Single Technology Appraisal

Fezolinetant for treating moderate to severe vasomotor symptoms caused by menopause [ID5071]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Fezolinetant for treating moderate to severe vasomotor symptoms caused by menopause [ID5071]

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 Astellas Pharma Ltd:**
   a. Full submission
   b. Summary of Information for Patients (SIP)

2. **Clarification questions and company responses**

3. **Patient group, professional group, and NHS organisation submissions** from:
   a. British Menopause Society

4. **Expert personal perspectives** from:
   a. Professor Waljit Dhillo – clinical expert, nominated by the Society for Endocrinology

5. **External Assessment Report** prepared by York Health Economics Consortium

6. **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

Single technology appraisal

Fezolinetant for treating vasomotor symptoms associated with the menopause [ID5071]

Document B

Company evidence submission

November 2024

File name	Version	Contains confidential information	Date
ID5071_Fezolinetant_NICE_Document B_Final_2.0 [REDACTED].docx	2.0	Yes	22nd November 2024

Instructions for companies

This is the template for submission of evidence to the National Institute for Health and Care Excellence (NICE) as part of the single technology appraisal (STA) process. Please note that the information requirements for submissions are summarised in this template; full details of the requirements for pharmaceuticals and devices are in the user guide.

This submission must not be longer than 150 pages, excluding appendices and the pages covered by this template. If it is too long it will not be accepted.

Companies making evidence submissions to NICE should also refer to the NICE health technology evaluation guidance development manual.

In this template any information that should be provided in an appendix is listed in a box.

Highlighting in the template (excluding the contents list)

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.

Grey highlighted text in the footer does not work as an automatic form field, but serves the same purpose – as prompt text to show where you need to fill in relevant details. Replace the text highlighted in [grey] in the header and footer with appropriate text. (To change the header and footer, double click over the header or footer text. Double click back in the main body text when you have finished.)

Contents

List of tables

List of figures

Abbreviations

Abbreviation	Definition
AESI	Adverse event of special interest
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
ATC	Anatomical therapeutic chemical
BLN	Baseline
BMI	Body mass index
BMS	British Menopause Society
BNF	British National Formulary
CBT	Cognitive behavioural therapy
CEM	Cost-effectiveness model
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CNS	Central nervous system
COMMA	Core outcomes in menopause
COS	Core outcome set
CSR	Clinical study report
DMC	Data Monitoring Committee
DSA	Deterministic sensitivity analysis
eCRF	Electronic case report form
EC	European Commission
ED	Early discontinuation
EMA	European Medicines Agency
EOT	End of treatment
(e)PRO	(Electronic) patient reported outcomes
EQ-5D-5L	EuroQoL five dimensions five level version
ER	Emergency room
FAS	Full analysis set
FDA	US Food and Drug Administration
FMP	Final menstrual period
FSH	Follicle-stimulating hormone
GEE	Generalised estimating equation
GPs	General practitioners
HCP	Healthcare professional
HCRU	Healthcare cost and resource use
HFS	Hot flash score
HRT	Hormone replacement therapy
HSUV	Health state utility values
HTA	Health technology assessment

ICER	Incremental cost effectiveness ratio
ICGP	Irish College of General Practitioners
INHB	Incremental net health benefit
INN	International non-proprietary name
INR	International normalised ratio
IPHA	Irish Pharmaceutical Healthcare Association
IRT	Interactive response technology
ISE	Integrated summary of effectiveness
ITT	Intention-to-treat
IUD	Intrauterine devices
KNDy	Kisspeptin, neurokinin B, and dynorphin
LH	Luteinising hormone
LS	Least squares
LY	Life year
LYG	Life year gain
MAR	Missing at random
Max	Maximum
MENQOL	Menopause-specific quality of life
MHRA	Medicines Healthcare Regulatory Agency
Min	Minimum
MMRM	Mixed models repeated measures analysis of covariance
N/A	Not applicable
NAMS	North American Menopause Society
NCPE	National Centre for Pharmacoeconomics
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHDI	National Institute for Health and Disability Insurance
NK3	Neurokinin-3
NK3R	Neurokinin-3 receptor
NKB	Neurokinin B
NMIC	National Medicines Information Centre
NR	Not reported
OECD	Organisation for Economic Co-operation and Development
OMA	Office for Market Access
ONS	Office for National Statistics
OR	Odds ratio
PAF	Pricing application form
PBAC	Pharmaceutical Benefits Advisory Committee
PGI-C VMS	Patient global impression of change for vasomotor symptoms
PHQ-4	Patient health questionnaire for anxiety and depression
PPS	Per protocol set
PROMIS SD SF 8b	Patient-reported outcomes measurement information system sleep disturbance – short form 8b

