treatment with relugolix CT (due to no response), this scenario results in relugolix CT saving costs compared to GnRH agonists. Thus, when its response rate is 1%, relugolix CT is dominant compared to GnRH agonists.

4.2.6.2 Subsequent treatment distributions

Following discontinuation from initial treatment with relugolix CT or GnRH agonists (C in Figure 4.1), patients receive either BSC or surgery; surgery can also be conservative (e.g., laparoscopy) or hysterectomy. The proportions of patients switching to these subsequent treatments are summarised in Table 4.5. For relugolix CT, these proportions were sourced from patient-level data collected in the SPIRIT extension study.[41] A total of 26 patients discontinued treatment due to either lack of efficacy (four cases) or AEs (22 cases). Of these 26 patients, five (16.7%) initiated BSC, and 21 (83.3%) underwent surgery. The distribution between conservative surgery and hysterectomy was sourced from Soliman et al. 2016,[42] a real-world evidence study in patients receiving treatment with leuprolide acetate for endometriosis. In the absence of data to inform this distribution for the GnRH agonists arm of the model, the company assumed the same distribution as in the relugolix CT arm. The company indicated that these assumptions were validated by UK clinical experts.

Table 4.5: Distribution of subsequent strategy to manage endometriosis following initial treatment discontinuation from relugolix CT or GnRH agonists

Treatment switch (from column to row)	Relugolix CT* and GnRH agonists**
BSC	16.7%
Conservative surgery	38.4%
Hysterectomy	45.0%

Based on Table 46 of the CS.[1]
* Sourced from Soliman et al. 2016 and trial data.[42, 43]
** Assumed to be equal to relugolix CT.
BSC = best supportive care; CS = company submission; CT = in combination with oestradiol and norethisterone acetate; GnRH = gonadotropin releasing hormone

The treatment effectiveness of BSC and surgery after discontinuation from relugolix CT (or GnRH agonists) was not discussed in the CS. However, as mentioned the previous Section of this report, since the modelled time horizon is lifetime and GnRH agonists are prescribed for at most 1 year, the CE results are mainly driven by the difference in treatment effectiveness between relugolix CT and BSC or surgery, rather than between relugolix CT and GnRH agonists. Therefore, these are expected to be relevant input parameters for the economic model. In the economic model sheet "Variable bank" there is an input parameter which accounts for the probability of response of BSC (after intervention/comparator) and equals to 18%. The EAG was unable to validate the source for this input parameter. Similar parameters should be expected for response to surgeries. However, the EAG was unable to locate these parameters in the model.

Patients who switched to BSC following discontinuation from initial treatment with relugolix CT or GnRH agonists are assumed to receive surgery (conservative or hysterectomy) when they do not respond to BSC. The distribution between conservative surgery and hysterectomy was also sourced from Soliman et al. 2016,[42] and it was 46% conservative surgery and 54% hysterectomy. Again, the same distribution as in the relugolix CT arm was assumed for GnRH agonists. However, changing these parameters does not seem to have any impact on the relugolix CT or GnRH agonists arms. Therefore, the EAG is unclear whether this is an error or another "legacy" from a previous model version. Finally note that response to surgery after BSC, relugolix CT or GnRH agonists should also be estimated from the correct population.

After hysterectomy (E in Figure 4.1) patients may also undergo a reoperation. The probability of reoperation is assumed to be 10% in the base-case, following input from a UK-based clinical expert. Furthermore, the company assumed an increased mortality risk following surgery. The 3-month (e.g., one model cycle) probability of dying after conservative surgery is estimated as 0.003% and as 0.038% after hysterectomy or oophorectomy.[44, 45]

As described in Figure 4.1, the modelled treatment pathway includes four types of interventions: relugolix CT, GnRH agonists, BSC and surgery (conservative surgery or hysterectomy). All these interventions are used in combination with analgesics following the proportions shown in Table 4.6. In the absence of data, the company assumed that the proportions of patients using analgesics in combination with GnRH agonists and surgery was the same as that in the relugolix CT arm of the SPIRIT trials.

The company indicated that add-back therapy is prescribed in addition to GnRH agonists for long-term use (not in situations where GnRH agonists are short-term used, e.g., prior to surgery) and after oophorectomy. The company assumed that patients treated with GnRH agonists initiate add-back

therapy after 3 months, and immediately after oophorectomy. This assumption was confirmed by clinical experts during an advisory board.

Table 4.6: Proportion of patients using analgesics

Intervention	Before response	After response	Source
Change from baseline response			
Relugolix CT	90.0%	28.9%	33, 34
GnRH agonists*	90.0%	28.9%	
BSC	72.0%	49.1%	
Surgery*	90.0%	28.9%	
Threshold response			
Relugolix CT	90.0%	28.3%	33, 34
GnRH agonists*	90.0%	28.3%	
BSC	72.0%	50.0%	
Surgery*	90.0%	28.3%	

Based on Table 44 and Table 45 of the CS.[1]
* Assumed to be equal to relugolix CT.
BSC = best supportive care; CS = company submission; CT = in combination with oestradiol and norethisterone acetate; GnRH = gonadotropin releasing hormone

EAG comment: The main concerns of the EAG regarding the distribution of subsequent treatments are the following:

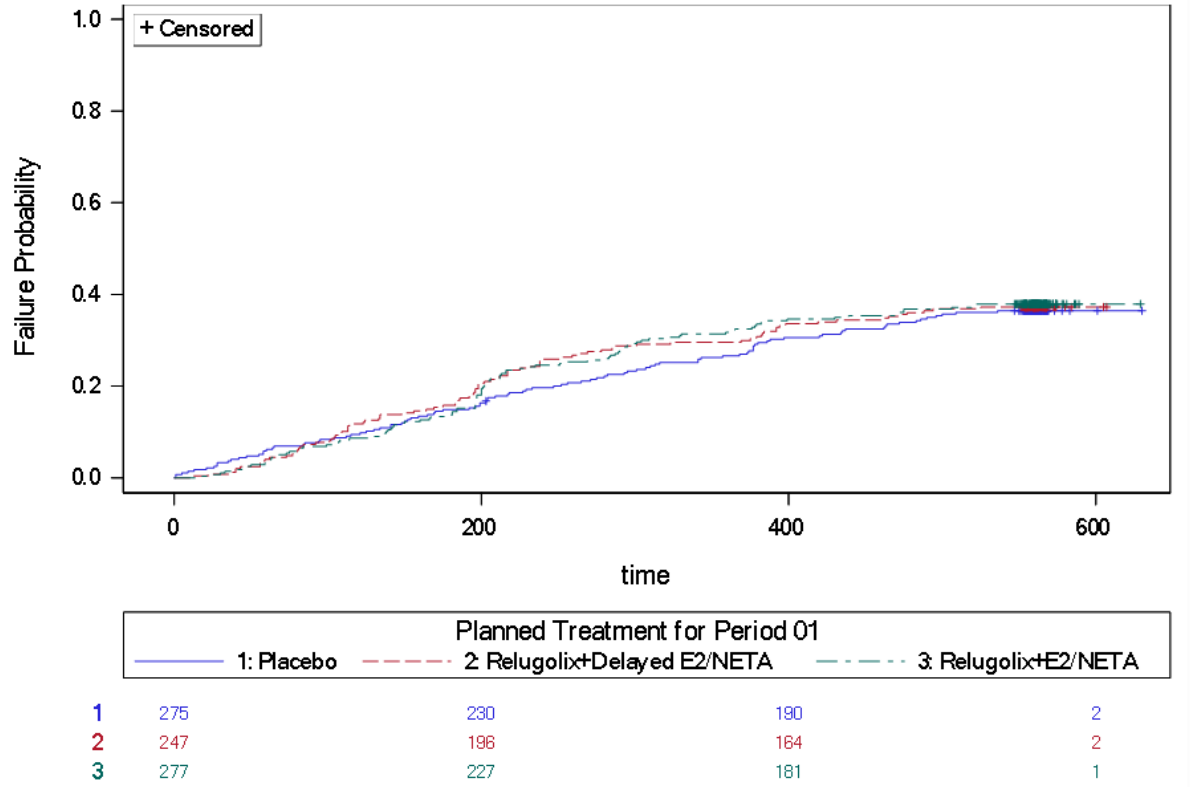
- The proportion of patients switching to BSC or surgery after discontinuation from relugolix CT were sourced from the SPIRIT extension study.[41] It is unclear whether the same proportions would be expected after GnRH agonists discontinuation. It is also unclear whether data on treatment discontinuation were collected during the SPIRIT 1 and 2 trials and why (not).
- The distribution between conservative surgery and hysterectomy after discontinuation with relugolix CT was sourced from Soliman et al. 2016.[42] This is a real-world evidence study in patients receiving treatment with leuprolide acetate for endometriosis, but it is not discussed whether the population in this study can be assumed to be equivalent to the DP population.
- Treatment effectiveness after discontinuation should have been explained in more detail in the CS. Response to BSC and to surgery after treatment discontinuation is expected to have more impact on model results than response to GnRH agonists given how short a time is spent on GnRH agonist treatment. How these have been implemented in the model remains unclear.
- The company explained that relugolix CT, GnRH agonists, BSC and surgery (conservative surgery or hysterectomy) are used in combination with *analgesics*. As discussed in Section 4.2.4, the EAG is concerned about the definition of BSC used by the company in the economic model. In response to clarification question B23,[3] the company indicated that BSC should have been defined as symptomatic treatment for pain management (i.e., NSAIDs). Since NSAIDs are analgesics, it is unclear 1) what type of analgesics are used in combination with the four types of interventions included in the model, 2) how BSC should be interpreted and 3) whether BSC can be considered equivalent to placebo in the SPIRIT trials.
- Analgesics would be expected to impact both costs and effectiveness results (assuming the use of analgesics would diminish pain to some extent, but they might also be associated with AEs).

However, in the model, they are only linked to the cost calculations. Because the use of analgesics is reduced after response, and response rates are higher for relugolix CT, including the effect of analgesics would most likely benefit relugolix CT although it is uncertain to what extent.

4.2.6.3 Treatment discontinuation

Treatment discontinuation rates (from relugolix CT, GnRH agonists and BSC) are applied at each model cycle (i.e., every 3 months). Discontinuation is assumed after loss of response or intolerability. Discontinuation rates for relugolix CT and BSC were estimated from a post-hoc analysis of discontinuation data obtained from the SPIRIT OLE study, as can be seen in Figure 4.2.[41] Discontinuation rates observed in the placebo arm of the SPIRIT OLE study were assumed for BSC.

Figure 4.2: Time to treatment discontinuation in the SPIRIT OLE Study



Based on Figure 39 in the CS.[1]
Note: patients who did not discontinue were censored at the last date of contact.
CS = company submission

Discontinuation rates over time were based on the observed hazard rates for time to treatment discontinuation at 3-month intervals and included discontinuation due to any reason. The estimated discontinuation rates were adjusted for events such as protocol deviation, under the assumption that these would not lead to discontinuation in clinical practice. Patients who discontinued treatment due to pregnancy or wish to get pregnant were excluded from the estimation of the discontinuation rates since treatment discontinuation in the model leads to either BSC or surgery, which are deemed as not feasible options for these patients. The discontinuation rates over time for relugolix CT, GnRH agonists and BSC are shown in Table 4.7. The same rates as in the relugolix CT arm were assumed for GnRH agonists. Note finally that in the company base-case it is assumed that the treatment duration with GnRH agonists is capped at 1 year; after that, patients are either treated with BSC or surgery.

Table 4.7: Discontinuation rates over time (months since treatment response)

	3 months	6 months	9 months	12 months	15+ months
Relugolix CT (and GnRH agonists)*	0.017	0.017	0.033	0.021	0.012
BSC**	0.017	0.017	0.033	0.021	0.012

Based on Table 48 of the CS and electronic model submitted after clarification.[1,37]
* Rates sourced from SPIRIT trials for relugolix CT, and the same rates were assumed for GnRH agonists.[33, 34]
** BSC discontinuation rates in the model are equal to those for relugolix CT. However, based on Figure 4.2, these are expected to be different.
BSC = best supportive care; CS = company submission; CT = in combination with oestradiol and norethisterone acetate

EAG comment:

The main concerns of the EAG regarding treatment discontinuation are the following:

- The same rates as in the relugolix CT arm were assumed for GnRH agonists. There is uncertainty regarding this assumption but the impact on the model results is minimal.
- The discontinuation rates over time for relugolix CT and BSC shown in Table 4.7 were based on observed hazard rates. However, it is unclear why exactly the same rates were used (after checking the model the EAG can confirm that these are identical) when based on Figure 4.2, discontinuation seems lower for BSC/placebo. Also note, again, that the placebo population in the trial is likely to be different from the BSC after relugolix CT, GnRH agonists or surgery in the model.
- The EAG does not understand why BSC, or surgery are deemed as not feasible options for patients who discontinued treatment due to pregnancy or wish to get pregnant, as the company stated. Best supportive care was defined by the company as analgesics only and surgery can also be conservative. If these options are infeasible, the EAG wonders how these women are supposed to be treated. It is also unclear why these patients were not modelled separately. The company indicated that these patients were excluded from the estimation of the discontinuation rates since treatment discontinuation, but it seems that they are completely excluded from the model. The impact of including these patients in the analyses is unknown, but it might be relevant if the proportion of patients seeking pregnancy is large. This could be informed by clinical experts in the absence of other source of data. Also, the EAG wonders whether this means that the decision problem should be redefined by excluding those women who got or wished to get pregnant.

4.2.6.4 Pain recurrence following surgery

As shown in Figure 4.1, patients who undergo either type of surgery are assumed to be at risk of recurrence of endometriosis related pain. The 3-month probabilities of pain recurrence are shown in Table 4.8, and they were estimated using data on healthcare claims from the Truven Health MarketScan Commercial Claims and Encounters Database between 2004-2013, following a study of post-surgery treatment outcomes in patients undergoing hysterectomy or laparoscopy for endometriosis.[46] Based on the same study, the company also indicated that in case of pain recurrence following conservative surgery, 80% of patients would be managed with BSC, 11.1% of patients would undergo hysterectomy and 8.9% of patients with another conservative surgery. This was validated by clinical experts.

Table 4.8: Rate of pain recurrence over time by type of surgery*

Type of surgery	3 months	6 months	9 months	12 months	15 months	18 months	21 months	24+ months
Conservative	0.322	0.073	0.079	0.086	0.094	0.104	0.116	0.131
Hysterectomy	0.021	0.002	0.002	0.002	0.002	0.002	0.002	0.002

Based on Table 49 of the CS[1]
* Rates sourced from Soliman et al. 2017[46]
CS = company submission

EAG comment: The main concerns of the EAG relate to the source used by the company to estimate the probabilities of pain recurrence after surgery (data on healthcare claims from the Truven Health MarketScan Commercial Claims and Encounters Database between 2004-2013). It is unclear (because it is not explained in the CS, and it is unexpected) why these rates were not derived from any of the SPIRIT trials. The source seems relatively old, and it is unknown whether the population and BSC in this study are the same as those used in the economic analyses (as in the SPIRIT trials). Check the impact on the results.

4.2.6.5 Treatment effect waning

Regarding relugolix CT treatment effect waning, the CS only mentions that there is a lack of data that would support a treatment waning effect for either relugolix CT or GnRH agonists (Table 41 in CS).[1]

EAG comment: The main concerns of the EAG relate to the lack of explorative analyses or discussion around treatment effect waning. Testing this assumption could be relevant in principle since relugolix CT is taken for many years. In the base-case analysis, patients on relugolix CT will continue treatment until response, discontinuation or until the age of menopause. That also indicates that response to relugolix CT treatment is assumed to be constant over time. The EAG asked the company to comment on the validity of this assumption and to present scenario analyses exploring the effect of long-term treatment effect waning of relugolix CT treatment. In response to clarification question B14, the company indicated that the efficacy of relugolix CT was demonstrated in the SPIRIT OLE study. At Week 52, 84.8% (95% confidence interval (CI): 80.06, 88.85) of patients met the dysmenorrhea responder definition and 73.6% (95% CI: 68.04, 78.74) of patients met the NMPP responder definition. According to the company, these rates of response were sustained through Week 104/EOT: 84.8% (80.06, 88.85) for dysmenorrhea and 75.8% (70.33, 80.74) for NMPP. The EAG is uncertain about the interpretation of "sustained response". However, the company did not present any scenario analyses exploring the effect of long-term treatment effect waning of relugolix CT treatment. This could have been done by assuming no treatment effect after Week 52 for example. The company concluded that the waning of relugolix CT treatment effect is captured through the discontinuation rates applied in the model, since at discontinuation, patients move from complete response to non-response. The EAG is uncertain about this too since as shown in Table 4.7, the company assumed a constant discontinuation rate after 15 months, which would imply a constant treatment effect after Week 60 approximately. The EAG would suggest the company that, in the absence of long-term follow-up data, assuming a 15-year sustained effect could be considered a strong assumption and its impact on the model results should be explored.

4.2.7 Adverse events (AEs) and complications

4.2.7.1 Treatment-related AEs

The CS states that AEs of Grade ≥3 and at a frequency of ≥1% were considered in the economic analysis. In the SPIRIT 1 and 2 trials, the only AE meeting the previous criteria was headache. However, other AEs such as hot flush, decreased libido, depression, increased blood pressure, and hair loss were highlighted by clinical experts and therefore included in the analyses. The analysis also distinguishes between acute (total probability at treatment start only) and constant risks (per 3-month cycle while on treatment). The acute and constant AE probabilities included in the economic analysis can be seen in Tables 4.9 and 4.10, respectively.

Table 4.9: Total (acute) AE probability (relugolix CT, BSC and GnRH agonists)*

AE	Relugolix CT	BSC	GnRH agonists (monotherapy)	GnRH agonists + add-back therapy	Source
Hot flush	0.24%	0.24%	0.74%	0.47%	13, 33, 34
Headache	1.67%	0.48%	1.71%	1.71%	
Depression	0.00%	0.24%	1.25%	0.40%	
Increased blood pressure	0.24%	0.00%	0.24%	0.24%	33, 34

Based on Table 52 of the CS.[1]
* Decreased libido and hair loss not included in table due to 0% probability.
AE = adverse event; BSC = best supportive care; CS = company submission; CT = in combination with oestradiol and norethisterone acetate; GnRH = gonadotropin releasing hormone; RR = risk ratio

Table 4.10: Three-month (cycle) AE probability (relugolix CT, BSC and GnRH agonists)*

AE	Relugolix CT	BSC	GnRH agonists (monotherapy)	GnRH agonists + add-back therapy	Source
Hot flush	0.11%	0.11%	0.34%	0.22%	13, 33, 34
Headache	0.78%	0.22%	0.79%	0.79%	
Depression	0.00%	0.11%	0.58%	0.18%	
Increased blood pressure	0.11%	0.00%	0.11%	0.11%	33, 34

Based on Table 53 of the CS.[1]
* Decreased libido and hair loss not included in table due to 0% probability.
AE = adverse event; BSC = best supportive care; CS = company submission; CT = in combination with oestradiol and norethisterone acetate; GnRH = gonadotropin releasing hormone; RR = risk ratio

Adverse event profiles of relugolix CT and BSC were sourced from the relugolix CT and placebo arms in the SPIRIT trials, respectively.[33, 34] Then, it is assumed that the AE profiles for GnRH agonists (monotherapy) and GnRH agonists + add-back therapy are different. This is because add-back therapy may not be given at GnRH agonist treatment initiation. In the base-case, the company assumed 100% of patients on GnRH agonists would receive add-back therapy after 3 months on treatment with GnRH agonists (assumption based on feedback collected during the global advisory board). The AE profiles for GnRH agonists (with and without add-back therapy) were derived by applying risk ratios (RRs) to the AE probabilities for BSC. These RR were derived from a Cochrane review on GnRH analogues for the treatment of endometriosis and can be seen in Table 4.11.[13] A Bucher approach was

used to estimate the RR for BSC versus GnRH agonists + add-back therapy,[47] since that was not available in the Cochrane review. Since the probability for decreased libido, hypertension and hair loss was 0% for BSC, the same probabilities observed for relugolix CT were assumed for GnRH agonists (with and without add-back therapy). Finally, the AE risks are weighted according to the proportion of patients receiving add-back therapy.