PROMIS SRI SF 8a	Patient-reported outcomes measurement information system sleep related impairment – short form 8a
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 control trial
RDI	Relative dose intensity
SAE	Serious adverse event
SAF	Safety analysis set
SD	Standard deviation
SE	Standard error
SEE	Structured expert elicitation
SLR	Systematic literature review
SNRI	Serotonin–norepinephrine reuptake inhibitor
SSRI	Selective serotonin reuptake inhibitor
STA	Single technology appraisal
STEER	Structured expert elicitation resources
SWAN	Study of women's health across the nation
TBL	Total bilirubin
TEAE	Treatment-emergent adverse event.
TVU	Transvaginal ultrasound
UK	United Kingdom
ULN	Upper limit of normal
US	United States
VAS	Visual analogue scale
VAT	Value added tax
VBA	Visual basic for application
VMS	Vasomotor symptoms
VTE	Venous thromboembolism
WHI	Women's health initiative
WPAI-VMS	Work productivity and activity impairment questionnaire specific to VMS
WTP	Willingness to pay
ZIN	The National Health Care Institute (Zorginstituut Nederland)

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

A note on terminology

It should be noted that while the vast majority of those experiencing menopause will be biological females who identify as women, Astellas acknowledges that there may be a proportion of this population who do not identify as women. As such, and in line with NICE guidelines on inclusive language, the term 'menopausal people' is used throughout to refer to the target population in this submission, with specific exceptions made where the relevant data or wording explicitly include women only.

B.1.1 Decision problem

Fezolinetant is an oral, first-in-class, non-hormonal selective neurokinin 3 (NK3) receptor antagonist that specifically targets moderate to severe vasomotor symptoms (VMS) associated with menopause (herein also referred to as vasomotor-predominant menopausal symptoms).[1]

Fezolinetant was licensed by the Medicines and Healthcare products Regulatory Agency (MHRA) on 14th December 2023 based on substantial evidence from phase 3 randomised controlled trials (RCTs) demonstrating its efficacy in statistically significantly reducing the frequency and severity of moderate to severe VMS while achieving a safety profile comparable to placebo (detailed in Section B.2).[2] These treatment benefits translated to clinically meaningful improvements in sleep, daily functioning, mental health and quality of life.

While other non-hormonal pharmacological alternatives may be used to treat VMS, including selective serotonin reuptake inhibitors (SSRIs; fluoxetine, paroxetine and citalopram), serotonin and norepinephrine reuptake inhibitors (SNRIs; venlafaxine), anticonvulsants (such as gabapentin and pregabalin) and non-pharmacological treatments such as cognitive behavioural therapy (CBT), the use of these therapies are sporadic, and often off-label.[3] Clonidine is the only other non-hormonal therapy with license wording that permits its use in "*the management of vasomotor conditions commonly associated with the menopause and characterised by flushing*",[4] however it is not recommended by the North American Menopause Society due to significant adverse events (AEs), being less effective than other therapies and having no recent studies showing greater benefit than placebo.[5] Clinical guidance in the UK (NG23) also does not recommend the use of clonidine as a front-line therapy to manage VMS alone.[3] This lack of recommendation also extends to the use of SSRIs and/or SNRIs, based on a lack of robust phase 3 evidence of treatment efficacy in moderate to severe VMS.

These concerns regarding the efficacy and safety profile of non-hormonal treatment options were reiterated by members of the British Menopause Society (BMS) during the draft scope consultation for fezolinetant, who stressed the unmet need for a routinely reimbursed (i.e., through the National Health Service [NHS]) non-hormonal treatment option, that is both more efficacious and tolerable than the non-licensed non-hormonal therapies that may currently be prescribed to menopausal people for the treatment of VMS.[6]

Fezolinetant therefore represents the only non-hormonal (NK3 antagonism) treatment option that is both licensed and has been clinically demonstrated to significantly reduce the frequency and severity of moderate to severe VMS.

In keeping with the licensed indication for fezolinetant,[2] the final scope considers menopausal people with moderate to severe VMS, as the target population relevant to this appraisal. However, fezolinetant is not expected to displace hormone replacement therapy (HRT) as the first-line therapy for treating moderate to severe VMS. As such, Astellas is seeking reimbursement in a population narrower than the licensed indication, namely menopausal people with moderate to severe VMS for whom HRT is not deemed suitable (HRT-unsuitable) for medical reasons, who may be further subdivided into:

- **HRT-contraindicated:** Menopausal people for whom HRT is contraindicated, including due to venous thromboembolism, cardiovascular disease, metabolic syndrome, severe hypertension, uncontrolled/complex diabetes mellitus, porphyria, etc.
- **HRT-caution:** Menopausal people for whom medical risk assessment of the specific caution has concluded that the risk of HRT outweighs the likely benefit.
- **HRT-stoppers:** Menopausal people who have previously received HRT but no longer take HRT.

Following advice from NICE during the scoping consultation, menopausal people who are HRT-averse i.e., those who made an informed choice not to take HRT following a consultation on the benefits and risks of HRT, have been excluded from the targeted population relevant to the decision problem, as it is expected that HRT remains the relevant comparator for these people.

UK clinical expert opinion received by Astellas strongly indicated that HRT remains the treatment of choice in the first-line setting for people with moderate to severe VMS because the benefits of HRT extend far beyond alleviating VMS alone.[7] HRT also addresses other menopausal issues, such as the genitourinary syndrome of menopause,[8, 9] and offers protective effects for cardiovascular health, bone density and cognitive function.[10, 11] Consequently, UK clinical experts expect that, in NHS primary care, fezolinetant will mainly be used in people with moderate to severe VMS for whom HRT is deemed unsuitable for medical reasons, due to the current unmet need for licensed, effective, and tolerable treatment options in UK clinical practice.

Similar feedback was received from members of the BMS during the scoping consultation stage, where it was noted that the current treatment landscape in the management of VMS for HRT-unsuitable menopausal people was unsatisfactory, and that effective non-hormonal treatment options were needed to manage VMS in this population. It was further noted that the need for safe and effective non-hormonal therapies extended beyond those who were deemed unsuitable due to a history of breast cancer, and would also include those with other hormone dependant cancers and those who have a history of significant side effects from taking HRT.[6] As such, Astellas seeks to optimise reimbursement in this subpopulation with the greatest unmet need.

The decision problem addressed in this submission is presented in Table 1, with any differences (and rationale) with the NICE final scope noted.

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	People with moderate to severe VMS associated with the menopause.	Menopausal people with moderate to severe vasomotor-predominant symptoms for whom HRT is deemed unsuitable for medical reasons: • HRT-contraindicated • HRT-caution • HRT-stoppers	UK clinical experts strongly indicated that HRT remains the treatment of choice in the first-line setting for menopausal people with moderate to severe VMS because the benefits of HRT extend far beyond alleviating VMS alone.[7, 9] Consequently, UK clinical experts expect that, in NHS primary care, fezolinetant will mainly be used in menopausal people with moderate to severe VMS for whom HRT is deemed unsuitable for medical reasons, due to the current unmet need for licensed, effective, and tolerable treatment options in clinical practice. As such, Astellas seeks to optimise reimbursement in this subpopulation. Menopausal people who are HRT-averse have been excluded from the targeted population, as it is expected that HRT remains the relevant comparator for these people.
Intervention	Fezolinetant	Fezolinetant	N/A – in line with the NICE final scope.
Comparator(s)	People for whom HRT is considered suitable: • Hormonal pharmaceutical treatments (such as oestrogen and progestogen combination, or oestrogen alone) People for whom HRT is not considered suitable: • No pharmacological treatment • Non-hormonal pharmacological treatments, for example:	Menopausal people for whom HRT is not deemed suitable for medical reasons: • No pharmacological treatment	Fezolinetant is a first-in-class, non-hormonal, pharmacological therapy, developed specifically to treat moderate to severe VMS. Current NICE clinical guidelines on the management of menopause (NG23) states, *"Do not routinely offer SSRI, SNRIs or clonidine as first-line treatment for VMS alone".*[3] This statement reflects the lack of robust phase 3 trial evidence for these treatments in moderate to severe VMS, which are not licensed for VMS in the UK (except for clonidine).

	○ Anti-depressants, such as SSRIs and SNRIs ○ Clonidine ○ Anti-convulsants, such as gabapentin and pregabalin • Non-pharmacological treatments such as CBT		UK clinical experts consulted by Astellas, including those consulted as part of NICE's early engagement process through the Office for Market Access (OMA),[7] highlighted the limited efficacy and unpleasant side effects of current non-hormonal treatments, such as SSRIs and SNRIs. These treatments are also not specifically targeted for VMS and are therefore less favourable compared with fezolinetant. Additionally, many patients who cannot take HRT, may also not tolerate the use of SSRIs or gabapentin. It is therefore expected that these non-hormonal therapies would be used in a later line compared with fezolinetant in the UK, if it is reimbursed. Additionally, UK clinical experts consulted as part of the NICE OMA meeting and the Draft Scoping Workshop indicated that psychological therapies and CBT do not represent relevant comparators in UK clinical practice, given that these treatments are typically used as add-on therapies. Given the above considerations, Astellas consider that no pharmacological treatment represents the most relevant comparator for this appraisal; a position which has been further validated, and agreed upon, with three UK clinical experts.[7, 12]
Outcomes	The outcome measures to be considered include: • Frequency of VMS • Severity of VMS • Sleep disturbance • Psychological symptoms (anxiety, low mood)	The outcome measures to be considered include: • Frequency of VMS • Severity of VMS • Sleep disturbance • Psychological symptoms (anxiety, low mood)	N/A – in line with the NICE final scope.