Table 4.11: Overview of relative risks for AEs linked to treatment with GnRH agonists*

Adverse event	RR (BSC versus GnRH agonists)**	RR (GnRH agonists versus GnRH agonists + add-back therapy)
Hot flush	3.08	1.59
Headache	3.55	1.00†
Depression+	5.21‡	3.13

Based on Table 51 of the CS.[1]
* RRs account for all severity grades of AEs but are assumed to apply to grade 3+ events only.
** RRs for BSC versus GnRH agonists were based on the comparison GnRH agonists versus placebo in line with the SPIRIT 1 and 2 trials, except the RR for depression.
+ Since RRs for depression were not available in the Cochrane analysis, emotional changes were used as a proxy.
† RR for headache reported was not statistically significant. An RR = 1 (no difference) was applied.
‡ Depression data not available for placebo. RR of GnRH agonists versus oral or injectable progestogens was used.
AE = adverse event; BSC = best supportive care; CS = company submission; GnRH = gonadotropin releasing hormone; RR = risk ratio

EAG comment: The EAG noticed a discrepancy between the approach mentioned in the CS to be used to incorporate the impact of AEs in the company base-case and the approach implemented in the electronic model. Specifically, on page 161 of the CS, the company noted that "*for the base case analysis all AEs are assumed to be constant and to persist while on treatment*".[1] However, in the electronic model the impact of AEs was included using the probability of each event only at the start of the treatment (acute event option).

4.2.7.2 Cardiovascular events

The company explained that an increase in total cholesterol levels and a decrease in high-density cholesterol levels was observed in the relugolix CT arm in the SPIRIT 1 and 2 trials. Both changes are associated with an increased risk of cardiovascular events as estimated by the Framingham Heart Study risk function.[29, 48] In the absence of data, patients on GnRH agonists were also assumed to experience the same changes in cholesterol levels. However, for GnRH agonists this risk should be applicable for at most 1 year (while they are on treatment). More details are provided in Section B.3.3 of the CS.[1]

General population mortality is also included in the model. Patients undergoing surgery are assumed to face an additional risk of death, which is applied to the subsequent cycle following surgery. However, no excess mortality due to cardiovascular events was included in the model. This was based on the low number of cardiovascular events (0.003 in the relugolix CT arm) observed in the trial. The impact of omitting this excess mortality due to cardiovascular events is expected to be minor.

The company assumed that the increased risk of cardiovascular events is completely associated with relugolix CT treatment and, therefore disappear following menopause, or when patients stop treatment.

While the company acknowledged that a residual risk of cardiovascular events may remain even after menopause, the company expected this to have little impact on the model results.

4.2.7.3 Change in bone mineral density and risk of fractures

The company explained that patients on active treatment might be at a higher risk of experiencing major osteoporotic fractures. However, in response to clarification question B18,[3] the company indicated that since no change in bone mineral density (BMD) was observed in the relugolix CT arm of the SPIRIT trials, no excessive fracture risk was in the end included in the model. Also, patients on GnRH agonists were assumed to maintain their BMD levels due to the use of add-back therapy. Therefore, the Section about changes in BMD and increased risk of fractures is not further summarised or commented on by the EAG in this report. We refer to Section B.3.3 of the CS for details.[1]

4.2.7.4 Complications related to surgery

Complications related to surgery identified by experts during an advisory board were also included in the analysis. These included the risk of complications for urinary tract infection, fistula and urinary retention/complication. These risks were derived from a prospective Finnish study on complications following 5,279 hysterectomies[49] and they are summarised in Table 4.12.[49] In the base-case analysis, complications related to surgery were assumed to be acute and then to persist for a period of 3 months (one model cycle), in line with the Finnish study.

Table 4.12: Three-month (cycle) probability of surgery-related complications

Complication	Risk type	Conservative surgery	Hysterectomy	Oophorectomy	Source
Urinary tract infection	Acute	0.00%	1.42%	1.42%	49, 50
Fistula	Acute	0.00%	0.04%	0.04%	49, 51
Urinary retention/complication	Acute	0.00%	0.99%	0.99%	49, 50
Impact of surgery on other organs (e.g., bowel problems)	Acute	3.00%	3.00%	3.00%	Expert opinion

Based on Table 58 of the CS.[1]
CS = company submission

4.2.8 Health-related quality of life (HRQoL)

In line with NICE's preferences, EQ-5D data was used to inform utilities for the health states. The EQ-5D-5L data collected from the SPIRIT 1 and 2 trials (shown in Table 18 of the CS),[1] were mapped to 3rd line using the age- and sex-specific NICE Decision Support Unit (DSU) mapping tool.[52, 53] Disutilities related to surgeries, surgical complications, and due to the incidence of AEs were extracted from the literature.

In terms of HRQoL, the health states were split in response to initial treatment (baseline utility score), complete response, partial response, and non-response to treatment. Utility values were estimated contingent on treatment response using three different ordinary least square (OLS) regression models which are described below in this Section. The regression models were fit to patient-level utility values at baseline and at Week 24 from the pooled modified intention-to-treat (mITT) population of the SPIRIT 1 and 2 trials with treatment arms combined (i.e., all patients). Table 4.13 presents the mean

EQ-5D-5L utility values as estimated at baseline and at Week 24 from the mITT population of the SPIRIT 1 and 2 trials.

Table 4.13: EQ-5D-3L utility values at baseline and Week 24 for the pooled mITT population

Timepoint	Number of subjects	Mean (95% CI)	Standard deviation
Baseline	821	0.58 (0.57, 0.60)	0.24
Week 24	684	0.80 (0.78, 0.81)	0.20

Based on Table 59 of the CS.[1]
CI = confidence interval; CS = company submission; EQ-5D-3L = European Quality of Life-5 dimensions 3 levels; mITT = modified intention-to-treat

The first of the three OLS models included an intercept term, a variable for baseline EQ-5D-5L utility value and an indicator variable for response status at Week 24 which was coded as "1" for responder and "0" for non-responder. The company noted that baseline mean NRS score, and age were not included as covariates in any of the model predictions to avoid multicollinearity issues. The second OLS model was defined similarly to the first, but also included a binary variable defining treatment arm which was coded as coded "1" for the relugolix CT arm and "0" for the placebo arm. The third OLS model incorporated an interaction term between the variable defining treatment arm and response status in the second OLS model specification.

The three OLS models were used to estimate utility scores using both definitions of treatment response as these are described in Section B.3.4 of the CS.[1] In the base-case analysis, the company selected to use predictions from the most parsimonious OLS model (first OLS model) grounded on the fact the variables defining treatment arm and interaction term for treatment arm and response status were found to be insignificant. Table 4.14 summarises the regression outputs from the first OLS model when using the two different definitions of treatment response. It was also noted that only assessments with non-missing values for all five dimensions of the EQ-5D-5L questionnaire at both baseline and Week 24 were included in the analysis.

Table 4.14: OLS regression output of mapped EQ-5D-3L utility values at Week 24

Covariate	Estimate	SE	95% CI	z-statistic	P-value
Treatment response defined as change from baseline response					
Intercept	0.5845	0.02547	[0.5346 - 0.6344]	22.95	<.0001
Baseline utility	0.2292	0.03374	[0.1631 - 0.2953]	6.79	<.0001
Response*	0.1650	0.01299	[0.1396 - 0.1905]	12.71	<.0001
Treatment response defined as threshold response					
Intercept	0.5928	0.02422	[0.5454 - 0.6403]	24.48	<.0001
Baseline utility	0.2002	0.03272	[0.1360 - 0.2643]	6.12	<.0001
Response*	0.1714	0.01304	[0.1459 - 0.1970]	13.15	<.0001

Based on Table 60 and Table 61 of the CS.[1]
* Non-responder is used as the reference group
CI = confidence interval; CS = company submission; EQ-5D-3L = European Quality of Life-5 dimensions 3 levels; OLS = ordinary least squares; SE = standard error

EAG comment: In the clarification letter, the EAG expressed concerns around the face validity of the utility values, presented in Table 4.13, which were estimated via the EQ-5D tool based on data from

the SPIRIT 1 and 2 trials, as there were no validity checks reported in the CS for these scores (question B19).[3] As an example, in clarification question B19, the EAG referred to the baseline utility score of 0.58, which was thought to be low. In their response, the company reported that this value was found to be within the range of the baseline utilities identified in the company's SLR ranging from 0.15 to 0.689 as these were presented in Table 121 of the Appendix H in CS.[5] The company further noted that the highest baseline value was 0.78 observed in a prospective observational study in France referring to the study by Grundström et al. 2019.[54] Firstly, the EAG noticed that the company erroneously referred to the study by Grundström et al. 2019, while likely describing the results and patient population in the analysis of Oppenheimer et al. 2021.[55] Secondly, the baseline of 0.78 was not included in the company's range of potential baseline utility values. Thirdly, the EAG is unclear why women in Oppenheimer et al. 2021 were considered less severe than the target population of relugolix CT. The company noted that these women did not have moderate-severe symptoms of endometriosis and were being treated with high-dose progestin and disregarded this evidence considering the patient population in the study to be less severe patients than patients treated with relugolix CT. As clarified in question B4, the target population of relugolix CT is adult patients with symptoms of endometriosis who have a history of previous medical or surgical treatment for their endometriosis, and not [redacted][3] Therefore, the EAG is still unclear why the utility values reported in the other studies included in Table 121 of Appendix H, including the study by Oppenheimer et al. 2021, were not considered relevant sources to validate the utility inputs or at least to be used as alternative sources in a scenario analysis. Fourthly, in the absence of a validation discussion, the relevant range of 0.15 to 0.689 presented by the company, may still be considered quite wide for validation purposes. Also considering the limitations of the SLR analysis presented by the company, and discussed in Section 3.1 of this report, the EAG is uncertain if the identified studies are appropriate to validate the utility values estimated in the relugolix CT trials and thinks that the estimates of the utility scores presented by the company are still subject of uncertainty. Finally, in their response to clarification question B19,[3] the company did not comment on the utility score of 0.80 estimated at 24 Weeks.

4.2.8.1 Health state utilities

To estimate response-related health state utilities the company used the OLS regression output presented in Table 4.15 above and a mean baseline utility value of 0.5838 that was observed across both treatment arms of the SPIRIT 1 and 2 trials. The company noted that the baseline value from the SPIRIT trials of 0.5838 corresponds to utility value of initial treatment. For partial response, the utility values were estimated using the average of response and non-response utility values. Table 4.15 presents the health state utility scores for both options of treatment response (i.e., change from baseline response and threshold response). Aligning with the approach used to define treatment effectiveness (see Section 4.2.6.1), the option of treatment response defined as a change from baseline response was employed to also estimate health state utilities. Following the clarification phase, the company also incorporated age-related utility decrements in the model, sourced from Szende et al. 2014 (see clarification response B20).[3, 56]

Table 4.15: Health state utility values as estimated via the OLS model

Response level	Utility	95% CI	P-value	Source
Initial treatment	0.5838	[0.5676 - 0.5999]	N/A	[33, 34]

Response level	Utility	95% CI	P-value	Source
Treatment response defined as change from baseline response				
Complete response	0.8839	[0.8697 - 0.8981]	<.0001	33, 34
Partial response	0.8014	[0.7761 - 0.8267]	N/A	Average utility of response and non-response
Non-response	0.7189	[0.6979 - 0.7399]	<.0001	33, 34
Treatment response defined as threshold response				
Complete response	0.8816	[0.8672 - 0.8960]	<.0001	33, 34
Partial response	0.7959	[0.7703 - 0.8215]	N/A	Average utility of response and non-response
Non-response	0.7102	[0.6891 - 0.7313]	<.0001	33, 34

Based on Table 62 of the CS.[1]
CI = confidence interval; CS = company submission; N/A = not applicable; OLS = ordinary least squares

The company further assumed that patients achieving response would experience the same utility levels irrespective of their treatment path. Therefore, the utility of response was applied to the health states of "response" to relugolix CT or GnRH agonists, "response to BSC", "post-hysterectomy stable", and response to "post-conservative surgery". Similarly, the utility of non-response was applied to health states of "BSC", "non-response BSC", "waiting time before surgery", "post-hysterectomy recurrence", and "PCS recurrence". Regarding non-responders, the company commented that the observed increase in the utility of non-responders from the baseline utility value of 0.5838 (initial treatment) to the utility value of 0.7189 or 0.7102 estimated from data of the SPIRIT trials (depending on the definition of response in Table 4.15 above), suggests that symptoms in non-respondents improved substantially following treatment with relugolix CT, although this increase was not sufficient to meet the criteria of treatment response.

EAG comment: The EAG has concerns around the company's approach to assume the same utility value for non-responders and the same utility value for responders irrespective of their treatment path and most importantly irrespective of the consecutive episodes of non-response. In the CS, the company noted that *"it is likely that the utility in patients who experience repetitive episodes of non-response regress back to the levels observed at baseline in the SPIRIT trials. The model does not account for the number of failures to respond to different strategies (medical treatment and surgeries), and thus the utility of non-response is thus used equally for both health states "Non-response", following initial treatment, and "Non-response BSC". To not decrease further the utility after failing another line of treatment is likely a conservative approach".*[1] Firstly, the EAG thinks it would be counterintuitive to expect that non-response to initial treatment (or BSC) would lead to the same utility value as the non-responsiveness of patients following surgery. Secondly, although the CS only refers to the potential bias related to the usage of the same utility value for all non-response health states, the EAG thinks that a similar bias may also hold for using the same utility values in all response health states. Specifically, if patients achieve response after multiple lines of treatment the health state utility score could be lower than patients achieving response after one or fewer lines of treatment. In response to EAG's question B21 related to this issue, the company adjusted the model to include health-state specific utility values but did not suggest or include any health-state specific values and they still used the original company base-case. The company further commented that "*since previous treatment failures are likely to contribute negatively to a patient's quality of life, assuming the same utility for all response health states*

is a conservative approach which is not expected to favour patients in the Relugolix CT arm versus the comparator arm".[3] To illustrate the impact of this assumption, the company presented a scenario analysis in which the health state utility values (HSUVs) for all subsequent treatments were set equal to the utility value of initial treatment health state of 0.5838. In this scenario the incremental cost-effectiveness ratio (ICER) dropped to about half of its original value showing the impact of the uncertainty around this approach. The EAG considers that this scenario may address some of the uncertainty, although unexpectedly, the EAG also noticed that the results are relatively insensitive to single changes in utility values used to define response or non-response health states. That can be supported by the fact that in the results presented by the company, about 88% of the gain in QALYs from relugolix CT treatment as compared to GnRH treatment is derived from a lower level of AEs and surgery-related complications for patients in the relugolix CT arm, and about 86% of the overall QALY gains from the lower level of the long-term surgery-related complications only (post-hysterectomy disutility). Further comments on this matter are included in the EAG comments in Section 4.2.8.3 and 5.1.

4.2.8.2 HRQoL data identified in the review

The company conducted an SLR to identify HRQoL data with details included in Appendix H of the CS.[5] In total, there were 12 records identified that met the specified eligibility criteria (Appendix G of the CS).[5] Of the 12 records, only five publications reported either direct utility/disutility values or reported use of generic QoL tools like EQ-5D as per the predefined set of criteria for inclusion of utility studies (Table 120 of Appendix N).[5] Two of the five studies sourced the utility data from other studies or databases. Four studies assessed the impact of medication treatments and surgery on EQ-5D index whereas one study compared only biopsy-proven endometriosis with healthy controls irrespective of any intervention. All included utility studies were subject to quality assessment based on the checklist adapted from Papaioannou et al. 2013[57] with further details presented in Table 122 of Appendix H.[5]

4.2.8.3 Disutility values

Disutilities due to treatment-related AEs

As reported in Section 4.2.7.1, the economic analysis incorporated AEs of Grade ≥3 that occurred at a frequency of ≥1% of patients in at least one treatment of interest (relugolix CT, BSC or GnRH agonist) of the SPIRIT trials. In the SPIRIT 1 and 2 trials, the only AE meeting these criteria was headache. However, the CS explained that following feedback from clinical experts, AEs related to hot flush, decreased libido, depression, increased blood pressure, and hair loss were also considered in the economic analysis. For each of these AEs, the company assigned disutility values which can be then applied either throughout the duration of treatment in case AEs are defined using a constant 3-month probability, or only at the time of the treatment initiation in case AEs are defined using the probability of event only at the start of the treatment (acute event). The annual disutility values with their respective source assigned to each of the AE are shown in Table 4.16.

The disutility of hot flush was obtained from a Canadian study for pre-menopausal women with uterine fibroids.[58] The disutility of headache was sourced from a CEA on moderate-to-severe migraine.[59] The disutilities for decreased libido and depression were derived from Wang et al. 2019, a CEA of moderate-to-severe endometriosis pain treatment in the United States (US).[32] No disutility was assumed for hypertension, whilst the disutility of hair loss was based on a study assessing health state utilities for non-small cell lung cancer.[60]

Table 4.16: AEs disutilities per year

Type of AE	Disutility	Source
Hot flush	-0.060	33, 34, 58
Headache	-0.340	59
Decreased libido	-0.049	32
Depression	-0.120	32
Hypertension	0	Assumption
Hair loss	-0.045	60

Based on Table 64 of the CS.[1]
AE = adverse event; CS = company submission

Disutilities due to surgery and surgical complications

The model also incorporates short- and long-term disutilities due to surgeries and disutilities due to surgical complications. Tables 4.16 and 4.18 presents all disutility values associated with the different types of surgeries and surgery-related complications.

The short-term disutilities due to surgeries were attributed to the immediate loss in QoL a patient can experience when undergoing surgery (referred to as acute disutility due to surgery in the CS) and were therefore applied only once in the model at the next cycle following surgery. The acute disutility associated with hysterectomy was estimated as the weighted average of the loss in QoL that patients experience when undergoing vaginal (-0.02), abdominal (-0.07) and laparoscopic hysterectomies (-0.04). The loss in QoL for each of the hysterectomy routes were collected from a RCT evaluating the CE of laparoscopic hysterectomy compared with standard hysterectomy,[61, 62] whilst the distribution of hysterectomies across different routes were extracted from Maresh et al 2002.[44] The acute disutility of oophorectomy was assumed equal to the acute disutility of hysterectomy as shown in Table 4.17, whereas the disutility of conservative surgery was set equal to that of laparoscopic hysterectomy (-0.04).

The long-term disutilities due to surgeries aimed to capture the permanent consequences on patients' QoL and were, therefore, implemented to all subsequent cycles following surgery. The long-term disutility of hysterectomy of -0.180 was applied in all post-hysterectomy health states (i.e., "stable", "recurrence" and "reoperation") and was derived from a global burden of disease report published by the World Health Organization (WHO).[63] This value was assumed to represent the disutility linked to infertility as women are not able to conceive post-hysterectomy. For conservative surgery there was no long-term disutility assumed as such a surgery would preserve the uterus, indicating no long-term impact.

Table 4.17: Surgery-related disutilities

Type of surgery	Disutility	Source
Short-term disutilities (acute disutilities)		
Hysterectomy	-0.0541	61, 62
Oophorectomy	-0.0541	Assumed equal to disutility associated with hysterectomy
Conservative surgery	-0.040	Assumed equal to disutility associated with laparoscopic hysterectomy[61, 62]
Long-term disutilities (post-surgery disutilities)		

Type of surgery	Disutility	Source
Post-hysterectomy	-0.180	[63]
Post-conservative surgery	0.000	Assumption
Based on Table 65 and Table 66 of the CS.[1] CS = company submission		

The disutilities due to surgical complications are summarised in Table 4.18. The surgical complications considered in the economic analysis consist of urinary tract infections, fistula, urinary retention/complications, and impact of surgery on other organs such as bowel problems. For urinary tract infections and urinary retention, the disutility value was derived from a CE study of surgical treatment for benign prostatic enlargement.[50] For fistula, the disutility was derived from an economic assessment comparing prostate cryotherapy to androgen deprivation therapy for the treatment of recurrent prostate cancer.[51] The CS noted that although both these studies included male patients, the disutility values were assumed to be representative for women as well. For the impact of surgery on other organs, the disutility value was extracted from a CEA of opioid-induced constipation in patients with advanced illness.[64] Similarly to the incidence of AEs, the model includes the option to implement the risk of surgery complications as acute (immediate following surgery) or constant (throughout the remainder time following surgery). For the base-case analysis, complications were applied as acute risks of events.