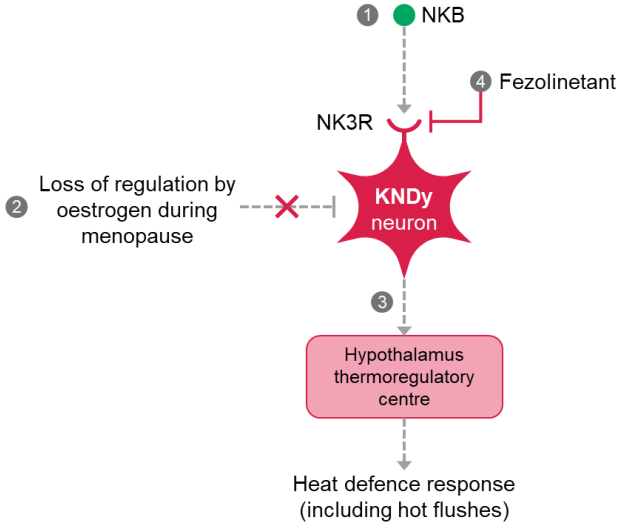
	• Adverse effects of treatment • Health-related quality of life	• Adverse effects of treatment • Health-related quality of life	
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The 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 economic analysis has been conducted in line with the NICE reference case	N/A – in line with the NICE final scope.
Special considerations including issues related to equity or equality	No special considerations raised.	The following equality issues should be considered relevant to this submission: • Differences in VMS prevalence and severity between ethnicities • Differences in VMS severity between socioeconomic levels • Location-dependent disparities in access to specialist menopause clinics	Equality considerations are further discussed in Section B.1.4.

Abbreviations: CBT: cognitive behavioural therapy; EQ-5D: EuroQol-5 Dimensions; HRT: hormone replacement therapy; N/A: not applicable; NHS: National Health Service; NICE: National Institute for Health and Care Excellence; OMA: Office for Market Access; PSS: personal social services; QALY: quality-adjusted life year; SNRI: serotonin and norepinephrine reuptake inhibitor; SSRI: Selective serotonin reuptake inhibitor; UK: United Kingdom; VMS: vasomotor symptoms.

B.1.2 Description of the technology being evaluated

A description of the technology being appraised, fezolinetant, is presented in Table 2.

Table 2: Technology being appraised

UK approved name and brand name	Fezolinetant (Veoza™)
Mechanism of action	Fezolinetant is an oral, NK3R antagonist that blocks neurokinin B (NKB) binding at the kisspeptin, neurokinin B, and dynorphin (KNDy) neuron, moderating its activity on the thermoregulatory centre of the brain to reduce the frequency and severity of VMS.[1] Typically, KNDy neuron activity is negatively regulated by oestrogen, but this control is lost during menopause due to declining oestrogen levels.[1] Consequently, KNDy neurons overstimulate the temperature control centre in the hypothalamus, resulting in hot flushes.[1] Fezolinetant interferes with NKB signalling, moderating KNDy neuron activity on the temperature control centre to reduce the occurrence of VMS.[13] **Figure 1. Fezolinetant mechanism of action**  **Abbreviations:** KNDy: kisspeptin/NKB/dynorphin; NK3R: neurokinin 3 receptor; NKB: neurokinin B. **Source:** Depypere *et al.,* (2021).[13]
Marketing authorisation/CE mark status	Marketing authorisation for fezolinetant was granted by the Medicines and Healthcare Products Regulatory Agency (MHRA) on 14th December 2023. In coming to its decision, MHRA relied on a European Commission (EC) decision on 7th December 2023 (EMEA/H/C/005851), in accordance with the advice from the Committee for Medicinal Products for Human Use (CHMP).

Indications and any restriction(s) as described in the SmPC	Fezolinetant is licensed for the treatment of moderate to severe VMS associated with menopause.[2]
	Contraindications:[2]
	• Hypersensitivity to the active substance or to any of the following excipients:
	  ○ Mannitol (E421)
	  ○ Hydroxypropyl cellulose (E463)

	○ Low-substituted hydroxypropyl cellulose (E463a) ○ Microcrystalline cellulose (E460) ○ Magnesium stearate (E470b) ○ Hypromellose (E464) ○ Talc (E553b) ○ Macrogol (E1521) ○ Titanium dioxide (E171) ○ Iron oxide red (E172) • Concomitant use of moderate or strong CYP1A2 inhibitors • Known or suspected pregnancy
Method of administration and dosage	Fezolinetant 45 mg is administered orally, once daily, at about the same time each day with or without food and taken with liquids. Tablets are to be swallowed whole and not broken, crushed, or chewed.
Additional tests or investigations	No additional tests or investigations are required.
List price and average cost of a course of treatment	List price per pack: £44.80 (28 tablets per pack).[14]
Patient access scheme (if applicable)	N/A – fezolinetant is intended for use in NHS primary care in line with UK clinical expert opinion and responses received from external stakeholders during the draft scope consultation.[6, 7]

Abbreviations: EC: European Commission; CHMP: Committee for Medicinal Products for Human Use; KNDy: kisspeptin/NKB/dynorphin; MHRA: Medicines and Healthcare Products Regulatory Agency; N/A: not applicable; NHS: National Health Service; NK3R: neurokinin 3 receptor; NKB: neurokinin B; SmPC: summary of product characteristics.