Table 4.18: Disutilities due to long-term surgical complications

Type of surgery	Disutility	Source
Urinary tract infection	-0.006	[50]
Fistula	-0.150	[51]
Urinary retention/complication	-0.006	[50]
Impact of surgery on other organs (e.g., bowel problems)	-0.017	[64]
Based on Table 67 of the CS[1] CS = company submission		

EAG comment: The main concerns of the EAG regarding the implementation of the disutilities in the model are the following:

a) The EAG noticed that to inform disutility values due to AEs or surgical complications the company used relatively older studies. In clarification question B3 the EAG expressed concerns around the appropriateness of the studies used to inform disutility values and requested further justification on their appropriateness.[3] In response to this question, the company only commented that the impact of these inputs on the results is relatively negligible without providing a proper and clear justification on the appropriateness of the sources used to inform these values. The EAG considers that the company failed to provide a satisfactory response to this question and does not agree with the company's viewpoint that the impact of these parameters is low to justify an appropriate search and selection process of potential sources. Furthermore, the EAG considers that this is not completely correct since the long-term disutilities can have a major impact on the model results.
b) An example reflecting upon the aforementioned EAG's concerns about the long-term disutility value of -0.180 used for patients undergoing hysterectomy is discussed next. This input parameter was informed from the WHO report on the Global Burden of Disease which was published in 2004,[63] with the original source for this input being the Global Burden of Disease report published

in 1990. Therefore, the EAG is uncertain if this value is still representative. Secondly, the EAG noticed that setting this value to half of its original value, would lead to an ICER which would be twice as high, indicating a non-negligible impact on the results in contradiction to the company's statement. As also mentioned in the EAG comments in Sections 4.2.8.1 and 5.1, the EAG noticed that the total QALYs results are relatively insensitive to single changes in HSUVs, used to define response or non-response health states. This was explained by the fact that in the results presented by the company, about 88% of the overall gain in QALYs from relugolix CT treatment as compared to GnRH agonists is obtained from a lower level of AEs and surgery-related complications for patients in the relugolix CT arm, and about 86% of the overall QALY gains from the lower level of the long-term surgery-related complications only (post-hysterectomy disutility). When setting the long-term disutility value to zero, disregarding thus any long-term disutility due to hysterectomy, the ICER for relugolix CT compared to GnRH agonists would be approximately 10 times higher the company's base-case ICER. This illustrates previously discussed concern about the appropriateness of the operationalisation of fertility in the model and its impact on the results.

c) Following up on the operationalisation of fertility, the company assumed a long-term disutility value of -0.180 in all post-hysterectomy health states (i.e., "stable", "recurrence" and "reoperation"), to capture the disutility linked to infertility as women are not able to conceive post-hysterectomy. However, in response to clarification question B7, which asked whether other important and clinically and economically relevant outcomes may have been omitted from the analysis, the company stated that "*disutility from infertility would only be expected to have an impact on the proportion of people actively trying to have a family*".[3] This particular issue has also been flagged by the EAG in the clarification question B22. The company in this response mentioned that "*they agree with the EAG that this parameter is uncertain, but it is difficult to determine what proportion of patients with a hysterectomy would be wishing to have children, particularly as this would change as the cohort ages. The intention was to capture QoL losses additional to those of infertility, which may include feelings of a loss of femininity associated with the loss of the uterus"*.[3] Furthermore, in clarification question B22, the EAG asked if patients on relugolix CT treatment should also be assumed to experience a similar disutility value of -0.180, since patients on relugolix CT treatment are not able to conceive while on treatment. The company noted that "*both GnRH and relugolix are contraceptive, and the disutility of infertility related to this would already have been captured within the trial EQ-5D values given that the women participating in the trials would have been aware of this"*.[3] The EAG does not agree with the company's answer, because, as also mentioned by the company above, childbearing wish may vary over age and this may not be captured in the EQ-5D trial data. In response to clarification question B7, the company went further in their response mentioning that "a *utility benefit for faster recovery of fertility following discontinuation of relugolix CT was considered too uncertain a parameter to include and would likely have little impact in results, given that the difference in time to regain in fertility between the two treatments is likely to be months rather than years"*.[3] The EAG is unclear around this statement considering that GnRH agonists are only provided for a maximum duration of 1 year compared to relugolix CT which can be administered until women's menopause, which may have a totally different impact on the patient's 'regain of fertility'.

d) The EAG considers that issues associated to infertility are quite relevant for patients with endometriosis, as it may also lead to a different choice of subsequent treatments. This is currently not appropriately incorporated in the economic analysis. To illustrate the potential impact of this, the EAG ran a *hypothetical* scenario analysis in which the disutility value associated with infertility was set to half of its original value, while assuming that this disutility would be experienced by half of the population following hysterectomy, and half of the population in the relugolix CT and GnRH agonists arms. In this scenario the ICER exceeded £30,000 per QALY gained. It should be noted

that in general the ICER is insensitive to changes in *all* input parameters, except for this long-term disutility (this is also discussed in Section 5.2.3 of this report). Therefore, it is crucial that the operationalisation of fertility in the model is exhaustively validated.

e) The EAG also noticed that the company used an additive approach to incorporate disutilities due to AEs and surgery-related complications. However, as per NICE's manual a multiplicative approach would be generally preferred.[65] The EAG would suggest the company to explore a multiplicative approach and justify their choice for a preferred method.

f) The company reported to be accounting for AEs related to headache, hot flush, decreased libido, depression, increased blood pressure, and hair loss in the economic analysis. However, considering that the probability of decreased libido and hair loss were set to zero as per Table 53 of the CS,[1] and the fact that disutility for hypertension was assumed to be zero as shown in Table 4.15 above, the EAG noticed that for relugolix CT only AEs related to hot flush and headaches were considered in the economic analysis, whilst for GnRH agonists AEs related to hot flush, headaches and depression. As noted in Section 2.1 of this report, based on the clinical expert opinion received by the EAG, relugolix CT should be used as a 2nd to 3rd line treatment for moderate to severe symptoms of fibroids as well as pain associated with endometriosis; but it would be carefully considered because of its unfavourable side-effect profile. The EAG is unclear why the company assumed a zero-disutility due to hypertension and no justification was provided in the CS.

4.2.8.4 Summary of utility values used in the CEA

A summary of all utility values used for the health states and all disutility values used for treatment-related AEs, surgeries and surgical complications is provided in Table 4.19.

Table 4.19: Summary of utility and disutility values used in the economic analyses

Health state/Type or surgery/Surgical complication	Utility/ Disutility	95% CI	Reference in CS[1]	Justification
Initial treatment	0.5838	(0.5676, 0.5999)	Section B.3.4 (page 160)	EQ-5D data SPIRIT 1 and 2 trials[33, 34]
Treatment response defined as change from baseline response (model base-case)				
Complete response	0.8839	(0.8697, 0.8981)	Section B.3.4 (page 160)	OLS models using EQ-5D data from the SPIRIT 1 and 2 trials[33, 34]
Partial response	0.8014	(0.7761, 0.8267)		
Non-response	0.7189	(0.6979, 0.7399)		
Treatment-related AEs disutilities				
Hot flush	-0.06	(-0.05, -0.07)	Section B.3.4 (page 162)	[33, 34, 58]
Headache	-0.34	(-0.31, -0.37)		[59]
Decreased libido	-0.05	(-0.04, -0.05)		[32]
Depression	-0.12	(-0.11, -0.13)		[32]
Hypertension	0	N/A		Assumption
Hair loss	-0.05	(-0.04, -0.05)		[60]

Health state/Type or surgery/Surgical complication	Utility/ Disutility	95% CI	Reference in CS[1]	Justification
Short-term surgery-related disutilities				
Conservative surgery	-0.04	(-0.04, -0.04)	Section B.3.4 (page 163)	[61, 62]
Hysterectomy	-0.05	(-0.05, -0.06)		
Oophorectomy	-0.05	(-0.05, -0.06)		
Long-term surgery-related disutilities				
Post-hysterectomy	-0.18	(-0.16, -0.20)	Section B.3.4 (page 163)	[63]
Post-conservative surgery	0	N/A		Assumption
Disutilities of long-term complications from surgery				
Urinary tract infection	-0.01	(-0.01, -0.01)	Section B.3.4 (page 164)	[50]
Fistula	-0.15	(-0.15, -0.17)		[51]
Urinary retention/ complication	-0.01	(-0.01, -0.01)		[50]
Impact of surgery on other organs (e.g., bowel problems)	-0.02	(-0.02, -0.02)		[64]

Based on Table 68 in the CS.[1]
AE = adverse event; CI = confidence interval; CS = company submission; EQ-5D = European Quality of Life-5 dimensions; N/A = not applicable; OLS = ordinary least squares

4.2.9 Resources and costs

The following cost categories were included in the analysis: drug acquisition and administration costs for relugolix CT, concomitant and add-back therapies, costs associated to visits to health care professionals, to surgery and its complications, costs for the treatment of AEs. Unit prices were based on the NHS Reference Costs, British National Formulary (BNF) and Personal Social Services Research Unit (PSSRU) in line with the NICE reference case.

4.2.9.1 Resource use and costs data identified in the review

According to the CS, a SLR was conducted to identify cost and healthcare resource use data for premenopausal women diagnosed with clinically confirmed endometriosis and those experiencing endometriosis-associated pain, without restrictions on definitions of pain severity. The SLR identified 11 unique relevant publications (nine full text, two conference abstracts).[5]

The studies were published between 2018 and 2022. Eight studies were cost and resource use-based studies, the other three were model based studies. Cost data was reported in seven of the 11 studies. Six

studies reported currency, and among those the majority of cost data was reported in United State Dollar (USD; 50%), followed by Australian Dollar (AUD: 33.3%) and Great British Pound (GBP; 16.6%).[5] Four studies reported direct costs data with a primary focus on drug cost and inpatient/outpatient related cost, two studies reported direct costs data about surgery related costs. Indirect costs data was reported in two studies, which highlighted the effect of pain severity on productivity costs and impact on office visits.[25, 66] One study reported cost data only in terms of odds ratio for association of endometriosis-related symptoms to sick leave and productivity loss.[67]

Resource use data were reported in six of the 11 studies. Three studies presented data on resource utilisation, focusing on patient visits and hospital stays. Two studies provided information on mean surgical durations and healthcare resource utilisation (HCRU). Additionally, one study reported HCRU data specifically for outpatient and emergency department (ED) visits.

In the study by Saine et al. 2020 significant differences were reported in various direct cost domains when comparing the use of non-opioid and opioid drug regimens.[68] Additionally, the same study revealed significantly higher healthcare utilisation in the opioid group compared to the non-opioid group across all measures, the most notable differences being the number of pharmacy visits (14.70 versus 5.74 visits), surgery visits (14.34 versus 12.32 visits), and outpatient visits (15.36 versus 13.69 visits).[69]

A detailed overview of the results can be found in Table 124 of Appendix I.[5]

EAG comment: According to the CS, Appendix I, the SRL focussed on premenopausal women diagnosed with clinically confirmed endometriosis and those experiencing endometriosis-associated pain, without restrictions on definitions of pain severity.[5] In response to the clarification letter B4, the company however defined the population included in the economic analyses as adult patients with symptoms of endometriosis who have a history of previous medical or surgical treatment for their endometriosis.[3] It is unclear to the EAG if this population was also used in the SLR.

The CS describes the results of the SLR and summarises the results in Table 124 of Appendix I.[5] However, they do not specify which of the studies were used to inform the model parameters. The EAG matched the references and noticed that only three studies were cited in the CS.[24, 25, 32] These studies are cited in the model structure Section to refer to clinical practice guidelines for endometriosis, but not in any of the tables presenting costs or resource use. The EAG also did not find these references back in the source column of the input sheet of the economic model. Therefore, it seems that none of the cost and resource use information found in the SLR was used to derive direct input values for the economic model.

4.2.9.2 Acquisition and administration costs

The drug acquisition costs for relugolix CT and GnRH agonists at list prices are summarised in Table 4.20. Drug acquisition costs are applied as long as patients are on active treatment. For GnRH agonists, a 50/50 split between the least costly short-acting and long-acting GnRH agonist (triptorelin) was assumed. No drug wastage was assumed.

Table 4.20: Drug acquisition costs and dosing

Treatment	Package cost (£)	Doses per package	Administrations per cycle	Total drug cost per cycle (£)	Source
Relugolix CT	72.00	28	91.31	234.80	[70]

Treatment	Package cost (£)	Doses per package	Administrations per cycle	Total drug cost per cycle (£)	Source
Short-acting GnRH agonist Triptorelin (3.75 mg)	69.00	1	3	225.02	NR
Long-acting GnRH agonist Triptorelin (11.25 mg)	207.00	1	1	207.00	[71]
Adjusted from Table 69 of CS.[1] CS = company submission; CT = in combination with oestradiol and norethisterone acetate; GnRH = gonadotropin-releasing hormone; NR = not reported					

Treatment administration costs are presented in Table 4.21. The relevant forms of administration for the intervention and comparator, as well as subsequent treatment (BSC) encompass intramuscular injection/SC injection, and oral administration. Relugolix CT is an orally administered tablet, thus, no administration costs are assumed. For GnRH agonists, the company assumed that this would be administered by a nurse based in a general practitioner (GP) setting. The company state that this assumption was validated with clinical experts who were asked to provide information regarding who administers GnRH agonist treatment, the duration required for treatment administration and the setting in which they would be administered (hospital or GP practice).

Table 4.21: Administration costs

Mode of administration	Administration - Resource use	Unit cost (£)	Source
Oral/ intranasal administration	Self-administered	0	Assumption
Intramuscular/SC injection	GP practice/specialty care nurse-administered	26	Cost of qualified nurse for 30 minutes, Unit Costs of Health & Social Care 2022[72]
Based on Table 70 of the CS[1] CS = company submission; GP = general practitioner; SC = subcutaneous			

EAG comment: The company base-case does not include drug wastage because it is assumed that BSC only includes NSAIDs, which are an oral tablet. The EAG agrees that drug wastage is not an issue with oral tables. This might become relevant for injections, which is the administration form of GnRH agonists, although the impact on the results is expected to be minimal.

In response to clarification question B23, the company explained that BSC should not be included in Table 4.19 (as originally reported in the CS) because it is not a comparator.[3] Best supportive care is part of the concomitant medication, the only comparator is GnRH agonists.

4.2.9.3 Concomitant medication

Concomitant medication is taken by patients in combination with medical treatment. For pain management, analgesics (i.e., NSAIDs) are included in the model (as BSC) and the dose frequency is

based on the SPIRIT trials that is driven by patients' response status (see Section 4.2.6.2 for details). In the model, it is assumed that patients do not require opioids, as these treatments are rarely prescribed in Europe. The company based this assumption on identified studies through a targeted literature review and their decision was confirmed/supplemented by clinical experts at an advisory board. Table 4.22 summarises the costs associated to NSAIDs. In the model, it is assumed that patients do not require opioids, as these treatments are rarely prescribed in Europe. The company based this assumption on identified studies through a targeted literature review and their decision was confirmed/supplemented by clinical experts at an advisory board.

Table 4.22: Concomitant medication costs

Concomitant medication	Package cost (£)	Doses per package	Administrations per cycle	Total cost per cycle (£)	Source
NSAIDs (ibuprofen 400 mg)	4.90	60	273.94	22.37	[73]
Based on Table 71 of the CS[1] CS = company submission; mg = milligram; NSAIDs = non-steroidal anti-inflammatory drugs					

EAG comment: In response to questions B10 and B23,[3] the company clarified that BSC is only comprised of symptomatic treatment for pain management, i.e. analgesics (NSAIDs). Other treatments originally reported in the CS are not presented in Table 4.22. The EAG noticed that removing the cost for the NSAIDs resulted in a minor change in costs and ICER.

Add-back therapy

Add-back therapy is the prescription in addition to GnRH agonists when used for longer term and after oophorectomy. With GnRH agonists, patients are initiated on add-back therapy typically at 3 months whereas after oophorectomy, add-back therapy is initiated directly following surgery. The company stated that this assumption was confirmed with clinical experts at an advisory board. In the base-case analysis, an equal split (50/50%) between tibolone and raloxifene as add-back therapy is considered. The company assumed that 100% of patients treated with GnRH agonists or oophorectomy get the add-back therapy. Add-back therapy is not prescribed when GnRH agonists are used short-term prior to surgery. Costs of add-back therapy are summarised in Table 4.23.

Table 4.23: Cost of add-back therapy

Add-back therapy	Cost per package (£)	Tablets per package	Tablets per day	Total drug cost per cycle (£)	Source
Tibolone 2.5 mg	14.13	84	1	15.36	[74]
Raloxifene 60 mg	4.55	28	1	14.84	[75]
Based on Table 72 of the CS[1] plus response to clarification letter B25[3] CS = company submission; mg = milligram					

EAG comment: In response to clarification question B25,[3] the company provided information on the dosing schedule of patients using the add-back therapy in the GnRH agonists arm. The dosing schedule for tibolone is 2.5 mg daily as an oral tablet and raloxifene it is 60 mg as a daily oral tablet. The company assumed that all GnRH agonist patients comply with the daily add-back therapy. The EAG noticed that differences in the assumption regarding an equal split (50/50%) between tibolone and raloxifene had a

minor effect on the results (cost difference less than £100). Removing add-back therapy completely had a minimal impact of less than £100 upon the total costs and the ICER.

Visit to health care professionals, tests, and procedures

Table 4.24 provides a summary of the costs of visits to healthcare professionals.

Table 4.24: Unit costs of healthcare professional visits

Healthcare provider	Cost per visit (£)	Source
Gynaecologist	181.26	[76, 77]
General Practitioner	42.00	[78]
Nurse	7.99	
Based on Table 74 of the CS[1] CS = company submission		

Table 4.25 shows that only the gynaecologist is involved for administration and monitoring until the time of evaluation of treatment response (6 months). It also shows that in the long-term follow-up (beyond the first three cycles) a nurse is required for both relugolix CT and GnRH agonists. The company assumed that the frequency of resource use is equal between relugolix CT and GnRH agonists. In the model, resource utilisation is calculated for each model cycle (quarterly).

Table 4.25: Resource used linked to administration and monitoring until treatment response evaluation and long-term follow-up, by treatment arm

Healthcare provider	Administration		Per cycle frequencies for long-term follow-up		Source
	Treatment initiation	6-month follow-up	Relugolix CT	GnRH agonist	
Gynaecologist	Yes	Yes	No	No	UK KOL input
General Practitioner	No	No	No	No	
Nurse	No	No	Yes	Yes	
Based on Table 75 and 76 of the CS[1] CS = company submission; CT = In combination with oestradiol and norethisterone acetate; GnRH = Gonadotropin-releasing hormone; KOL = key opinion leaders; UK = United Kingdom					

The company reported that clinical experts explained that patients who are being treated with a pharmacological treatment such as GnRH agonists are not subject to any additional monitoring with tests and procedures. Patients who undergo surgery, undergo on average one ultrasound and one MRI scan.[1] The cost of these tests are provided in Table 4.26.

Table 4.26: Cost of test and procedures

Test/procedure	Cost (£)	Source
Ultrasound	181.00	[79]
MRI scan	114.00	
Based on Table 78 of the CS.[1] CS = company submission; MRI = magnetic resonance imaging		

EAG comment: The CS, Table 78 included a blood test and a dexa scan as part of the additional tests and procedures.[1] However, in the model the resource use for these tests was 0. The EAG therefore did not include the costs of blood tests and dexa scans in Table 4.25.