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

Overview of VMS

- Menopause (cessation of menstruation) is a natural biological process which marks the natural decline of reproductive capacity. The mean age of natural menopause in the UK is 51 years, with 1% of people experiencing menopause before 40.[15]
- VMS (hot flushes and night sweats) are the most common and bothersome symptoms, occurring in up to 80% of postmenopausal people, of which approximately 25–38% have moderate to severe symptoms requiring treatment and 25% report a severe impact on QoL.[16-19]
- Women with VMS have an average of 17 hot flushes and 11 night sweats per week, although some women can experience over 100 hot flushes and 50 night sweats per week.[20, 21] Considering only moderate to severe VMS in postmenopausal women, the average frequency is 10.57 episodes per day.[22]
- Although the pathophysiology of VMS is not fully understood, it likely represents an interplay between multiple central and peripheral physiological systems, and reproductive hormones are widely considered to play a key role.[23]
- Individual experience of VMS can vary based on a combination of genetic background, psychological factors, and other health conditions or treatments.[24]
- In primary care, for menopausal people aged ≥45 years, postmenopause is typically diagnosed after 12 months of amenorrhoea based solely on menstrual history, without the need for laboratory testing.[25]

Burden of VMS

- The occurrence of VMS can cause discomfort, anxiety, stress, self-consciousness, embarrassment, and sleep loss, which can ultimately have a considerable negative impact on quality of life (QoL).[20, 26, 27]
- Poor sleep quality is one of the most frequently reported areas of life affected during menopause and there is a significant association between moderate to severe hot flushes and poor sleep.[20]
- VMS have also been reported to be linked to symptoms of both anxiety and depression.[28-30]
- Work productivity is also impacted; in a UK-based survey of 4,014 women, 44% reported that their ability to work had been affected by menopausal symptoms.[30]
- VMS are also associated with a social impact; women with VMS and their partners report less intimacy, loss of closeness, arguments, and frustration, which can create considerable tension in relationships.[31]

Clinical pathway of care and unmet need

- Effective management of menopausal people with VMS is pivotal for their overall health and quality of life. The NICE clinical guideline on menopause (NG23) is the main source for recommendations on the management of VMS in the UK, based on expert opinion and low-to-moderate quality evidence.[3] Additionally, the NICE guidelines on early and locally advanced breast cancer (NG101) provides recommendations on the management of VMS in menopausal people with breast cancer.[32]
- HRT is the mainstay treatment of menopausal symptoms, offering comprehensive relief for VMS and a spectrum of menopausal challenges. Non-hormonal pharmacological treatments may serve as alternatives for menopausal people who are deemed HRT-unsuitable.[3, 32] However, there is an unmet need for rapid and sustained symptom relief. Non-hormonal pharmacological alternatives like SSRIs, SNRIs, clonidine, gabapentin and pregabalin have limitations due to variable efficacy in relieving VMS and high discontinuation rates resulting from unpleasant adverse effects.[3, 33] Additionally, these treatments do not target the neurological nor pathophysiological cause of VMS, underscoring the need for novel and VMS-targeted treatments.

- Fezolinetant is a novel non-hormonal therapy that targets the underlying neurological pathway that is linked with VMS, the inhibition of which reduces VMS frequency and severity. Pivotal phase 3 trials demonstrated efficacy and onset of relief from day one of treatment.[34]
- Anticipated for use in NHS primary care, fezolinetant holds the potential to transform the treatment pathway, offering a well-tolerated and effective option for menopausal people with moderate to severe vasomotor-predominant menopausal symptoms deemed HRT-unsuitable for medical reasons. Its unique mechanism of action positions it ahead of existing alternatives, representing a step change in the management of VMS.

B.1.3.1 Overview of VMS

The menopause is a natural biological process that usually affects people over 45 years and which marks a major health milestone that is characterised by the natural decline of reproductive capacity. Ovarian function is regulated by the gonadotrophins luteinising hormone (LH) and follicle-stimulating hormone (FSH), which are produced by the pituitary gland. Over time, the ovaries become less responsive to these gonadotrophins, leading to reduced secretion of reproductive hormones, particularly oestradiol. As a result, physiological changes occur, and menstrual periods eventually cease (amenorrhoea). Menopausal people are typically diagnosed as postmenopausal after 12 months of amenorrhoea, whilst the term 'menopause' refers specifically to the last menstrual period and is rarely given as a diagnosis.[3] Perimenopause is the period of time before the menopause when the clinical, biological and endocrinological features of approaching menopause start, such as menopausal symptoms and menstrual irregularity, and ends 12 months after the last menstrual period.[35] Perimenopause is also referred to as 'menopause transition'.

The change in hormone levels that occur during perimenopause, particularly the decline in oestrogen levels, lead to the development of acute menopausal symptoms, including VMS, sleep disturbance, depression and mood changes, and urogenital symptoms. Of these, VMS (hot flushes and night sweats) are the most common and bothersome symptoms, occurring in up to 80% of postmenopausal people, of which approximately 25–38% have moderate to severe symptoms requiring treatment and 25% report a severe impact on QoL.[16-19] Hot flushes (also referred to as hot flashes) are commonly defined as transient periods of intense heat, experienced predominantly around the head, neck, chest, and upper back, and are often accompanied by profuse sweating and flushing of the skin.[23] Many hot flushes are often followed by cold chills and accompanied also by palpitations and a sense of anxiety.[36] Night sweats are hot flushes that occur at night that often interfere with sleep.[36]

The frequency and severity of VMS vary among women and over time.[36] Symptoms are usually infrequent and mild in the early stages of perimenopause, and become more frequent and severe in the late perimenopause and early postmenopausal years.[36-38] Postmenopausal VMS occur in as many as 3 out of 4 women in the first two years after their final menstrual period, and decline slowly over the course of years.[39] As highlighted above, a considerable number of women who experience VMS have moderate to severe symptoms.[18, 38] In a real-world observational study in the UK, the annual prevalence of moderate to severe VMS was estimated to be approximately 21%.[40] Despite the considerable burden of VMS, the number of approved treatments remain limited for menopausal people with moderate to severe VMS for whom HRT is unsuitable.

Pathophysiology

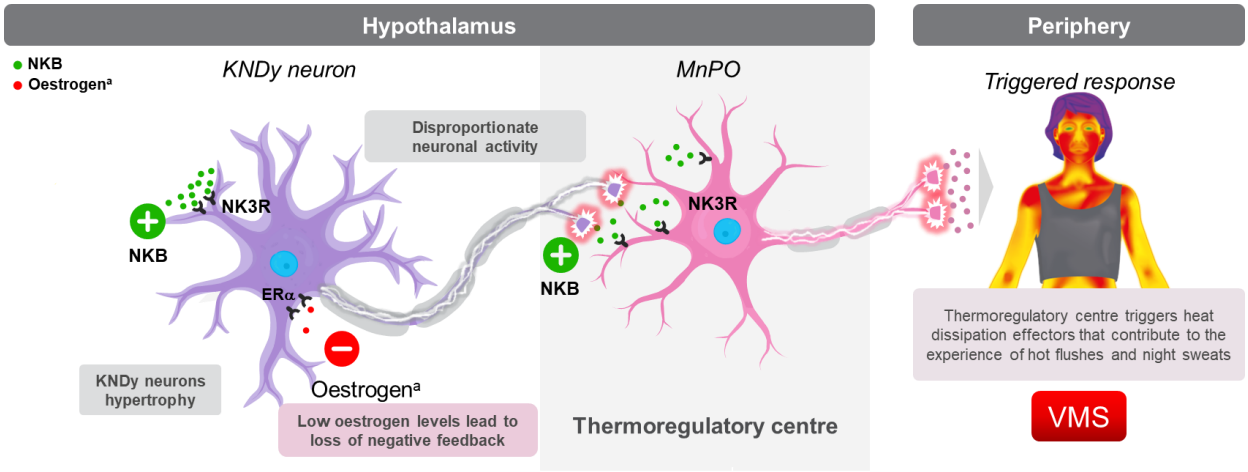
Although the pathophysiology of VMS is not fully understood, it likely represents an interplay between multiple central and peripheral physiological systems.[23] VMS are thought to occur due

to dysfunction in the temperature control centre, resulting in rapid, exaggerated, and potentially unnecessary activation of the heat dissipation response, including sweating, skin reddening, and increased heart rate.[36] Located in the hypothalamus, the temperature control centre is innervated by neurons that coordinate various functions, including thermoregulation, sex hormone levels, circadian rhythms, and sleep.[1, 41]

Reproductive hormones are widely considered to play a key role in VMS. Onset of VMS coincides with fluctuation and progressive decline of oestrogen and progesterone during perimenopause, and symptom relief is brought about by HRT with oestrogen.[23] However, although there is evidence linking circulating oestrogen levels to the occurrence of VMS, symptoms can occur both when oestrogen levels fluctuate (during perimenopause) and when they become diminished (postmenopause).[36] Several other mechanisms have also been implicated, including serotonin, norepinephrine, and opioid signalling, as well as vascular processes.[23]