Cost of surgery

Costs of different surgical procedures were sourced from pricelists available through the NHS and were validated with clinical experts. The cost and proportions of different routes of hysterectomy are summarised in Table 4.27. The cost of laparoscopy is derived from the NHS England 2022/23 national tariff workbook (Annex A) unit costs for "Major, Intermediate and Minor Laparoscopic or Endoscopic, Upper Genital Tract Procedures", by taking an average of the three unit costs.[79] The company assumed that the costs of conservative surgery are the same as the cost of laparoscopic hysterectomy. The company stated that there are no costs of oophorectomy available in the NHS England 2022/23 national tariff workbook. As an alternative, the company calculated the cost of conservative surgery using the unit price for hysterectomy (£4,381.69) from the NHS England 2022/23 national tariff workbook and weighting this by multiplying it with the proportion (0.61 = 2,275/3,703) of hysterectomy (3,703) and oophorectomy (2,275) costs from the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS).[79, 80]

Table 4.27: Cost and frequencies of different surgeries

Type	Cost (£)	Proportion of patients	Procedure	Source
Routes of hysterectomy				
Vaginal	4,414.00	30%	Major open upper genital tract procedures, average of CC scores 0-5+, currency codes MA07G, MA07F, MA07E; weighted average of elective, day case, and outpatient unit costs	44, 79
Abdominal	4,414.00	67%		
Laparoscopic	3,337.00	3%	Major, laparoscopic or endoscopic, upper genital tract procedure, average of CC scores 0-2+, currency codes MA08B, MA08A; weighted average of elective, day case, and outpatient unit costs	44, 79
Cost of surgery				
Conservative surgery (assumed same as laparoscopic hysterectomy)	3,337.00		Major, laparoscopic or endoscopic, upper genital tract procedure, average of CC scores 0-2+, currency codes MA08B, MA08A; weighted average of elective, day case, and outpatient unit costs	79
Oophorectomy	2,691.96		Unit price for hysterectomy and oophorectomy from the NHS England 2022/23 national tariff workbook multiplied with the proportion of hysterectomy and oophorectomy (2275/3703).	79, 80
Hysterectomy	4,381.69		Based on the cost and frequencies as presented in first part of the Table: 4,414 * 0.97 + 3,337 * 0.03	

Based on Table 79 and Table 80 of the CS[1]

CS = company submission; NHS = National Health Service

4.2.9.4 Health state costs

In the model, the frequency of monitoring and disease management related healthcare resource use is not driven by health states but instead by whether patients are on active pharmacological treatment (relugolix CT or GnRH agonists), or they have undergone surgery. The assumptions underpinning healthcare resource use applied in the model are presented in Table 4.25.

4.2.9.5 AE costs, surgery, and complications costs

Table 4.28 presents the cost for AE, complications from surgery and cardiovascular events and fractures included in the economic analyses. In the model, the costs of AEs are applied in the cycle in which they occur. The frequencies and probabilities of experiencing AE are presented in Section 4.2.7 of this report.

Table 4.28: Costs of AEs and complications

Type	Cost (£)	Cost detail	Source
Costs of AEs related to medical treatment			
Hot flush	0	No cost incurred as it is assumed that this will be self-managed, and no treatment sought	Assumption
Headache			
Decreased libido	42	Based on unit cost for a GP, per surgery consultation lasting 9.22 minutes, excluding travel	[78]
Depression			
Blood pressure			
Hair loss			
Cost of complications related to surgery			
Urinary tract infection	457.35	Based on the non-elective short stay unit cost for kidney or urinary tract infections, without interventions, with CC Score 0-1 (LA04S)	[77]
Fistula	4,039.00	Based on cost of fistula	[51]
Urinary retention/complication	612.62	Based on the non-elective short stay unit cost for kidney or urinary tract infections, without interventions, with CC Score 4-7 (LA04Q)	[77]
Impact of surgery on other organs (e.g., bowel problems)	1,020.76	Based on the non-elective short stay unit cost for diagnostic colonoscopy, 19 years and over (FE32Z)	[77]
Cost of cardiovascular events and fractures			
Cardiovascular event	2,648.00	Based on an average of the costs for angina (EB13A-D), actual or suspected myocardial infarction (EB10A-E), stroke (AA35A-F), heart failure or shock (EB03A-E), transient ischaemic attack (AA29C-F), peripheral vascular disorders (YQ50A-E)	[77]

Type	Cost (£)	Cost detail	Source
Hip fracture	8,686.10	Based on an average of the elective costs for hip fracture without interventions, with CC Score 0-3 to 12+ (HE11H to HE11E)	[77]
Based on Tables 81, 82 and 83 of the CS[1] AE = adverse events; CS = company submission			

4.2.9.6 Societal costs

The company also provided information on the societal burden of the disease and explored the impact on the results in a scenario analysis. The NICE guidelines recommend using a healthcare perspective. Therefore, the Section about societal cost is not summarised or commented on by the EAG. We refer to Sections B.3.5 and B.3.11 of the CS for details.[1]

4.2.10 Disease severity

The NICE reference case stipulates that the committee will regard all QALYs as being of equal weight. However, the committee may consider the severity of the condition, as determined by the absolute and proportional QALY shortfall (including discounting at the reference case rate), as decision modifier. Severity can be then taken into account quantitatively in the CEAs through QALY weighting, based on the absolute and proportional shortfall, as shown in Table 4.29. Whichever implies the greater severity level will be considered, and if either the proportional or absolute QALY shortfall falls exactly on the cut-off between two severity levels, the higher level will apply.[65]

Table 4.29: QALY weightings for disease severity

QALY weight	Proportional QALY shortfall	Absolute QALY shortfall
1.0	Less than 0.85	Less than 12
1.2	From 0.85 to 0.95	From 12 to 18
1.7	At least 0.95	At least 18
QALY = quality-adjusted life year		

The company indicated that the QALY shortfall for relugolix CT was calculated using the online calculator tool published by Schneider et al. 2021,[81] assuming a 100% of females and an average age of 33.9 years. Based on this calculation, the company concluded that a QALY weighting of 1 should be used for this appraisal. The company also mentioned that, since this is the first NICE evaluation for interventions treating pain associated with endometriosis, it was not possible to provide a summary list of QALY shortfall calculations used in previous appraisals.

EAG comment: The QALY shortfall results presented by the company were validated by the EAG with the Institute for Medical Technology Assessment (iMTA) Disease Burden Calculator (iDBC). In addition, the iDBC tool also estimates the likelihood of the applicable QALY weight based on the PSA results provided in the company's model, which can be used to estimate the severity adjusted probability of being cost-effective.[82] The iDBC tool can be found here: https://imtamodels.shinyapps.io/iDBCv2_1/. The QALY shortfall calculations conducted by the EAG were in line with those presented by the company.

5. COST EFFECTIVENESS RESULTS

5.1 Company's CE results

Table 5.1 shows the company's base-case deterministic discounted CE results as presented in the original CS. In response to the clarification letter, the company made some modifications to their model. These are discussed in Section 6. These results indicated that relugolix CT was both more costly and more effective than GnRH agonists. In particular, relugolix CT accrued 0.71 incremental QALYs at £1,182 additional costs. Therefore, the ICER was £1,670 per QALY gained. The disaggregated discounted QALYs, disutilities and costs are shown in Tables 5.2, 5.3 and 5.4, respectively.

Table 5.1: Company base-case deterministic CE results

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)
Relugolix CT	11,473	11.80	9.75				
GnRH agonists	10,291	11.54	9.05	1,182	0.26	0.71	1,670

Based on Table 88 in the CS.[1]
CE = cost effectiveness; CS = company submission; CT = in combination with oestradiol and norethisterone acetate; GnRH = gonadotropin releasing hormone; ICER = incremental cost-effectiveness ratio; Inc. = incremental; LYG = life years gained; QALY = quality-adjusted life year

Table 5.2: Disaggregated QALYs results

Health state	QALY relugolix CT	QALY GnRH agonists	Incr.	Absolute increment	% absolute increment
Initial treatment	0.07	0.07	0.00	-0.09	0%
Response	4.44	0.34	4.09	4.00	4340%
Partial response	0.06	0.06	0.00	-0.09	0%
Non-response	0.24	0.30	-0.06	-0.15	-59%
Best supportive care	0.57	0.96	-0.38	-0.48	-408%
Pre-surgery waiting time	0.35	0.54	-0.19	-0.28	-197%
Post conservative surgery	0.71	1.12	-0.41	-0.50	-431%
Post hysterectomy	4.16	7.13	-2.97	-3.06	-3145%
Total	**10.62**	**10.52**	**0.09**	**0.00**	**100%**

Based on Table 126 in Appendix J of the CS.[5]
CS = company submission; CT = in combination with oestradiol and norethisterone acetate; GnRH = gonadotropin releasing hormone; Incr. = increment; QALY = quality-adjusted life year

Table 5.3: Disaggregated QALYs lost due to disutilities

Health state	QALY relugolix CT	QALY GnRH agonists	Increment	Absolute increment	% absolute increment
Adverse events	-0.001	-0.002	0.000	-0.613	0%
Complications	0.000	0.000	0.000	-0.613	0%
Surgery (Acute)	-0.012	-0.018	0.006	-0.607	99%

Health state	QALY relugolix CT	QALY GnRH agonists	Increment	Absolute increment	% absolute increment
Surgery (Long-term)	-0.851	-1.457	0.606	-0.007	1%
Total	**-0.864**	**-1.477**	**0.613**	**0.00**	**100%**

Sourced from company's electronic model[83]
CS = company submission; CT = in combination with oestradiol and norethisterone acetate; GnRH = gonadotropin releasing hormone; QALY = quality-adjusted life year

Table 5.4: Disaggregated cost results (£)

Item	Cost relugolix CT	Cost GnRH agonists	Increment	Absolute increment	% absolute increment
Drug costs	5,226.66	872.69	4,353.96	3,172.22	368%
Administration costs	0.00	223.80	-223.80	-1,405.55	-19%
Other medication treatment costs	516.20	664.51	-148.31	-1,330.05	-13%
Healthcare visits	1,552.06	2,116.81	-564.74	-£1,746.49	-48%
Medical tests and exams	293.17	451.11	-157.94	-£1,339.69	-13%
Cardiovascular events	6.66	0.85	5.81	-£1,175.93	0%
Adverse events	0.10	0.63	-0.53	-£1,182.27	0%
Conservative surgery	1,474.12	2,259.65	-785.53	-£1,967.28	-66%
Hysterectomy	2,365.58	3,642.12	-1,276.54	-£2,458.28	-108%
Complications following surgery	38.24	58.88	-20.64	-£1,202.39	-2%
Total	**£11,473**	**£10,291**	**£1,182**	**£0.00**	**100%**

Based on Table 127 in Appendix J of the CS.[5]
Note: Costs associated to fractures and indirect costs were assumed to be £0 and they are not shown in the table.
CS = company submission; CT = in combination with oestradiol and norethisterone acetate; GnRH = gonadotropin releasing hormone

Overall, the new technology is modelled to affect QALYs by:

- Increasing the number of QALYs in "response" health states.
- Reducing the number of QALYs post-hysterectomy.
- In all other health states, the difference in QALYs is not substantial.

Overall, the technology is modelled to affect costs by:

- Its higher unit price compared to current treatments.
- Decreasing costs associated to surgery and health care visits.

EAG comment: The health state difference in QALYs between relugolix CT and GnRH agonists is 0.09, as shown in Table 5.2. The main gain in QALYs for relugolix CT is therefore due to the long-term disutilities after surgery applied in the model. As discussed in Section 4.2.8, this parameter is a key driver of the model results, and it is crucial to carefully address the uncertainty around its value and its implementation in the model before drawing any conclusions from this CEAs.

5.2 Company's sensitivity analyses

5.2.1 PSA

The company conducted a PSA where key input parameters were sampled simultaneously from their corresponding probability distributions over 1,000 iterations. The input parameters and the probability distributions used in the PSA can be found in Table 129 in Appendix M of the CS.[5] The average PSA results are summarised in Table 5.4, where it can be observed that these are nearly identical to the deterministic ones shown in Table 5.1.

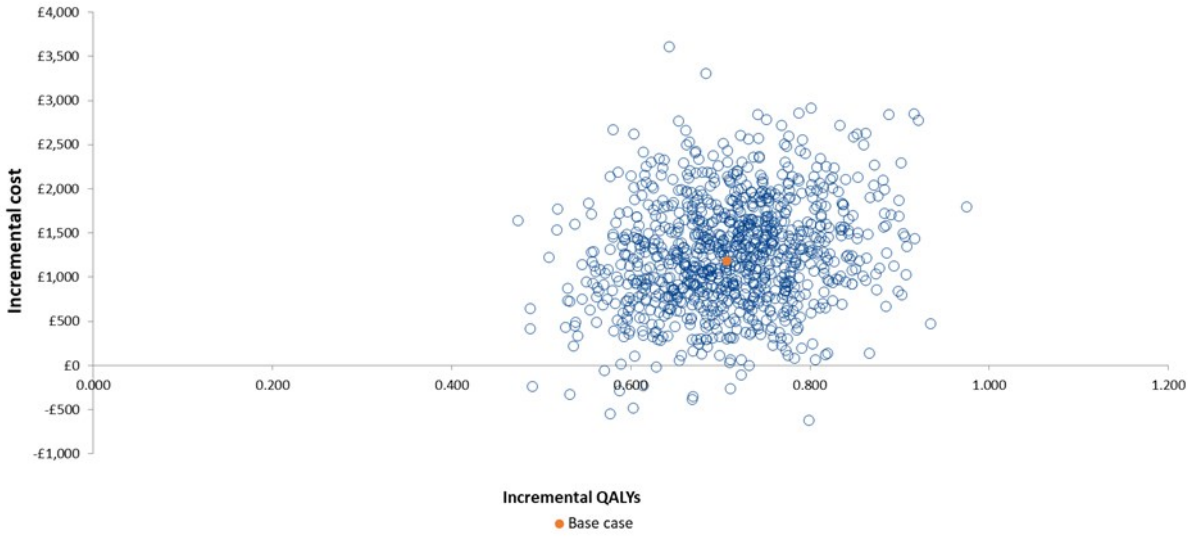
Table 5.5: Company PSA results

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)
Relugolix CT	£11,440	11.79	9.75				
GnRH agonists	£10,258	11.53	9.04	£1,182	0.26	0.70	£1,677

Based on Table 89 in the CS.[1]
CS = company submission; CT = in combination with oestradiol and norethisterone acetate; GnRH = gonadotropin releasing hormone; ICER = incremental cost-effectiveness ratio; Inc. = incremental; LYG = life years gained; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year

The company also plotted the PSA outcomes on a CE-plane, which is shown in Figure 5.1. All of the PSA outcomes were located on the eastern quadrants of the CE-plane, where relugolix CT is more effective than GnRH agonists. Nearly all the outcomes were in the northern quadrant where relugolix CT is also more costly. All of them were below the common thresholds of £20,000–£30,000 per QALY gained, as can be seen in the cost effectiveness acceptability curve (CEAC) plot in Figures 5.2. Therefore, at the thresholds of £20,000 and £30,000 per QALY gained, the estimated probability that relugolix CT is a cost-effective alternative to GnRH agonists was 100%.

Figure 5.1: PSA CE-plane (relugolix CT versus GnRH agonists)



Based on Figure 40 in the CS.[1]
CE = cost effectiveness; CS = company submission; CT = in combination with oestradiol and norethisterone acetate; GnRH = gonadotropin releasing hormone; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year

Figure 5.2: PSA CEAC (relugolix CT versus GnRH agonist)

Based on Figure 41 in the CS.[1]
Note: RYEQO = relugolix CT
CEAC = cost effectiveness acceptability curve; CS = company submission; CT = in combination with oestradiol and norethisterone acetate; GnRH = gonadotropin releasing hormone; PSA = probabilistic sensitivity analysis

EAG comment: The company's PSA results are nearly identical to the deterministic ones. The company concluded that the little variation in incremental costs and QALYs indicates a minor impact of the parameter uncertainty on the CE results and, therefore, that the analysis is robust. The EAG disagreed with this conclusion and suggested that the little variation might be caused because many input parameters were not included in the PSA, and for those that were included a fixed 10% variation from the mean *for all parameters* was assumed. In clarification question B30,[3] the EAG asked the company to include uncertainty ranges around each input parameter separately. These ranges should be implemented according to the parameter's source, since this uncertainty is unlikely to be the same for all parameters. At a minimum, the fixed standard error (SE) for all parameters should be removed from the model and let all parameters to have their own SE. Special attention is required to those parameters for which a non-symmetric CI is to be expected (for example, a hazard ratio (HR)). In their response, the company explained that the fixed SE was replaced as suggested where the SE was either directly reported or calculated from CIs were available. However, it was still unclear for how many parameters this was done. For all the remaining parameters a fixed SE of 10% is still assumed (even though this value can be changed in the model). The EAG considers that this issue was partially addressed by the company.

Also, in clarification question B31,[3] the EAG asked the company to include in the PSA those parameters that were originally excluded (for no obvious reason). The company answered that this was corrected but did not provide any overview of the parameters that were (or were not) included in the PSA nor provided any justification.

Some parameters associated to regression equations were included in the PSA but were sampled independently instead of using the variance/covariance matrices. In response to clarification question B32,[3] the company explained that the covariance matrix from the Framingham Risk Function for Cardiovascular Events was not available and therefore could not be implemented in the PSA. However, SEs from the estimated beta-coefficients were included.

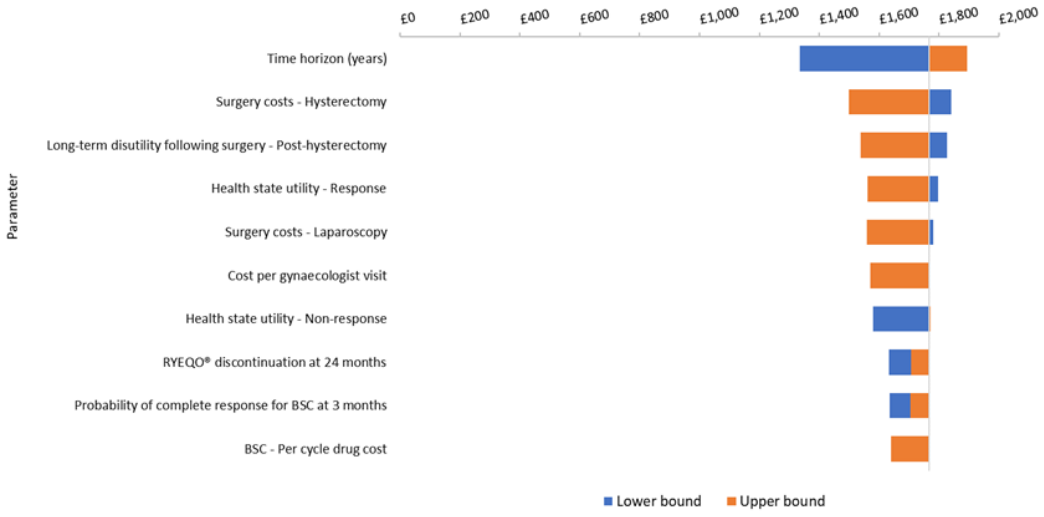
The company fixed an error detected when increasing the SE to 20%. This was caused by the starting age parameter which could occasionally drop below 18 years and, in that case, the lookup function for mortality (on Mortality sheet) produced an error. In addition, age was included in the PSA, the model time horizon and age at menopause. The model time horizon was adjusted so that it never goes beyond 100 years, and it is constrained to be at least 1 year long. A restriction on the sampled value was implemented to ensure that age at menopause was at least 1 year more than age at baseline (to ensure that the model runs for at least 1 year).

Finally, in response to clarification question B11,[3] the company updated the model to include the ITC point estimate (as OR) and its associated CrI. In particular, the point estimates from the analysis of the network for overall pelvic pain, with random effects and weakly informative priors, were used to derive the response rates of GnRH agonists using an OR of 1.1 with (0.032, 41) as 95% CrI.