Owing to the role of the hypothalamus in body temperature control, deregulation in hypothalamic signalling during the menopause is likely to contribute to VMS.[42] The thermoregulatory centre in the brain is innervated by KNDy (kisspeptin, neurokinin B [NKB] and dynorphin) neurons that are inhibited by oestrogen and stimulated by NKB. During the menopause, declining oestrogen levels disrupt this balance, leading to overstimulation of the thermoregulatory centre, resulting in VMS (Figure 2).[13] The role of NKB in VMS is supported by the observation that NKB administered to healthy premenopausal people induces hot flushes.[43] Therefore, the regulation of hypothalamic function by NK3 receptor signalling not only provides further insight into VMS pathophysiology, but also presents a potential therapeutic target for the management of VMS.

Figure 2: Pathophysiology of VMS: altered thermoregulatory activity



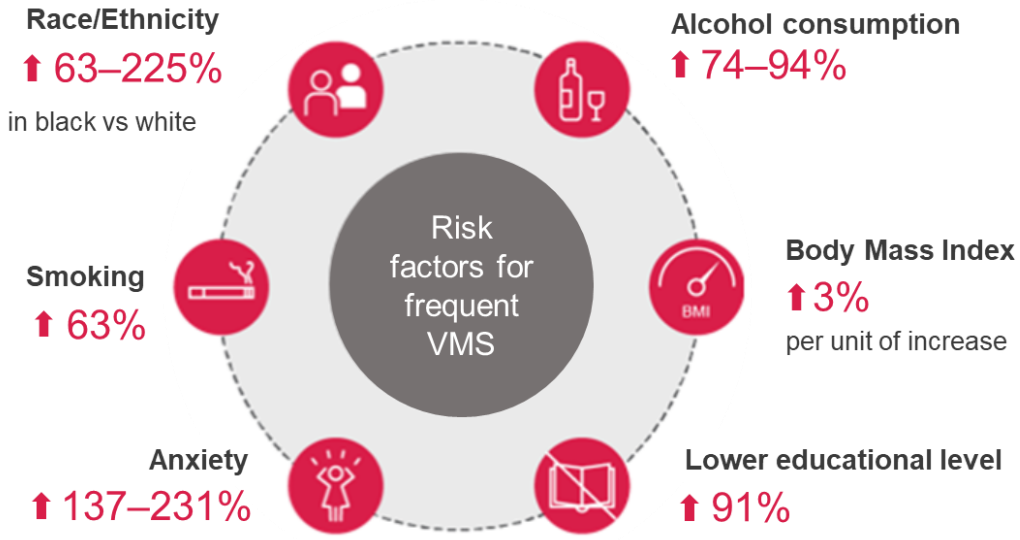
Abbreviations: ERα: oestrogen receptor alpha; KNDy: kisspeptin, neurokinin B and dynorphin; MnPO: median preoptic nucleus; NK3R: neurokinin 3 receptors; NKB: neurokinin B; VMS: vasomotor symptoms.
Source: Adapted from Depypere H *et al.*, (2021)[13]

Risk factors for VMS

VMS are multidimensional, and individual experience of VMS can vary based on a combination of genetic background, psychological factors, and other health conditions or treatments.[24] While women with VMS have an average of 17 hot flushes and 11 night sweats per week, some women can experience over 100 hot flushes and 50 night sweats per week.[20, 21] For moderate to severe vasomotor-predominant symptoms in postmenopausal people, the average frequency is 10.57 episodes per day.[22] Although there is variation in the reported frequency of VMS, there are

several documented risk factors for greater symptom frequency. These include demographic factors, health conditions, and lifestyle factors, such as smoking, obesity and alcohol consumption (Figure 3).[38, 44, 45]

Figure 3: Risk factors associated with VMS frequency



Footnote: Percentages show the increase in relative risk for VMS for each risk factor from two longitudinal studies involving 3,198 and 1,455 women. Note that published evidence for the role of alcohol consumption in the risk of VMS is mixed.
Abbreviations: VMS: vasomotor symptoms.
Sources: Gold *et al*., (2006);[38] Tepper *et al*., (2016).[44]

Similar to being a risk factor for increased VMS frequency, race and ethnicity can also impact the likelihood of people experiencing VMS, owing to differences in biological and hormonal changes in people of different races and ethnicity. In a UK study, women of African and Caribbean descent were reported as having the highest prevalence and longest duration of VMS, with symptoms being more severe.[46] They are also more likely to report problems of sleep disturbance.[46] These results are corroborated by findings from the US-based Study of Women's Health Across the Nation (SWAN), which indicated that between the different ethnicity groups, African Americans are far more likely to experience VMS, second only to Hispanic women, with 46.5% and 49.4% reporting hot flushes and/or night sweats respectively, compared with 28.9% of Caucasian women.[38] Asian women report fewer and less severe VMS than their white counterparts, although they do appear to suffer from more psychological symptoms, joint and muscle pain.[46]