5.2.2 DSA

The company conducted a deterministic sensitivity analysis (DSA) to explore the sensitivity of the model results to changes in individual parameter values. The company indicated that parameters that could not be varied without compromising the integrity of the Markov model were excluded from the DSA, including the distribution of subsequent treatment strategy after discontinuation of medical therapy or recurrence of pain. In addition, binary variables (such as definition of treatment response and stopping rule) were also not included in the DSA. The parameters that were included in the DSA were varied by assuming a 10% deviation from the base-case value (except for the annual discount rates). The results of the DSA were presented by the company in the form of tornado diagram showing the 10 parameters with the largest influence on the ICER, which can be seen in Figure 5.3. None of the changes resulted in an ICER above £2,000 per QALY gained.

Figure 5.3: DSA tornado diagram for relugolix CT versus GnRH agonists



Based on Figure 42 in the CS.[1]
Note: RYEQO = relugolix CT
BSC = best supportive care; CS = company submission; CT = in combination with oestradiol and norethisterone acetate; DSA = deterministic sensitivity analysis; GnRH = gonadotropin releasing hormone

EAG comment: As mentioned with the PSA, the company also assumed a 10% deviation from the mean for all parameters. This was corrected after clarification.

5.2.3 Scenario analysis

The company presented in total the results of seven scenario analyses to assess the robustness of the model results to changes in certain modelling assumptions. A summary of the results of these scenarios is provided in Table 5.5. These included exploring an alternative definition of response, changing the time point for evaluation of complete response, considering different durations of GnRH agonists treatment, adjusting GnRH agonists and HRT dose intensities, increasing the waiting time for surgery, and applying a societal perspective for the economic analysis. All the scenarios resulted in ICERs below £2,000 per QALY gained, indicating that the modelling assumptions explored by the company had a minor effect on the ICER.

Table 5.6: Summary of company scenario analyses

Scenario	Description (base-case)	Description (scenario)	Inc. costs	Inc. QALYs	ICER (£/QALY)
Base-case	-	-	**£1,182**	**0.71**	**£1,670**
1. Definition of response	Change from baseline: NRS score reduction from baseline of both 2.8 for dysmenorrhea and 2.1 for NMPP and no increase of analgesic use.	Threshold: achieving or maintaining a threshold below 4 in NRS scale (mild pain) for both NMPP and dysmenorrhea and no increase of analgesic use	£1,315	0.76	£1,742
2. Timepoint for evaluation of complete response	6 months	3 months	£778	0.48	£1,622
3. Duration of GnRH agonist treatment	12 months	6 months	£1,270	0.73	£1,739
4. Duration of GnRH agonist treatment	12 months	24 months	£803	0.62	£1,288
5. GnRH agonist and HRT dose intensity	100%	50%	£1,205	0.71	£1,703
6. Waiting time for surgery	6 months	12 months	£1,210	0.71	£1,711
7. Perspective for analysis	Payer	Societal	£101	0.71	£143

Based on Table 90 in the CS.[1]
CS = company submission; GnRH = gonadotropin releasing hormone; HRT = hormone replacement therapy; ICER = incremental cost-effectiveness ratio; Inc. = incremental; NMPP = non-menstrual pelvic pain; NRS = numerical rating scale; QALY = quality-adjusted life year

EAG comment: The main concerns of the EAG relate to:

- The number of scenarios seems a priori insufficient. Change in utilities (and disutilities), in effectiveness parameters, treatment effect waning, impact of fertility for example were not explored in detail by the company.
- The plausibility of the selected scenarios was not discussed.
- The EAG is aware that the ICERs are small in general. However, the EAG noticed that including treatment effect waning and adjusting the impact of the long-term disutility associated to loss of fertility can have a substantial impact on the ICER. Note that the changes in the model that would result in a decrease of incremental QALYs can make the ICER increase quite quickly. Therefore, it is crucial to assess the robustness of the model to changes that can decrease the estimated incremental QALYs.

5.3 Model validation and face validity check

Some of the validation efforts conducted on the economic model were briefly discussed in the validation Section of the CS (B.3.14).[1] The validation efforts discussed in Section B.3.14 of the CS referred to an internal validation conducted by the developers. The model structure and clinical assumptions were discussed and ratified during an advisory board which included UK clinical experts and industry representatives. In addition, the company indicated that key opinion leader (KOL) engagement was enhanced by means of primary research interviews with consultant gynaecologists. In these interviews the model assumptions, particularly those pertaining to HRU, were discussed in more detail. Finally, feedback was also elicited from a sample of five KOLs via email. Other validation aspects, such as the validation of some input parameters or how KOL feedback was used to validate other modelling features are scattered over Document B of the CS.[1] In addition, more details about model validation were provided by the company in response to some clarification questions.[3] In the remaining of this Section, the validation efforts performed on the model, as presented by the company, are categorised according to the types of validation used in the Assessment of the Validation Status of Health-Economic decision models (AdViSHE) tool.[84]

5.3.1 Validation of the conceptual model

5.3.1.1 Face validity testing (conceptual model)

On page 136 of the CS,[1] the company indicated that the model structure, including clinical/treatment pathway, and key assumptions, was validated with clinical experts during the so-called global advisory board, but no additional details were provided.

5.3.1.2 Cross-validity testing (conceptual model)

According to the company, this submission represents the first NICE assessment for interventions treating pain associated with endometriosis. Therefore, there are no previous NICE appraisals that can be used for cross-validation. However, it is unclear whether other CE models like those identified in Section 4.1, could have been used to cross-validate at least some parts of the model.

5.3.2 Input data validation

5.3.2.1 Face validity testing (input data)

Input parameters were estimated from different sources of data. Some brief discussion about the face validity of some input parameters is presented in an unstructured way throughout the CS. This is summarised below:

- The validity of the ITC is discussed in detail in Section 3.4 of this report.
- The choice of the comparator (GnRH agonists), the alternative definition of treatment response, the distribution of the subsequent strategies to manage endometriosis (following treatment discontinuation), or the distribution of strategies to manage endometriosis in case of pain recurrence following conservative surgery were validated by clinical experts, but no additional details were provided.
- Cost and resource use assumptions regarding GnRH agonists and surgical procedures were also validated by clinical experts. Again, additional details on the validation procedure were not provided.

The EAG noticed that some clinical parameters, such as (but not exclusively) the disutilities for headaches and hair loss, were derived from relatively old studies. In clarification question B3, the EAG asked the company to explain if that may be related to the EAG concerns on the appropriateness of the literature searches (as discussed above) or if there was a reason these studies were considered more appropriate over more recent ones.[3] In their response, the company indicated that the impact of disutility and costs of AEs on the overall model results was small. While the EAG considers that this is correct in general, the company did not really answer the EAG's question. An important exception, as previously discussed in this report, is the long-term post-hysterectomy disutility, which, depending on its value and assumptions such as the proportion of women to whom the disutility is applicable or whether this proportion is fixed or constant over time, can have a large impact on the ICER. The reference cited as source for this disutility value is also old going back to a global WHO report from 1990.[63] Therefore, the EAG considers that this remains as a potential issue requiring further clarification, justification and exploratory analyses.

5.3.2.2 Model fit testing

In relation to model fit testing, the company did not report any validation efforts in the CS. Therefore, it is unclear if proper validation of the efficacy inputs included in the model, time to response, time to treatment discontinuation, etc. was conducted or not.

5.3.3 Validation of the computerised model (technical verification)

As explained in Section 4.2.2, the model implementation provided by the company contains some "legacy assumptions" from an original global model that includes certain functionalities that are not applicable to this NICE submission. Examples of these functionalities are the inclusion as additional comparators of BSC and surgery, the definition of BSC not including hormone therapy, input parameters related to the proportion of patients receiving each type of surgery, cost items such as blood tests and dexa scans, and possibly more. The model version received after clarification still contained these "legacy" functionalities, but it was not mentioned which ones. The EAG would suggest the company to remove all "legacy" parameters from the model to make it fit for this submission. Otherwise, the EAG would need to check the model cell by cell in order to identify them. This is extremely inefficient, slows the validation of the model down tremendously and it is practice unfeasible.

5.3.3.1 External review

No details on this type of validation efforts other than "*external ratification went into the final model and this written submission*" were reported in the CS.[1] In response to clarification question B26,[3] the company indicated that the quality control and model review was performed by a second modeller, but it is not clear whether this modeller was involved in the initial development of the model or not. An error in the calculation of life years (LYs) was identified by the EAG. The company corrected this error after clarification.

5.3.3.2 Extreme value testing

No details about quality-control procedures for code verification were provided by the company in the CS. Therefore, it was unclear whether extreme value or other types of testing were performed on the model. These could have been conducted following the guidance of the Technical Verification (TECH-VER) tool for example.[85] In response to clarification question B26,[3] the company mentioned that a formal technical quality control protocol was conducted at a late stage during the development of the model. Specific examples of black-box tests were provided in the company's answer.

5.3.3.3 Testing of traces

Markov traces can be found in the model sheets named "Markov Comp1" and "Markov Comp3". The model includes standard checks to test that the distribution of patients across health states always add up to 100%. No discussion about the face validity of the traces was provided by the company.

5.3.3.4 Unit testing

It is unknown whether code verification included checks of the model results, calculations, data references, model interface, or Visual Basic for Applications code.

5.3.4 Operational validation (validation of model outcomes)

5.3.4.1 Face validity testing (model outcomes)

Although it is not explicitly mentioned in the CS, the EAG assumed that model results were presented to experts who provided some sort of validation. This was confirmed by the company in response to clarification question B26.[3] However, the company only provided the following example: "*Clinical experts also concluded that patients may undergo a maximum of 2 conservative surgeries during their lifetime, which was reflected in the CE model with an average number of conservative surgeries of 0.5 for Relugolix CT, 0.7 for BSC, and 0.8 for GnRH agonists*". The EAG appreciated the additional clarification, but it seems insufficient to consider the model outcomes as properly validated. In addition, the EAG is uncertain about how to interpret this example since BSC is not a comparator in the model, but a subsequent treatment after discontinuation (therefore, it should be part of either the relugolix CT or the GnRH agonists arm).

Furthermore, as discussed in Section 4.2.6.1, the EAG noticed that the model produces what could be considered counterintuitive results. One might expect that the CE of relugolix CT would increase with the proportion of patients achieving complete response (the more response, the better). However, this is not the case. By increasing the proportion of patients achieving complete response to relugolix CT, the ICER increases (modestly). This result makes sense because the increase in incremental costs due to higher response rates is relatively larger than the increase in incremental QALYs. Results become counterintuitive, according to the EAG, when the proportion of patients achieving complete response to relugolix CT decreases. In the extreme scenario where the proportion of patients achieving complete response in the relugolix CT arm is assumed to be only 1%, relugolix CT still produces 0.011 incremental QALYs compared to GnRH agonists, possibly due to the assumed long-term effect (note that no waning in effect is assumed). However, because the 99% of patients discontinue treatment with relugolix CT (due to no response), this scenario results in relugolix CT saving costs compared to GnRH agonists. Thus, when its response rate is 1%, relugolix CT is dominant compared to GnRH agonists. The EAG would suggest the company to explore this further since in principle this does not seem to be valid results.

5.3.4.2 Cross validation testing (model outcomes)

Comparisons with other technology appraisals

As mentioned above (validation of the conceptual model), this submission represents the first NICE assessment for interventions treating pain associated with endometriosis. Therefore, company indicated that there are no previous NICE appraisals that can be used for cross-validation.

Comparisons with other models (not necessarily technology appraisals)

As mentioned above (validation of the conceptual model), it is unclear whether other CE models like those identified in Section 4.1, could have been used to cross-validate at least some outcomes of the model.

5.3.4.3 Validation against outcomes using alternative input data

This type of validation was not explicitly reported by the company unless it was considered part of the scenario analyses.

5.3.4.4 Validation against empirical data

Comparison with empirical data used to develop the economic model (dependent validation)

This type of validation was not reported by the company.

Comparison with empirical data not used to develop the economic model (independent validation)

This type of validation was not reported by the company.

6. EVIDENCE ASSESSMENT GROUP'S ADDITIONAL ANALYSES

6.1 Exploratory and sensitivity analyses undertaken by the EAG

6.1.1 Explanation of the company adjustments after the request for clarification

Following the clarification questions from the EAG,[3] the company made the following amendments to the originally submitted CE model:

- Clarification question B6: introduction of a post-menopause health state and model extension to a lifetime time horizon (base-case change). The utility and costs associated to the post-menopause health state are assumed to be 1 and £0, respectively. The utility values are adjusted by an age-related utility decrement, as mentioned below. The post-menopause health state costs are set to £0 based on the assumption that no direct medical costs related to endometriosis are assumed to be incurred after menopause. The model also includes the option of running a scenario to explore the impact of removing surgery for patients close to the age of menopause.
- Clarification question B11: including the odds ratio from the ITC to inform relative treatment effect of GnRH-agonists versus relugolix CT has been added to the model (base-case change).
- Clarification question B20: including age-related utility decrements (base-case change).[56]
- Clarification question B21: including a functionality to input individual utility values for each health state.
- Clarification question B29: correction of an error in the calculation of LYs (base-case change).

After the changes made by the company, the updated base-case ICER was £1,715 per QALY gained, due to a small increase in the incremental costs and a small decrease in the incremental QALYs. Therefore, the effect of all these changes on the original base-case results was minor. The company also presented the updated results of the scenario analyses but given the impact on the ICERs, these are not presented here.

In a second step of amendments, the company corrected several issues relating to the implementation of parameter uncertainty in the PSA and one-way sensitivity analysis (OWSA), as discussed in Section 5.2.1 of this report. However, these changes had no effect on the base-case results.

6.1.2 Explanation of the EAG adjustments

Table 6.1 summarises the key issues related to the CE categorised according to the sources of uncertainty as defined by Grimm et al. 2020:[86]

- Transparency (e.g., lack of clarity in presentation, description, or justification).
- Methods (e.g., violation of best research practices, existing guidelines, or the reference case).
- Imprecision (e.g., particularly wide CIs, small sample sizes, or immaturity of data).
- Bias and indirectness (e.g., there is a mismatch between the DP and evidence used to inform it in terms of population, intervention/comparator and/or outcomes considered).
- Unavailability (e.g., lack of data or insight).

Identifying the source of uncertainty can help determine what course of action can be taken (i.e., whether additional clarifications, evidence and/or analyses might help to resolve the key issue). Moreover, Table 6.1 lists suggested alternative approaches, expected effects on the CE, as well as additional evidence or analyses that might help to resolve the key issues.

Based on all considerations in the preceding Sections of this report (and summarised in Table 6.1), the EAG deemed unfeasible to define a new base-case. Assessing most of the uncertainties identified in the

CEA would require major changes to the economic model (e.g., inclusion of missing comparators or an alternative operationalisation of infertility), which cannot be conducted with the current evidence. The EAG would suggest the company to consider the key issues presented in Table 6.1. Resolving these issues should help the company defining a new base-case that could be appropriate for the current decision problem.

6.1.3 EAG exploratory scenario analyses

No exploratory analyses were conducted by the EAG.

6.1.4 EAG subgroup analyses

No subgroup analyses were performed by the EAG.

Table 6.1: Overview of key issues identified by the EAG related to the CE

Key issue	Section	Source of uncertainty	Alternative approaches	Expected impact on ICER[a]	Resolved in EAG scenario analyses	Required additional evidence or analyses
Relevant comparators may be missing from the economic analyses	4.2.4	Transparency bias and indirectness	Include missing comparators in the model	+/-	No	Update SLR and economic model to include relevant comparators
The link between the clinical effectiveness evidence and health economic analyses should be stronger	4.2.9 5.2.1	Transparency methods bias and indirectness	Inform CE parameters using clinical effectiveness results as much as possible	+/-	No	Update evidence synthesis and economic model to include missing outcomes from the final NICE scope
The operationalisation of (in)fertility in the model needs to be carefully considered	4.2.2 4.2.3 4.2.8 5.1	Transparency bias and indirectness	Link (in)fertility concerns to treatment pathway Estimate model input parameters for this group of patients where possible When a utility decrement due to infertility is applied, consider a more recent estimate, apply it only to the proportion of women actively seeking to have a	+	No	Update model structure Targeted search for more recent estimates This seems *the* key driver of the model results, therefore, carefully address the uncertainty around implementation in the model

Key issue	Section	Source of uncertainty	Alternative approaches	Expected impact on ICER[a]	Resolved in EAG scenario analyses	Required additional evidence or analyses
			family, and make it age dependent			
The definition and role of BSC in the model should be clarified	4.2.4 4.2.6	Transparency	Provide a clear definition of BSC, placebo and analgesics, and how these are used in the model	+/-	No	Update evidence synthesis (if available) and economic model to include effectiveness estimates of BSC in the correct population (BSC *after* treatment discontinuation)
The number of relevant scenario analyses to test key modelling assumptions is insufficient	4.2.6 4.2.8	Transparency methods bias and indirectness	Explore impact of potentially relevant assumptions such as treatment effect waning, infertility (after update in operationalisation as mentioned in previous issue) or a multiplicative approach when implementing disutilities	N/A	No	Conduct additional scenario analyses
Model validation efforts are needed to improve model transparency and credibility of results	5.3.3 5.3.4	Transparency methods	Remove all "legacy" parameters and assumptions from the model	N/A	No	Update economic model

Key issue	Section	Source of uncertainty	Alternative approaches	Expected impact on ICER[a]	Resolved in EAG scenario analyses	Required additional evidence or analyses
			Explore (and explain) counterintuitive model results (e.g., relugolix CT seems to be more cost effective with lower response rates)			

[a] Likely conservative assumptions (of the intervention versus all comparators) are indicated by '-'; while '+/-' indicates that the bias introduced by the issue is unclear to the EAG and '+' indicates that the EAG believes this issue likely induces bias in favour of the intervention versus at least one comparator

BSC = best supportive care; CT = in combination with oestradiol and norethisterone acetate; EAG = Evidence Assessment Group; ICER = incremental cost-effectiveness ratio; N/A = not applicable; NICE = National Institute for Health and Care Excellence; SLR = systematic literature review

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

The EAG would like to refer to the key issues presented in Table 6.1.

6.3 EAG's preferred assumptions

As mentioned in Section 6.1.2 of this report, the EAG was unable to define a new base-case. Assessing most of the uncertainties identified in the CEAs would require major changes to the economic model, which cannot be conducted with the current evidence. The EAG would suggest the company to consider the key issues presented in Table 6.1.

6.4 Conclusions of the CE Section

Searches were conducted on 5 December 2022 to retrieve published economic models, available economic evidence including economic evaluations, costs, and resource use, as well as relevant utility data for patients with endometriosis-associated pain.[1, 5] The EAG noted that searches reported for both MEDLINE and Embase, carried the same limitations as the clinical effectiveness searches, with regard to the condition facet and use of the overly restrictive pain facet, therefore the EAG concludes that it is likely that relevant studies may have been missed. Since no CE models to address the impact of relugolix CT treatment were identified by the company, a de novo model was built. Finally, the SLR did not identify any previous NICE technology appraisals for endometriosis treatment.

The company's base-case complied with the NICE reference case. It is only unclear if the disutility values considered in the economic model are representative of the UK population.

The key issues highlighted by the EAG throughout this report (and summarised in Table 6.1) were the following:

1) Relevant comparators (as reported in the NICE scope) may be missing from the economic analyses. Therefore, the results presented by the company are likely to be invalid.
2) The link between the clinical effectiveness evidence and health economic analyses should be stronger. The ITC conducted by the company informs only one parameter of the model, and its impact on the model results is negligible. It is unclear whether the studies identified in the cost/resource use search have been used in the model or not. Also, relevant outcomes identified in the final NICE scope might be included in the model.
3) The operationalisation of (in)fertility in the model needs to be carefully considered for example by linking women (in)fertility concerns to the treatment pathway. When a utility decrement due to infertility is applied, a more recent estimate should be considered, which should be applied only to the proportion of women actively seeking to have a family (instead to all of them) and make it age dependent. This seems the key driver of the model results, therefore, the uncertainty around the implementation in the model should be carefully addressed.
4) The definition and role of BSC in the model should be clarified. This should include providing a clear definition of BSC, placebo and analgesics, and how these are used in the model.
5) The number of relevant scenario analyses to test key modelling assumptions is insufficient. The company may conduct additional scenario analyses to explore impact of potentially relevant assumptions such as treatment effect waning, infertility (after update in operationalisation as mentioned in previous issue) or a multiplicative approach when implementing disutilities.
6) Model validation efforts are needed to improve model transparency and credibility of results. The company should remove all "legacy" parameters and assumptions from the model and explore (and explain) counterintuitive model results (e.g., relugolix CT seems to be more cost effective with lower response rates)

The first concern of the EAG in this submission regarding the CE evidence relates to the unclear model description before and after clarification. There is text in Document B of the CS that does not match the model implementation. The company explained that these correspond to "legacy assumptions" from an original global model that includes certain functionalities that are not applicable to this NICE submission, such as additional comparators as BSC and surgery or the definition of BSC. However, after clarification these have not been completely corrected. Some assumptions/functionalities mentioned in the CS could not be validated by the EAG. The EAG also wonders whether the model structure might have been simplified by removing the "Waiting time before surgery" health state and allow patients to transition from BSC to surgery directly. Finally, it is not clear either if there is an upper limit on the number of times patients can "loop" over the model, for example, if there is a limit for the number of surgeries a patient can undergo.

Best supportive care is part of the modelled treatment pathway (after treatment discontinuation with relugolix CT or GnRH agonists). The role and definition of BSC is also not clear. The CS defines BSC as a treatment option that includes hormonal therapy with or without analgesics, but this is incorrect.[1] After clarification, the company confirmed that BSC consists of symptomatic treatment for pain management (NSAIDs, i.e., analgesics only) and that other hormonal treatments should be disregarded. Moreover, it seems that patients are allowed to receive BSC multiple times after being a non-responder. It is therefore unclear why patients would be re-treated with BSC when it has already failed. Also, for some input parameters, the company assumed that the modelled BSC (NSAIDs*, after failing treatment with GnRH agonists or relugolix CT*) is equivalent to placebo in the SPIRIT trials. The company explained that BSC in England and Wales for the patient population in this appraisal is symptomatic treatment for pain management, such as analgesics, which is the same as the definition of BSC in the SPIRIT trials. However, the EAG is uncertain about the validity of this assumption. Following Giudice et al. 2022,[38] it seems that placebo is not defined as NSAIDs since the use of analgesics (opioids and non-opioids) in the trials is explicitly mentioned separately and in fact one of the aims of the trials was to reduce the use of them. The company also explained that relugolix CT, GnRH agonists, BSC and surgery (conservative surgery or hysterectomy) are used in combination with analgesics. Since NSAIDs are analgesics (and it is supposed to represent BSC), it is unclear 1) what type of analgesics are used in combination with the four types of interventions included in the model, 2) how BSC should be interpreted and 3) whether BSC can be considered equivalent to placebo in the SPIRIT trials. However, it is also unclear what the impact on the CE results might be.

Some potentially relevant health economic outcomes such as fertility, hospital admissions, overall pain, recurrence of endometriosis, or complications of treatment were not (completely) included in the economic model. Of key importance is the operationalisation of fertility in the model. A utility decrement due to infertility is applied in the model to all women post hysterectomy. The EAG considers that this should only be applied to the proportion of women actively seeking to have a family. However, the EAG is uncertain how this proportion can be estimated. Since both GnRH agonists and relugolix CT are contraceptive, the company considered that the utility decrement related to infertility would have been captured within the EQ-5D values collected in the trials, given that women participating in the trials were aware of this. The EAG considers that even if this utility decrement would have been captured, it should be emphasised that GnRH agonists are only administered for a maximum of 1 year, whereas in the model women can remain on relugolix CT treatment for 16 years. Therefore, the impact of the utility decrement due to infertility would be much longer for women on relugolix CT compared to those on GnRH agonists. The company also explained that a utility benefit for recovery of fertility following discontinuation of relugolix CT was deemed too uncertain to be parameterised, but it was expected to have little impact on the CE results, given that the difference in time to regain in fertility

between the two treatments is likely to be months rather than years. The EAG is uncertain about the company's expectation given that relugolix CT and GnRH agonists may have a different timeframe. For example, if a woman stops treatment after 1 year or after 10 years (only possible with relugolix CT) it is likely that the time to regain fertility would not be the same. Also, as suggested by the NICE algorithm for diagnosing and managing endometriosis, fertility concerns for patients seem to guide the clinical pathway to some extent. This does not seem to be captured by the current model structure. The company also indicated that women who discontinued treatment due to pregnancy (or wish to get pregnant) were excluded from the estimation of the discontinuation rates since treatment with BSC or surgery are deemed as not feasible options for these patients. Based on this, the EAG is unclear whether these patients have been properly included in the model.

Another major EAG concern is related to the exclusion of relevant comparators in the economic model. This is linked to the poor quality of the searches resulted in the omission of relevant comparators such as oral peptide GnRH antagonists, off licence nasal or parenteral GnRH analogues and addback HRT, which could all be used in the UK clinical practice as per clinical expert opinion. Therefore, since the economic analyses presented by the company in this submission are lacking relevant comparators, their results are likely to be invalid.

The EAG also has several concerns regarding the implementation of treatment effectiveness in the economic model. First, the model is insensitive to the choices between the two definitions of complete response, and the choice between complete and partial response. The EAG wonders whether the model could have been simplified by excluding one of the definitions of complete response and partial response. Second, the company initially assumed equal effectiveness between relugolix CT and GnRH agonists (based on no statistical significance). This was changed after clarification and the results of the ITC were included in the economic model. It should be emphasised though that the ITC conducted by the company informs only one parameter of the model, and its impact on the model results is negligible. Third, treatment effectiveness after discontinuation should have been explained in more detail in the CS. Response to BSC and to surgery after treatment discontinuation is expected to have more impact on model results than response to GnRH agonists. How these have been implemented in the model remains unclear. Fourth, the same discontinuation rates as in the relugolix CT arm were assumed for GnRH agonists. There is uncertainty regarding this assumption but the impact on the model results is minimal. The discontinuation rates over time for relugolix CT and BSC were based on observed hazard rates. However, it is unclear why exactly the same rates were used (after checking the model the EAG can confirm that these are identical) when discontinuation seems lower for BSC/placebo. Fifth, it is not explained in the CS why data on healthcare claims from the Truven Health MarketScan Commercial Claims and Encounters Database between 2004-2013 were used to estimate the probabilities of pain recurrence after surgery instead of the SPIRIT trials. That source seems relatively old, and it is unknown whether the population and BSC in this study are the same as those used in the economic analyses. Finally, there are concerns regarding the lack of explorative analyses or discussion around treatment effect waning. Testing this assumption could be relevant in principle since relugolix CT is taken for many years. In the base-case analysis, patients on relugolix CT will continue treatment until response, discontinuation or until the age of menopause. That also indicates that response to relugolix CT treatment is assumed to be constant over time. The company concluded that the waning of relugolix CT treatment effect is captured through the discontinuation rates applied in the model, since at discontinuation, patients move from complete response to non-response. The EAG is uncertain about this since the company assumed a constant discontinuation rate after 15 months, which would imply a constant treatment effect after week 60 approximately. The EAG considers that, in the absence of long-

term follow-up data, assuming a 15-year sustained effect could be considered a strong assumption and its impact on the model results should be explored.

There are also concerns regarding the HRQoL implementation in the economic model. First, to define HSUVs across the different subsequent treatment options, the company assumed that patients achieving response would experience the same utility levels irrespective of their treatment path. Therefore, utility of response as estimated from the SPIRIT 1 and 2 trials was applied to the health states of "response" to relugolix CT or GnRH agonists, "response to BSC", "post-hysterectomy stable", and response to "post-conservative surgery". Similarly, the utility of non-response was applied to health states of "BSC", "non-response BSC", "waiting time before surgery", "post-hysterectomy recurrence", and "PCS recurrence". Regarding non-responders, the company commented that the observed increase in the utility of non-responders from the baseline utility value of 0.5838 (initial treatment) to the utility value of 0.7189 estimated from data of the SPIRIT trials suggests that symptoms in non-respondents improved substantially following treatment with relugolix CT, although this increase was not sufficient to meet the criteria of treatment response. Second, the EAG questioned the company's approach to assume the same utility value for non-responders and the same utility value for responders irrespective of their treatment path. The company has noted that this is likely a conservative approach.[1] The EAG thinks this assumption may not be reflective of patients' QoL as for example it would be counterintuitive to expect that non-response to initial treatment or BSC would lead to the same utility value as the non-responsiveness of patients following surgery. In that line of reasoning, a similar bias may also hold for using the same utility values in all response health states. Specifically, if patients achieve response after multiple lines of treatment the health state utility score could be lower than patients achieving response after one or fewer lines of treatment. In response to this issue, in the clarification phase, the company only adjusted the model to include health-state specific utility values but did not suggest or include any health-state specific values, while they commented that "*since previous treatment failures are likely to contribute negatively to a patient's quality of life, assuming the same utility for all response health states is a conservative approach which is not expected to favour patients in the relugolix CT arm versus the comparator arm*".[3] When trying to assess the uncertainty around this approach, the EAG noticed that the results are relative insensitive to changes in utility values used to define response or non-response health states. This can be explained by the fact that in base results presented by the company, 88% of the gain in QALYs from relugolix CT treatment as compared to GnRH treatment is derived from a lower level of AEs and surgery-related complications for patients in the relugolix CT arm, and about 86% of the overall QALY gains from the lower level of the long-term surgery-related complications only (post-hysterectomy disutility). The EAGs main concern relates to the implementation of the disutilities in the model. The EAG noticed that to inform disutility values due to AEs or surgical complications the company used relatively older studies. In response to this question, the company only commented that the impact of these inputs on the results is relatively negligible without providing a proper and clear justification on the appropriateness of the sources used to inform these values. The EAG thinks the company failed to provide a satisfactory response to this question and does not agree with the company's viewpoint that the impact of these parameters is low to justify an appropriate search and selection process of potential sources. As for example the long-term disutility value of 0.180 used for patients undergoing hysterectomy, is quite influential in the results of the economic analysis as explained above. However, this input was informed from the WHO report on the Global Burden of Disease which was published in 2004, [63] with the original source for this input being the Global Burden of Disease report published in 1990. Therefore, the EAG is uncertain if this value is representative of the UK population. Finally, the EAG agrees that "*disutility from infertility would only be expected to have an impact on the proportion of people actively trying to have a family*" and this is not currently distinguished in the economic analysis. The company responded that "*they agree with the EAG that this*

parameter is uncertain, but it is difficult to determine what proportion of patients with a hysterectomy would be wishing to have children, particularly as this would change as the cohort ages. The intention was to capture QoL losses additional to those of infertility, which may include feelings of a loss of femininity associated with the loss of the uterus". Furthermore, related to this issue the EAG expressed concerns that patients on relugolix CT treatment may also experience similar disutility as patients on relugolix CT treatment are not able to conceive. The company reacted that "*both GnRH and relugolix are contraceptive, and the disutility of infertility related to this would already have been captured within the trial EQ-5D values given that the women participating in the trials would have been aware of this".* The EAG does not agree with the company's answer, because a childbearing wish may vary over age (as also flagged by the company) and this may not be captured in the EQ-5D trial data. In response to clarification question B7, the company went further on in their response mentioning that "a *utility benefit for faster recovery of fertility following discontinuation of relugolix CT was considered too uncertain a parameter to include and would likely have little impact in results, given that the difference in time to regain in fertility between the two treatments is likely to be months rather than years".* The EAG is unclear around this company's statement considering that GnRH agonists are only provided for a maximum duration of 1 year compared to relugolix CT which can be administered until women's menopause, which of course may have a totally different impact on the patient's 'regain of fertility'. To summarise, the EAG thinks that the issue of infertility is quite a relevant outcome for patients with endometriosis as it may also lead to a different choice of subsequent treatments and does not think that the company has appropriately incorporated this issue in the current economic analysis. To illustrate the impact of this in the current economic model, the EAG ran a hypothetical scenario analysis in which the disutility value associated with infertility was set to half of its original value while it was assumed that this disutility value would be experienced by half of the population following hysterectomy and half of the population on the relugolix CT and GnRH arms. In this scenario the ICER of relugolix CT versus GnRH would exceed £30,000 per QALY gained. Furthermore, when setting the disutility value to zero, so disregarding any long-term disutility due to hysterectomy, the ICER for relugolix CT versus GnRH treatment would be approximately 10 times higher the ICER presented by the company in the base-case analysis.

The main EAG concern regarding costs and resource use relates to the lack of transparency and clarity due to the "legacy" assumptions and parameters from previous submissions that were not corrected in this one. Nevertheless, it should be noted that costs and resource use parameters have a minor impact on the model results.

The company's base-case results indicated that relugolix CT was both more costly and more effective than GnRH agonists. In particular, relugolix CT accrued 0.71 incremental QALYs at £1,182 additional costs. Therefore, the ICER was £1,670 per QALY gained. Results also indicate that the majority of gain in QALYs for relugolix CT (0.606 from the total 0.71) is due to the long-term disutilities after surgery (to account for the loss of fertility) applied in the model. Therefore, as previously mentioned, this is a key driver of the model results, and it is crucial to carefully address the uncertainty around its value and its implementation in the model before drawing any conclusions from this CEAs. The company's PSA results are nearly identical to the deterministic ones. All of the PSA outcomes were located on the eastern quadrants of the CE-plane, where relugolix CT is more effective than GnRH agonists. Nearly all the outcomes were in the northern quadrant where relugolix CT is also more costly. All of them were below the common thresholds of £20,000–£30,000 per QALY gained. At these thresholds, the estimated probability that relugolix CT is a cost-effective alternative to GnRH agonists was 100%. The company concluded that the little variation in incremental costs and QALYs indicates a minor impact of the parameter uncertainty on the CE results and, therefore, that the analysis is robust.

The EAG disagreed with this conclusion and suggested that the little variation might be caused because many input parameters were not included in the PSA, and for those that were included, a fixed 10% variation from the mean *for all parameters* was assumed. In clarification question B30,[3] the EAG asked the company to include uncertainty ranges around each input parameter separately. These ranges should be implemented according to the parameter's source, since this uncertainty is unlikely to be the same for all parameters. At a minimum, the fixed SE for all parameters should be removed from the model and let all parameters to have their own SE. Special attention is required to those parameters for which a non-symmetric CI is to be expected (for example, a HR). In their response, the company explained that the fixed SE was replaced as suggested where the SE was either directly reported or calculated from CIs were available. However, it was still unclear for how many parameters this was done. For all the remaining parameters a fixed SE of 10% is still assumed (even though this value can be changed in the model). The EAG considers therefore that this issue was partially addressed by the company. Also, in clarification question B31,[3] the EAG asked the company to include in the PSA those parameters that were originally excluded (for no obvious reason). The company answered that this was corrected but did not provide any overview of the parameters that were (or were not) included in the PSA nor provided any justification. The company conducted a DSA to explore the sensitivity of the model results to changes in individual parameter values. The parameters that were included in the DSA were varied by assuming a 10% deviation from the base-case value (except for the annual discount rates). The results of the DSA were presented by the company in the form of a tornado diagram showing that none of the changes resulted in an ICER above £2,000 per QALY gained. As mentioned with the PSA, the company also assumed a 10% deviation from the mean for all parameters. This was corrected after clarification. Regarding scenario analyses, the EAG considered the number of scenarios insufficient. The plausibility of the selected scenarios was not discussed. The EAG is aware that the ICERs are small in general. However, the EAG noticed that including treatment effect waning and adjusting the impact of the long-term disutility associated to loss of fertility can have a substantial impact on the ICER. Note that the changes in the model that would result in a decrease of incremental QALYs can make the ICER increase quite quickly. Therefore, it is crucial to assess the robustness of the model to changes that can decrease the estimated incremental QALYs. Finally, the EAG noticed that the model produces what could be considered counterintuitive results. One might expect that the CE of relugolix CT would increase with the proportion of patients achieving complete response (the more response, the better). However, this is not the case. By increasing the proportion of patients achieving complete response to relugolix CT, the ICER increases (modestly). This result makes sense because the increase in incremental costs due to higher response rates is relatively larger than the increase in incremental QALYs. Results become counterintuitive, according to the EAG, when the proportion of patients achieving complete response to relugolix CT decreases. In the extreme scenario where the proportion of patients achieving complete response in the relugolix CT arm is assumed to be only 1%, relugolix CT still produces 0.011 incremental QALYs compared to GnRH agonists, possibly due to the assumed long-term effect (note that no waning in effect is assumed). However, because the 99% of patients discontinue treatment with relugolix CT (due to no response), this scenario results in relugolix CT saving costs compared to GnRH agonists. Therefore, when its response rate is 1%, relugolix CT is dominant compared to GnRH agonists. The EAG would suggest the company explore this further since in principle these do not seem to be valid results.

Considering the above, the EAG was unable to define a new base-case. Assessing most of the uncertainties identified in the CEAs would require major changes to the economic model, which cannot be conducted with the current evidence. The EAG would suggest the company consider the key issues presented in Table 6.1. It should be emphasised though that the current submission deals with a comparison between a long-term intervention (relugolix CT – given for 16 years) against a short-term

intervention (GnRH agonists – given for at most 12 months). The focus of the clinical effectiveness Sections of this submission lies on estimating the relative treatment effects of relugolix CT against GnRH agonists for at most 12 months since treatment initiation. When these relative treatment effects are parameterised in the company's economic model, which has a lifetime time horizon, the comparison between relugolix CT versus GnRH agonists is largely irrelevant, since the CE results are then driven by the relative treatment effects of relugolix CT compared to the subsequent treatments after GnRH agonists discontinuation, namely BSC and surgeries. With the currently modelled clinical pathway, BSC and surgeries can be considered as the comparator options for 15 years in the economic model (all treatments are assumed to stop when women become 50 years old). The clinical effectiveness Sections of this submission did not assess how to estimate the relative treatment effects of relugolix CT versus BSC or surgery (after treatment discontinuation) through evidence synthesis. Instead, these relative treatment effects were directly sourced from the SPIRIT trials. Therefore, any changes to the clinical effectiveness Sections, including for example an updated SLR and ITC, are expected to have a minimal impact on the CE results unless the missing comparators can be applied in the long-term. With the current model structure, the EAG anticipates that the only change that can have a major impact on the model results is the operationalisation of loss of fertility (because it carries a long-term effect in the model).

7. REFERENCES

[1] Gedeon Richter UK Ltd. *Relugolix–estradiol–norethisterone acetate for treating symptoms of endometriosis [ID3982; GID-TA10873]: submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Document B - Company evidence submission*, 2023

[2] National Institute for Health and Care Excellence. *Relugolix–estradiol–norethisterone acetate for treating symptoms of endometriosis [ID3982]: Final scope*. London: NICE, 2023 [accessed 18.10.23]

[3] National Institute for Health and Care Excellence. *Relugolix–estradiol–norethisterone acetate for treating symptoms of endometriosis [ID3982]: Response to request for clarification from the ERG*, 2023 [accessed 9.11.23]

[4] NHS England. *Menopause [Internet]*, 2022 [accessed 26.9.23] Available from: https://www.nhs.uk/conditions/menopause/

[5] Gedeon Richter UK Ltd. *Relugolix–estradiol–norethisterone acetate for treating symptoms of endometriosis [ID3982; GID-TA10873]: submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Document B: Appendices C-M - Company evidence submission*, 2023

[6] Secosan C, Balulescu L, Brasoveanu S, Balint O, Pirtea P, Dorin G, et al. Endometriosis in menopause-renewed attention on a controversial disease. Diagnostics (Basel) 2020; 10(3):134

[7] Kalaitzopoulos DR, Samartzis N, Kolovos GN, Mareti E, Samartzis EP, Eberhard M, et al. Treatment of endometriosis: a review with comparison of 8 guidelines. BMC Womens Health 2021; 21(1):397

[8] McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline statement. J Clin Epidemiol 2016; 75:40-6

[9] Sampson M, McGowan J, Cogo E, Grimshaw J, Moher D, Lefebvre C. An evidence-based practice guideline for the peer review of electronic search strategies. J Clin Epidemiol 2009; 62(9):944-52

[10] National Institute for Health and Care Excellence. *Relugolix–estradiol–norethisterone acetate for treating symptoms of endometriosis [ID3982]: Clarification letter*, 2023

[11] Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. *Cochrane handbook for systematic reviews of interventions version 6.3 (updated February 2022) [Internet]*: Cochrane, 2022 [accessed 4.3.22] Available from: https://training.cochrane.org/handbook

[12] Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking reviews in health care [Internet]*. York: University of York, 2009 [accessed 23.3.11] Available from: http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm

[13] Veth VB, van de Kar MMA, Duffy JMN, van Wely M, Mijatovic V, Maas JWM. Gonadotropin-releasing hormone analogues for endometriosis. Cochrane Database Syst Rev 2023; Issue 6. Art. No.: CD014788. DOI: 10.1002/14651858.CD014788.pub2

[14] National Institute for Health and Care Excellence. *Relugolix–estradiol–norethisterone acetate for treating symptoms of endometriosis [ID3982]: Response to request for clarification from the ERG: Additional clarification questions – company response* 2023 [accessed 17.11.23]

[15] National Institute for Health and Care Excellence. *Endometriosis: diagnosis and management. NICE guideline NG73 [Internet]*. London: NICE, 2017 [accessed 26.9.23] Available from: https://www.nice.org.uk/guidance/ng73/chapter/Context

[16] Lehmann M, Lueza B, Lilja M. *Report of indirect treatment comparisons of Ryeqo® and GnRH agonists for the treatment of moderate-to-severe endometriosis (Version 1.1) [PDF provided by the company]*: Gedeon Richter UK Ltd, n.d. [accessed 9.11.23]

[17] Whiting P, Savović J, Higgins JP, Caldwell DM, Reeves BC, Shea B, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. J Clin Epidemiol 2016; 69:225-34

[18] Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011; 343

[19] Becker C, Bokor A, Heikinheimo O, Horne A, Jansen F, Kiesel L, et al. ESHRE guideline: endometriosis. Hum Reprod Open 2022; 2022(2):hoac009

[20] Murad MH, Wang Z, Chu H, Lin L. When continuous outcomes are measured using different scales: guide for meta-analysis and interpretation. BMJ 2019; 364:k4817

[21] Gedeon Richter UK Ltd. *Report of indirect treatment comparisons of Ryeqo® and GnRH agonists for the treatment of moderate-to-severe endometriosis: sensitivity analysis using Standardized Mean Differences [PDF provided by the company]*, 2023 [accessed 9.11.23]

[22] Gedeon Richter UK Ltd. *Report of indirect treatment comparisons of Ryeqo® and GnRH agonists for the treatment of moderate-to-severe endometriosis: sensitivity analysis using Mean Differences [PDF provided by the company]*, 2023 [accessed 9.11.23]

[23] Dias S, Sutton AJ, Welton NJ, Ades AE. *NICE DSU Technical Support Document 3: Heterogeneity: subgroups, meta-regression, bias and bias-adjustment*. Sheffield: Decision Support Unit, ScHARR, 2011. 76p. Available from: https://www.sheffield.ac.uk/sites/default/files/2022-02//TSD3-Heterogeneity.final-report.08.05.12.pdf

[24] Grand TS, Basarir H, Jackson LJ. The cost-effectiveness of oral contraceptives compared to 'no hormonal treatment' for endometriosis-related pain: an economic evaluation. PLoS One 2019; 14(1)

[25] Bohn JA, Bullard KA, Rodriguez MI, Ecker AM. Stepwise approach to the management of endometriosis-related dysmenorrhea: a cost-effectiveness analysis. Obstet Gynecol 2021; 138(4)

[26] Arakawa I, Momoeda M, Osuga Y, Ota I, Koga K. Cost-effectiveness of the recommended medical intervention for the treatment of dysmenorrhea and endometriosis in Japan. Cost Eff Resour Alloc 2018; 16(1):1-10

[27] Wu B, Yang Z, Tobe RG, Wang Y. Medical therapy for preventing recurrent endometriosis after conservative surgery: a cost-effectiveness analysis 2018; epub doi: 10.1111/1471-0528.14786

[28] Dai Y, Shi B, Huang X, Duan J, Qiu Y, Ha C, et al. Cost-effectiveness analysis of dienogest compared with gonadotropin-releasing hormone agonist after conservative surgery for endometriosis in China. Clin Ther 2021; 43(8):1276-84

[29] Institute for Clinical Economic Review (ICER). *Final evidence report: elagolix for treating endometriosis [Internet]*, 2018 Available from: https://icer.org/wp-content/uploads/2020/10/ICER_Elagolix_Final_Evidence_Report_080318.pdf

[30] Leonardi M, Martin E, Reid S, Blanchette G, Condous G. Deep endometriosis transvaginal ultrasound in the workup of patients with signs and symptoms of endometriosis: a cost analysis. BJOG 2019; 126(12)

[31] Sanghera S, Barton P, Bhattacharya S, Horne AW, Roberts TE. Pharmaceutical treatments to prevent recurrence of endometriosis following surgery: a model-based economic evaluation. BMJ OPEN 2016; 6:e010580

[32] Wang ST, Johnson SJ, Mitchell D, Soliman AM, Vora JB, Agarwal SK. Cost-effectiveness of elagolix versus leuprolide acetate for treating moderate-to-severe endometriosis pain in the USA. J Comp Eff Res 2019; 8(5)

[33] Myovant Sciences GmbH. *SPIRIT 1 (MVT-601-3101) Clinical Study Report. Data on file [PDF provided by the company]*, 2021

[34] Myovant Sciences GmbH. *SPIRIT 2 (MVT-601-3102) Clinical Study Report. Data on file [PDF provided by the company]*, 2021

[35] Australian Institute of Health and Welfare. *Endometriosis: emergency department presentations [Internet]*: AIHW, Australian Government, 2023 [accessed 9.11.23] Available from: https://www.aihw.gov.au/reports/chronic-disease/endometriosis-in-australia/contents/treatment-management/emergency-department-presentations

[36] Australian Institute of Health and Welfare. *Endometriosis: hospitalisations [Internet]*: AIHW, Australian Government, 2023 [accessed 9.11.23] Available from: https://www.aihw.gov.au/reports/chronic-disease/endometriosis-in-australia/contents/treatment-management/hospitalisations

[37] Gedeon Richter UK Ltd. *ID3982 relugolix CT UK adapted CE clarification model 17112023KM [noCON]: Revised model provided at clarification [Excel file]*, 2013 [accessed 17.11.23]

[38] Giudice LC, As-Sanie S, Arjona Ferreira JC, Becker CM, Abrao MS, Lessey BA, et al. Once daily oral relugolix combination therapy versus placebo in patients with endometriosis-associated pain: two replicate phase 3, randomised, double-blind, studies (SPIRIT 1 and 2). Lancet (London, England) 2022; 399(10343)

[39] European Society of Human Reproduction and Embryology (ESHRE). *Information on endometriosis: patient leaflet based on the ESHRE guideline on endometriosis [Internet]*. Strombeek-Bever, Belgium: ESHRE, 2022 Available from: https://www.eshre.eu/Guideline/Endometriosis

[40] Rafique S, Decherney AH. Medical management of endometriosis. Clin Obstet Gynecol 2017; 60(3):485-96

[41] Myovant Sciences GmbH. *SPIRIT OLE (MVT-601-3103) Clinical Study Report. Data on file [PDF provided by the company]*, 2022

[42] Soliman AM, Bonafede M, Farr AM, Castelli-Haley J, Winkel CC. Analysis of adherence, persistence, and surgery among endometriosis patients treated with leuprolide acetate plus norethindrone acetate add-back therapy. J Manag Care Spec Pharm 2016; 22(5):573-87

[43] ClinicalTrials.gov, Myovant Sciences GmbH. *SPIRIT EXTENSION: Efficacy and safety extension study of relugolix in women with endometriosis-associated pain [Full Text View]. NCT03654274. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US)*, 2023 [accessed 4.10.23] Available from: https://clinicaltrials.gov/ct2/show/NCT03654274

[44] Maresh MJ, Metcalfe MA, McPherson K, Overton C, Hall V, Hargreaves J, et al. The VALUE national hysterectomy study: description of the patients and their surgery. Bjog 2002; 109(3):302-12

[45] Chapron C, Querleu D, Bruhat MA, Madelenat P, Fernandez H, Pierre F, et al. Surgical complications of diagnostic and operative gynaecological laparoscopy: a series of 29,966 cases. Hum Reprod 1998; 13(4):867-72

[46] Soliman AM, Du EX, Yang H, Wu EQ, Haley JC. Retreatment rates among endometriosis patients undergoing hysterectomy or laparoscopy. J Womens Health (Larchmt) 2017; 26(6):644-54

[47] Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol 1997; 50(6):683-91

[48] D'Agostino RB, Sr., Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation 2008; 117(6):743-53

[49] 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

[50] Armstrong N, Vale L, Deverill M, Nabi G, McClinton S, N'Dow J, et al. Surgical treatments for men with benign prostatic enlargement: cost effectiveness study. BMJ 2009; 338:b1288

[51] Boyd KA, Jones RJ, Paul J, Birrell F, Briggs AH, Leung HYCP. Decision analytic cost-effectiveness model to compare prostate cryotherapy to androgen deprivation therapy for treatment of radiation recurrent prostate cancer. BMJ OPEN 2015; 5(10):e007925

[52] Hernández Alava M, Pudney S, Wailoo A. Estimating the relationship between EQ-5D-5L and EQ-5D-3L: results from a UK population study. Pharmacoeconomics 2023; 41(2):199-207

[53] National Institute for Health and Care Excellence. *NICE health technology evaluations: the manual. Process and methods [PMG36] [Internet]*. London: NICE, 2022 [accesed 26.9.23] Available from: https://www.nice.org.uk/process/pmg36/chapter/introduction-to-health-technology-evaluation

[54] Grundström H, Gerdle B, Alehagen S, Berterö C, Arendt-Nielsen L, Kjølhede P. Reduced pain thresholds and signs of sensitization in women with persistent pelvic pain and suspected endometriosis. Acta Obstet Gynecol Scand 2019; 98(3):327-36

[55] Oppenheimer A, Verdun S, Perot M, Du Cheyron J, Panel P, Fauconnier A. Do high-dose progestins impair sexual function in women treated for endometriosis? A prospective observational longitudinal study. Acta Obstet Gynecol Scand 2021; 100(5):850-9

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

[57] Papaioannou D, Brazier JE, Paisley S. *NICE DSU Technical Support Document 9: The identification, review and synthesis of health state utility values from the literature*. Sheffield: Decision Support Unit, ScHARR, 2011. 64p. Available from: https://www.sheffield.ac.uk/sites/default/files/2022-02/TSD9-HSUV-values_FINAL.pdf

[58] Hux M, Ng C, Ortega GL, Ferrazzi S, Goeree R. Utility values for pre-menopausal women suffering from symptomatic uterine fibroids. Expert Rev Pharmacoecon Outcomes Res 2015; 15(1):181-9

[59] Xu R, Insinga RP, Golden W, Hu XH. EuroQol (EQ-5D) health utility scores for patients with migraine. Qual Life Res 2011; 20(4):601-8

[60] Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins JCP. Health state utilities for non small cell lung cancer. Health Qual Life Outcomes 2008; 6:84

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

[62] Sculpher M, Manca A, Abbott J, Fountain J, Mason S, Garry RCP. Cost effectiveness analysis of laparoscopic hysterectomy compared with standard hysterectomy: results from a randomised trial. BMJ 2004; 328(7432):134

[63] World Health Organization. *Global burden of disease 2004 update: disability weights for diseases and conditions [PDF provided by the company]*, 2004

[64] Earnshaw SR, Klok RM, Iyer S, McDade C. Methylnaltrexone bromide for the treatment of opioid-induced constipation in patients with advanced illness--a cost-effectiveness analysis. Aliment Pharmacol Ther 2010; 31(8):911-21

[65] National Institute for Health and Care Excellence. *NICE health technology evaluations: the manual. Process and methods [PMG36] [Internet]*. London: National Institute for Health and Care Excellence, 2022 [accessed 8.2.22]. 181p. Available from: https://www.nice.org.uk/process/pmg36

[66] Armour M, Lawson K, Wood A, Smith CA, Abbott J. The cost of illness and economic burden of endometriosis and chronic pelvic pain in Australia: a national online survey. PLoS One 2019; 14(10):e0223316

[67] Sperschneider ML, Hengartner MP, Kohl-Schwartz A, Geraedts K, Rauchfuss M, Woelfler MM, et al. Does endometriosis affect professional life? A matched case-control study in Switzerland, Germany and Austria. BMJ OPEN 2019; 9(1):e019570

[68] As-Sanie S, Soliman AM, Evans K, Erpelding N, Lanier R, Katz NP. Healthcare utilization and cost burden among women with endometriosis by opioid prescription status in the first year after diagnosis: a retrospective claims database analysis. J Med Econ 2020; 23(4):371-7

[69] Chibber S, Baker R, Byrd T, Marsh E. Managing pain: trends in endometriosis-related emergency department visits using the Nationwide Emergency Department Sample (NEDS) database (2010-2014). Reprod Sci 2018; 25(Suppl 1):146A

[70] British National Formulary (BNF). *BNF: Relugolix with estradiol and norethisterone acetate (Myfembree) [Internet]*. London: BNF Publications, 2023 [accessed 26.9.23] Available from: https://bnf.nice.org.uk/drugs/relugolix-with-estradiol-and-norethisterone-acetate/

[71] British National Formulary (BNF). *BNF: Triptorelin: medicinal forms [Internet]*. London: BNF publications, 2023 [accessed 26.9.23] Available from: https://bnf.nice.org.uk/drugs/triptorelin/medicinal-forms/

[72] EndoAct Canada. *Discussion: the need for action on endometriosis in Canada [Internet]*, 2022 [accessed 26.9.23] Available from: https://endoact.ca/wp-content/uploads/2022/02/EndoActCanadaBackgroundMaterial.pdf

[73] British National Formulary (BNF). *BNF: Ibuprofen [Internet]*. London: BNF publications, 2023 [accessed 26.9.23] Available from: https://bnf.nice.org.uk/drugs/ibuprofen/

[74] British National Formulary (BNF). *BNF: Tibolone [Internet]*. London: BNF publications, 2023 [accessed 26.9.23] Available from: https://bnf.nice.org.uk/drugs/tibolone/

[75] British National Formulary (BNF). *BNF: Raloxifene hydrochloride [Internet]*. London: BNF publications, 2023 [accessed 26.9.23] Available from: https://bnf.nice.org.uk/drugs/raloxifene-hydrochloride/

[76] NHS England. *2020/21 National cost collection data publication [Internet]*, 2022 [accessed 26.9.23] Available from: https://www.england.nhs.uk/publication/2020-21-national-cost-collection-data-publication/

[77] NHS England. *2021/22 National cost collection data publication [Internet]*, 2023 [accessed 26.9.23] Available from: https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/

[78] Jones K, H W, Birch S, Castelli A, Chalkley M, Dargan A, et al. Unit Costs of Health and Social Care 2022. 2023;

[79] NHS England. *2022/23 National tariff workbook (Annex A) [Internet]*, 2022 [accessed 26.9.23] Available from: https://www.england.nhs.uk/publication/national-tariff-payment-system-documents-annexes-and-supporting-documents/

[80] Menon U, McGuire AJ, Raikou M, Ryan A, Davies SK, Burnell M, et al. The cost-effectiveness of screening for ovarian cancer: results from the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS). Br J Cancer 2017; 117(5):619-627

[81] Schneider P, McNamara S, Love-Koh J, Doran T, Gutacker N. *QALY shortfall calculator [Internet]*, 2021 [accessed 11.12.23] Available from: https://shiny.york.ac.uk/shortfall

[82] Versteegh MM, Ramos IC, Buyukkaramikli NC, Ansaripour A, Reckers-Droog VT, Brouwer WBF. Severity-adjusted probability of being cost effective. Pharmacoeconomics 2019; 37(9):1155-63

[83] Gedeon Richter UK Ltd. *ID3982 relugolix Gedeon Richter CEM 06102023KM [CON]: Model provided with company submission [Excel file]*, 2013 [accessed 9.10.23]

[84] Vemer P, Corro Ramos I, van Voorn GA, Al MJ, Feenstra TL. AdViSHE: a validation-assessment tool of health-economic models for decision makers and model users. Pharmacoeconomics 2016; 34(4):349-61

[85] Büyükkaramikli NC, Rutten-van Mölken M, Severens JL, Al M. TECH-VER: a verification checklist to reduce errors in models and improve their credibility. Pharmacoeconomics 2019; 37(11):1391-408

[86] Grimm SE, Pouwels X, Ramaekers BLT, Wijnen B, Knies S, Grutters J, et al. Development and Validation of the TRansparent Uncertainty ASsessmenT (TRUST) Tool for assessing uncertainties in health economic decision models. Pharmacoeconomics 2020; 38(2):205-216

in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

Relugolix–estradiol–norethisterone acetate for treating symptoms of endometriosis [ID3982]

EAG appendix to External Assessment Report

Produced by Kleijnen Systematic Reviews (KSR) Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University

Authors Robert Wolff, Managing Director, KSR Ltd, United Kingdom (UK)
Nigel Armstrong, Health Economics Manager, KSR Ltd, UK
Isaac Corro Ramos, Health Economics Researcher, Institute for Medical Technology Assessment (iMTA), EUR, the Netherlands (NL)

Correspondence to Robert Wolff, Kleijnen Systematic Reviews Ltd
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, YO19 6FD
United Kingdom

Date completed 21/02/2024

Source of funding: This report was commissioned by the National Institute for Health Research (NIHR) Evidence Synthesis Programme as project number STA 13/56/85.

Declared competing interests of the authors None.

Evidence Assessment Group's (EAG's) appendix

Counterintuitive model results.

Scenario 1: as discussed in Section 4.2.6.1 of the EAR, one might expect that the CE of relugolix CT would increase with the proportion of patients achieving complete response (the more response, the better). However, this is not the case. By increasing the proportion of patients achieving complete response to relugolix CT, the ICER increases (modestly). This result makes sense because the increase in incremental costs due to higher response rates is relatively larger than the increase in incremental QALYs. Results become counterintuitive, according to the EAG, when the proportion of patients achieving complete response to relugolix CT decreases. In the extreme scenario where the proportion of patients achieving complete response in the relugolix CT arm is assumed to be only 1%, relugolix CT still produces 0.011 incremental QALYs compared to GnRH agonists, possibly due to the assumed long-term effect (note that no waning in effect is assumed). However, because 99% of patients discontinue treatment with relugolix CT (due to no response), this scenario results in relugolix CT saving costs compared to GnRH agonists. Thus, when its response rate is 1%, relugolix CT is dominant compared to GnRH agonists. The EAG would suggest the company to explore this further since in principle this does not seem to be valid results.

Scenario 2: By setting the proportion of patients who discontinue from relugolix CT equal to 100% at 9 or 12 months, relugolix CT dominates (more QALYs and less costs) GnRH agonists. This seems counterintuitive since, as explained in the response to query 3, the OR = 1.1 implies that GnRH agonists are more effective than relugolix CT during the first year in the model.

Company cost effectiveness results (deterministic and probabilistic) following changes made in the clarification responses.

These are provided in the tables below. Please note that main changes with respect to the original base-case are due to the increased time horizon (in the original model it was until menopause and now it is lifetime).

Table 1: Company base-case deterministic CE results

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)
Relugolix CT	11,487	23.11	17.16				
GnRH agonists	10,280	23.10	16.46	1,207	0.01	0.70	1,715
Based on electronic model after clarification. CE = cost effectiveness; CS = company submission; CT = in combination with oestradiol and norethisterone acetate; GnRH = gonadotropin releasing hormone; ICER = incremental cost-effectiveness ratio; Inc. = incremental; LYG = life years gained; QALY = quality-adjusted life year							

Table 2: Company probabilistic CE results

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)
Relugolix CT	11,084	23.14	17.27				
GnRH agonists	10,065	23.13	16.61	1,018	0.01	0.66	1,535
Based on electronic model after clarification.							

Technologies	Total costs (£)	Total LYG	Total QALYs	Inc. Costs (£)	Inc. LYG	Inc. QALYs	ICER (£/QALY)
CE = cost effectiveness; CS = company submission; CT = in combination with oestradiol and norethisterone acetate; GnRH = gonadotropin releasing hormone; ICER = incremental cost-effectiveness ratio; Inc. = incremental; LYG = life years gained; QALY = quality-adjusted life year							

Further details on the EAG's key issue 5 about the link between clinical and economic evidence.

Regarding Key Issue 5 (in Table 1.6 of the EAR) the EAG mentioned that "*The ITC conducted by the company informs only one parameter of the model, and its impact on the model results is negligible*". The signpost to Section 5.2.1 however seems incorrect. The EAG would suggest that this is replaced by 4.2.6.1. The EAG also noticed that in Section 4.2.6.1 of the EAR, this issue is not discussed since it is only mentioned that the company used an odds ratio (OR) from the ITC of the network for OPP instead of assuming equal effectiveness.

The only parameter linking the ITC with the model is the OR for OPP (relugolix vs. GnRH agonists), which is equal to 1.1 with (0.032, 41) as 95% CrI. The EAG noticed that these values were not presented in Section 3.4 of the EAR either. The OR =1.1 implies that GnRH agonists are more effective than relugolix, but this OR is applied only for 1 year in the model (hence its minor impact on the results).

Outcomes related to dysmenorrhoea and NMPP are (implicitly) incorporated in response to treatment, but not as part of any ITC, since the response values observed in the trial were used in the model.

The EAG addressed the issue (but from a more general perspective) of linking evidence and model at the end of Section 6.4 of the EAR, in the last paragraph. The EAG concluded for example that "*any changes to the clinical effectiveness sections, including for example an updated SLR and ITC, are expected to have a minimal impact on the CE results unless the missing comparators can be applied in the long-term*".

Single Technology Appraisal

Relugolix–estradiol–norethisterone acetate for treating symptoms of endometriosis [ID3982]

EAG report – factual accuracy check and confidential information check

"Data owners may be asked to check that confidential information is correctly marked in documents created by others in the evaluation before release." (Section 5.4.9, NICE health technology evaluations: the manual).

You are asked to check the EAG report to ensure there are no factual inaccuracies or errors in the marking of confidential information contained within it. The document should act as a method of detailing any inaccuracies found and how they should be corrected.

If you do identify any factual inaccuracies or errors in the marking of confidential information, you must inform NICE by **5pm on 8 January 2024** using the below comments table.

All factual errors will be highlighted in a report and presented to the appraisal committee and will subsequently be published on the NICE website with the committee papers.

Please underline all confidential information, and information that is submitted as 'confidential' should be highlighted in turquoise and all information submitted as 'depersonalised data' in pink.

Issue 1

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Linzagolix (Theramex) is mentioned throughout the EAG report as a comparator that should have been considered in the CS. However, Linzagolix cannot be considered a comparator for the following reasons: 1. Linzagolix is not currently licensed for endometriosis in the UK (it is currently only licensed for uterine fibroids) (https://products.mhra.gov.uk/substance/?substance=LINZAGOLIX%20CHOLINE) 2. Linzagolix is not scoped by NICE for endometriosis and is currently awaiting development (https://www.nice.org.uk/guidance/awaiting-development/gid-ta11376) 3. Linzagolix is not currently available for use on the NHS in the UK for any indication, with no list price available online Locations where this inaccuracy occurs include, but may not be limited to, the following: • Page 11-12, Table 1.3, row 2 column 2 and row 5 column 2 • Page 22, section 2.3, under 'EAG comment' • Page 32, section 3.1.2, EAG comment, bullet 2 • Page 46, section 4.2.4, under 'EAG comment' • Page 91, paragraph 2	Remove all mentions of Linzagolix as a suggested comparator or that Linzagolix is used in clinical practice, along with text specifically related to the omission of available GnRH antagonist comparators where appropriate.	To accurately reflect the availability of linzagolix in the UK	Where required, wording has been amended to reflect that linzagolix is not currently available for use in the NHS of England and Wales

Issue 2

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Table 2.1, page 19: overall pelvic pain and adverse effects of treatment are missing from the list of outcomes. While these were also missing in the original submission, they were subsequently added at the clarification stage.	Please add 'overall pelvic pain' to the list of outcome measures in the clinical effectiveness section and 'adverse effects of treatment' to the list of outcome measures in the CE model	Factual accuracy, alignment with the updated information provided at clarification.	The EAG comment in Table 2.1 was amended accordingly.

Issue 3

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 21, section 2.1, under 'EAG comment': *'...relugolix CT should be used as a 2nd to 3rd line treatment for moderate to severe symptoms of fibroids as well as pain associated with endometriosis.'*	Please amend text to read: *'...relugolix CT should be used as a 2nd to 3rd line treatment for pain associated with endometriosis.'*	Relevance to this submission for endometriosis.	Not a factual inaccuracy. Wording reflects clinical opinion received by the EAG.

The company does not feel that uterine fibroids is relevant to a submission on endometriosis.			

Issue 4

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 21, section 2.1, under 'EAG comment': *'It was noted that the treatment should be carefully considered because of its unfavourable side-effect profile.'* It is unclear what intervention relugolix CT is being unfavourably compared to and no reference for this statement is given.	Please specify what intervention(s) relugolix CT is being compared to in this statement and add a reference to support this comparison.	To be accurate and valid, comparisons between interventions should be specific and have supporting references. This amendment will remove the potentially misleading 'hanging comparison' and improve the accuracy of the text.	The EAG agrees that this statement should be clarified. However, at the time of submitting the EAG response to FAC, no clarification has been received.

Issue 5

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 24, section 2.2: '*According to Tables 41 and 86 of the CS...'* The information cited as coming from Table 86 of the CS in fact comes from Table 87	Please amend '*86'* to *'87'*.	Typographical error.	Amended accordingly.

Issue 6

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 21, section 2.2: *'According to the company, a dual X-ray absorptiometry scan is recommended 1-year following relugolix CT treatment...'.* The word '*following'* is misleading in this context and should read '*after starting'.*	Please change '*following'* to '*after starting'*.	Currently, the text may be misinterpreted as referring to 1 year after the treatment period has finished as opposed to 1 year after treatment had commenced. This should be changed for clarity.	Amended accordingly.

Issue 7

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 23, Table 3.1: states that no date span for the database searches are unclear in the original CS. However, date spans for the updated searches were provided at clarification as follows: Embase: 01 April 2022 to 01 November 2022 PubMed: 01 April 2022 to 01 December 2022 Clinical trial databases/Google: 01 April 2022 to 01 December 2022	We suggest adding the date spans for the search updates to Table 3.1 and specifying in the text that date spans were unclear for the **original** searches	Alignment with the updated information provided at clarification.	Amended accordingly.

Issue 8

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 25, section 3.1.1.2.1, bullet 1: the word 'as' appears to be missing. *'alternative treatments such **[as]** non-steroidal anti-inflammatory drugs (NSAIDs) and hormonal contraceptives,'*	Please add the word *'as'* as indicated in the description.	Improve the accuracy and readability of the document.	Amended accordingly.

Issue 9

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 32, EAG comment, bullet 1: it is stated that analgesia was not listed as a comparator for the SLR searches. However, analgesics were included in the search criteria for the updated SLR, as shown in Table 3.2 of the EAG report	Please remove *'However, analgesia is not listed as a comparator'.*	Factual accuracy.	Not a factual inaccuracy. The highlighted EAG comment does not refer to the SLR searches but to the SLR inclusion criteria which are reproduced in Table 3.2.

and Appendix D (Table 91) of the CS.			

Issue 10

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 35,section 3.4.1.2: *'The nature of the data varied with mean 'score' used for SPIRIT 1 and 2 as opposed to percentage in each of five categories ranging of the Biberoglu-Behrman (B&B) score for the two comparator trials, all of which were graphically estimated.'* The B&B scale (percentages in each of the five categories) was only used for Strowitzki et al 2010; the data had to be graphically estimated. TPP data for Lang et al comprised mean scores (as per SPIRIT 1 and	Please amend to: '*The nature of the data varied with mean 'score' used for the SPIRIT studies and Lang et al 2018, as opposed to the Biberoglu-Behrman (B&B) score (percentages in each of the five categories) for Strowitzki et al 2010. Data for Strowitzki et al 2010 had to be graphically estimated.'*	To accurately describe the nature of the TPP data taken from the comparator studies	Amended accordingly.

2) and was not graphically estimated			

Issue 11

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 40: *'For the health state utilities data collected from the SPIRIT 1 and 2 trials using EQ-5D data as per NICE reference case.'*	Please amend to: *'For the health state utilities, quality-of-life was prospectively measured in the SPIRIT 1 and 2 trials using the EQ-5D measure as per the NICE reference case.'*	Improve the accuracy and readability of the document.	This was amended as suggested by the company.

Issue 12

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 40: *'For disutility values other sources were used with some uncertainty if they representative of the UK population*'	Please amend to: *'For disutility values other sources were used with some uncertainty of whether they are representative of the UK population'*	Typographical error.	This was amended as suggested by the company.

Issue 13

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 41: *'Initial response to treatment can be selected at either 3 months or 3 months.'*	Please amend to: *'Initial response to treatment can be selected at either 3 months or 6 months.'*	Typographical error.	This was amended as suggested by the company.

Issue 14

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 41 *'Therefore, 6 months was selected for the base-case. Patients may then transition to either complete response, partial response or non-response.'*	Please amend to: *'Therefore, 6 months was selected for the base-case. Patients may then transition to either complete response or non-response.'*	At 6 months, patients can only be complete responders (i.e., response on both dysmenorrhea and NMPP) ir non-responders. Partial response status is determined at 3 months and captures patients with response to *either* NMPP or dysmenorrhea, if these patients do not become complete responders by month 6, then they are	This was amended as suggested by the company.

		categorised as non-responders. There are no partial responders at 6 months thus this statement is inaccurate and should be updated to improve the accuracy and readability of the document.	

Issue 15

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 42 *'For example, the EAG could not find in the model an option to change the share between hysterectomy and oophorectomy'*	We would request that the EAG please remove this sentence from this paragraph.	Whilst we recognise that the EAG were unable to locate this option in the model, this sentence could be considered to indicate that this option is not available in the model, which is incorrect. The split between hysterectomy and oophorectomy can be changed in cell G262 of the input-sheet. Please note that	This was amended as suggested by the company.

		this split only affects the cost of surgery.	

Issue 16

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 42 *'After pain recurs after hysterectomy, the company indicated that patients transition to either BSC or undergo an additional surgery. However, in the model description BSC can only be reached from conservative surgery, not from hysterectomy.'*	Please amend to *'after pain recurs after hysterectomy, patients transit into post-hysterectomy recurrence where they are treated with BSC. Some patients subsequently undergo an oophorectomy and transit into post-hysterectomy reoperation.'*	Improve the accuracy and readability of the document.	This was amended as suggested by the company.

Issue 17

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 42 *'In the model sheet "Variable bank" there is one parameter which is supposed to change the proportion of patients treated with surgery who switch to BSC following discontinuation of treatment. This value is set to 0%, but it seems that changing this input does not change results at all.'*	We would request that the EAG please remove this sentence from this paragraph.	We believe the EAG is referring to a parameter that is obsolete as it impacts results when surgery is used as a comparator, hence the absence of change in the results when modifying this parameter.	Not a factual inacuracy. This is an example of what the EAG considers an issue: the model contains obsolete parameters and "legacy assumptions" that hamper the EAG review and validation tasks.

Issue 18

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 42 *'The EAG is unclear whether similar parameters are used to define what happens post-hysterectomy.'*	We would request that the EAG please remove this sentence from this paragraph.	Whilst we recognise that this was an area of uncertainty for the EAG, the inclusion of this sentence may be considered misleading for readers.	This was amended as suggested by the company.

		The transition from post-hysterectomy stable to post-hysterectomy recurrence (the health state in which patients get treated with BSC) is determined by the rate of pain-recurrence (Row 148 of worksheet 'input-sheet'). For patients in 'post-hysterectomy recurrence', the variable that impacts the flow into re-operation is found in cell D153 of the input-sheet.	

Issue 19

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 45: *'The company conducted an email survey to UK-based healthcare professionals, who gave guidance on to which GnRH agonists are most commonly used for the*	*Please amend to: 'The company conducted an email survey with UK-based healthcare professionals, who gave guidance as to which GnRH agonists are most commonly used for the treatment of endometriosis associated pain.'*	Typographical error.	This was amended as suggested by the company.

treatment of endometriosis associated pain.'			

Issue 20

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Table 4.4, page 47, second row under "Threshold response": **Threshold response** CR: 3 months — 47.4% CR: 3 months — 63.4% PR: 3 months — 25.8% "3 months" (highlighted in red font) should be "6 months"	Please change from "3 months" to "6 months" as described	Typographical error	This was amended as suggested by the company.

Issue 21

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 48: *'The proportion of patients switching to these subsequent treatments are summarised in Table 4.5.'*	Please amend to: *'The proportion of patients switching to these subsequent treatments is summarised in Table 4.5'.*	Typographical error	This was amended as follows: 'The **proportions** of patients switching to these subsequent treatments are summarised in Table 4.5.'

Issue 22

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Table 4.7, page 52: The table purports to show discontinuation rates for relugolix CT and BSC. In fact, it shows discontinuation rates for relugolix CT and GnRH agonists (as per Table 48 in the CS). Similarly, in the final paragraph on page 51, and	Please change "BSC" to "GnRH agonists" in the table and the text	Accuracy of reporting	Table 4.7 was intended to show indeed discontinuation rates for relugolix CT and BSC as shown in Figure 4.2. the table has been amended nevertheless to emphasise that: • "The same rates as in the relugolix

in the second bullet point below Table 4.7, the EAG states that “The discontinuation rates over time for relugolix CT and BSC are shown in Table 4.7”. This is incorrect as the table in fact shows discontinuation rates for relugolix CT and GnRH agonists.			CT arm were assumed for GnRH agonists”, as mentioned in page 52. • “It is unclear why exactly the same rates were used (after checking the model the EAG can confirm that these are identical) when based on Figure 4.2, discontinuation seems lower for BSC/placebo”, as mentioned in page 53.

Issue 23

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 54: *'This is because add-back therapy may not be given at GnRH agonists treatment initiation'*	Please amend to: *'This is because add-back therapy may not be given at GnRH agonist treatment initiation'*	Typographical error.	This was amended as suggested by the company.

Issue 24

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Table 4.17, page 61: the title of this table is given as '*Surgery-related disutilities (short –term)'* However, this table shows both short- and long-term disutility data.	Please remove '(short term)' from the title of Table 4.17	Accuracy and clarity	This was amended as suggested by the company.

Issue 25

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Table 4.18, page 62: the title of this table does not make clear that the disutilities are for *long-term* surgical complications	Please change the table title to *'Disutilities due to long-term surgical complications'*	Accuracy and clarity	This was amended as suggested by the company.

Issue 26

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 63: '*The EAG is unclear around this company's statement considering that GnRH agonists are only provided for a maximum duration of 1 year compared to relugolix CT which can be administered until women's menopause, which may have a totally different*	Please amend to: *'The EAG is unclear around this statement considering that GnRH agonists are only provided for a maximum duration of 1 year compared to relugolix CT which can be administered until women's menopause, which may have a totally different impact on the patient's 'regain of fertility'.'*	Typographical error	This was amended as suggested by the company.

impact on the patient's 'regain of fertility'.'			

Issue 27

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 63: *'To illustrate the potential impact of this, the EAG run a hypothetical scenario analysis in which the disutility value associated with infertility was set to half of its original value, while assuming that this disutility would be experienced by half of the population following hysterectomy, and half of the population in the relugolix CT and GnRH agonists arms.'*	Please amend to: ''*To illustrate the potential impact of this, the EAG ran a hypothetical scenario analysis in which the disutility value associated with infertility was set to half of its original value, while assuming that this disutility would be experienced by half of the population following hysterectomy, and half of the population in the relugolix CT and GnRH agonists arms.'*	Typographical error	This was amended as suggested by the company.

Issue 28

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Table 4.19, page 64, initial treatment row: the page number given is incorrect (this information is found on page 160 of the CS, not page 157).	Change 'page 157' to 'page 160' in this cell	Typographical error	This was amended as suggested by the company.

Issue 29

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 68, EAG comment below Table 4.22: *'Other treatments originally reported in the CS are not presented in Table 4.21'* The table number is incorrect; it should be Table 4.22	Please change *'Table 4.21'* to '*Table 4.22'* in this sentence	Typographical error	This was amended as suggested by the company.

Issue 30

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 68: *'Removing add-back therapy completely changed the costs and ICER estimates in less than £100.'*	Please amend to: *'Removing add-back therapy completely had a minimal impact of less than £100 upon the total costs and the ICER'*	More precise language which will improve the readability of the report	This was amended as suggested by the company.

Issue 31

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 69: *'It also shows that in the long-term follow-up (beyond the first three cycles) nurse is required for both relugolix CT and GnRH agonists'*	Please amend to*: 'It also shows that in the long-term follow-up (beyond the first three cycles) a nurse is required for both relugolix CT and GnRH agonists'*	Typographical error	This was amended as suggested by the company.

Issue 32

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 70: '*The company stated that there are no cost of oophorectomy available in the NHS England 2022/23 national tariff workbook.'*	Please amend to '*The company stated that there are no costs of oophorectomy available in the NHS England 2022/23 national tariff workbook.'*	Typographical error	This was amended as suggested by the company.

Issue 33

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Table 4.27, page 71, conservative surgery row: *'(assumed similar prices as laparoscopy)'*	Please amend text to *'(assumed same as laparoscopic hysterectomy)'*	Accuracy of reporting	This was amended as suggested by the company.

Issue 34

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 72: *'The assumptions about the used healthcare resources applied in the model are presented in Table 4.25.'*	Please amend to '*The assumptions underpinning healthcare resource use applied in the model are presented in Table 4.25.'*	More precise language that will imrpove readability of the report	This was amended as suggested by the company.

Issue 35

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Footnote to Table 5.4, page 75: *'Note: Costs associated to fractures and indirect costs were assumed to be £0 and they are now shown in the table.'*	Please change *'now'* to *'not'* in this text	Typographical error	This was amended as suggested by the company.

Issue 36

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 76: *'Nearly all the outcomes where in the northern quadrant were relugolix CT is also more costly.'*	Please amend to: *'Nearly all the outcomes were in the northern quadrant where relugolix CT is also more costly.'*	Typographical error	This was amended as suggested by the company.

Issue 37

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 93: *'To illustrate the impact of this in the current economic model, the EAG ran an hypothetical scenario analysis in which the disutility value associated with infertility was set to half of its original value while it was assumed that this disutility value would be experienced by half of the population following hysterectomy and half of the*	Please amend to: *'To illustrate the impact of this in the current economic model, the EAG ran a hypothetical scenario analysis in which the disutility value associated with infertility was set to half of its original value while it was assumed that this disutility value would be experienced by half of the population following hysterectomy and half of the population on the relugolix CT and GnRH arms.*	Typographical error	This was amended as suggested by the company.

population on the relugolix CT and GnRH arms.'			

Issue 38

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 93: *'Furthermore, when setting the disutility value to zero, so disregarding any long-term disutility due to hysterectomy, the ICER or relugolix CT versus GnRH treatment would be approximately 10 times higher the ICER presented by the company in the base-case analysis.*	Please amend to: *'Furthermore, when setting the disutility value to zero, so disregarding any long-term disutility due to hysterectomy, the ICER for relugolix CT versus GnRH treatment would be approximately 10 times higher the ICER presented by the company in the base-case analysis.*	Typographical error	This was amended as suggested by the company.

Issue 39

Description of problem	Description of proposed amendment	Justification for amendment	EAG comment
Page 95: *'Therefore, any changes to the clinical effectiveness Sections, including for example an updated SLR and ITC, are expected to have a minimal impact on the CE results if unless the missing comparators can be applied in the long-term.'*	Please amend to: *'Therefore, any changes to the clinical effectiveness Sections, including for example an updated SLR and ITC, are expected to have a minimal impact on the CE results unless the missing comparators can be applied in the long-term.*	Typographical error	This was amended as suggested by the company.