Diagnosis

In the UK, NHS primary care practitioners are ideally placed to diagnose and treat people experiencing menopausal symptoms. For people aged ≥45 years, postmenopause is typically diagnosed after 12 months of amenorrhoea based solely on menstrual history, without the need for laboratory testing.[25] However, serum FSH measurements may be considered in some cases (including people aged ≥45 years with atypical symptoms, 40–45 years with menopause symptoms including a change in cycle and <40 years with a suspected diagnosis of premature ovarian insufficiency) to support diagnosis, providing no combined hormonal contraception or HRT is currently being taken.[25] In addition to VMS, other symptoms of menopause are

considered in diagnosis, including cognitive impairment and mood disorders such as anxiety or low mood, sleep disturbance and urogenital symptoms.

During the consultation with their general practitioner (GP), a personal or family medical history is routinely taken to identify potential risk factors for future disease and to assess suitability for therapeutic interventions needed.[25]

Physical examination may include recording of body mass index and blood pressure. Clinical examination of the breasts and pelvic examination are not routinely recommended but could be performed if clinically indicated.[25]

Classification of VMS

There is a lack of consistency in how VMS is classified in the literature, making it difficult to accurately estimate the burden of VMS in the real-world setting.[47] UK clinical expert opinion emphasised that VMS frequency is the most objective measure of the health effects of treatments and most relevant to menopausal people experiencing vasomotor-predominant symptoms, as it is easily definable and captured, while severity is inherently subjective and not routinely captured in UK clinical practice.[7] Patient expert opinion agreed that frequency of VMS matters most to menopausal people, as they find the recurrence of their symptoms to be more debilitating, and that severity of VMS was subjective due to the difficulty patients have in grading the severity of their VMS symptoms.[12] Some studies of VMS (as described in subsequent sections), define the severity of VMS by a certain frequency of VMS per day, or in combination with both frequency and a severity score. The European Medicines Agency (EMA) provides definitions of symptom severity based on sensation of heat, sweating, and interruption of activity (Table 3); these definitions were used in the phase 3 clinical trials for fezolinetant.[48]

Table 3: Definitions of the severity of VMS according to the EMA

Severity of VMS	Definition
Mild	Sensation of heat without sweating
Moderate	Sensation of heat with sweating, able to continue activity
Severe	Sensation of heat with sweating, causing cessation of activity

Abbreviations: EMA: European Medicines Agency; VMS: vasomotor symptoms.
Source: EMA, 2005.[48]

Duration of VMS

The overall duration of VMS is impacted by the menopausal stage during which symptoms first appear. In the SWAN study (n=1,449), it was estimated that the average total duration of frequent VMS of any severity (defined as hot flushes or night sweats occurring at least six days in the previous two weeks) from the onset of symptoms was 7.4 years (including 4.5 years after the final menstrual period).[49] Additionally, women who first report VMS at early stages of menopause (premenopause or early perimenopause) have VMS for the longest duration of time (>11.8 years) and the longest persistence after the final menstrual period (9.4 years).[49]

B.1.3.2 Burden of VMS

The occurrence of VMS can cause physical discomfort, anxiety, stress, self-consciousness, embarrassment, and sleep loss, which can ultimately have a considerable negative impact on QoL.[26, 50, 51]

In a qualitative study of women with VMS, 81% of patients reported that symptoms significantly affected their self-care.[52] Furthermore, the unpredictability of VMS can be a source of anxiety and result in the avoidance of social and leisure situations.[26, 53] Although menopause in general has been shown to impact QoL, VMS are recognised as a key contributing factor.[20, 26, 27] In a study of 2,489 women in the UK, the menopause transition did not negatively impact QoL (measured using the Short Form-36 [SF-36]), whereas VMS had a negative impact on QoL.[50] Greater symptom severity was also associated with the largest declines in functioning. For example, women experiencing severe VMS demonstrated a 3.3 point decline in physical functioning on the SF-36 scale, compared with women who did not experience VMS.[50] Individual domains of the SF-36 are scored from 0–100, with 100 signifying no restrictions, and 0 signifying total loss of function.[54] Within this scale, a change ≥3 points is considered clinically significant.[55] Compared with perimenopausal and postmenopausal women without symptoms, postmenopausal women who reported VMS experienced significant reductions in most domains of the SF-36, including physical functioning, emotional role limitation, and vitality.[50] Among postmenopausal women, VMS frequency was inversely proportional to QoL.[50]

Sleep disturbance and mental health

Poor sleep quality is one of the most frequently reported areas of life affected during menopause.[51] Women transitioning through menopause and those who are postmenopausal are 60–67% more likely to experience sleep difficulties than premenopausal women.[56, 57] As VMS frequently occur during both the day and night, VMS specifically are a key contributor to insomnia and sleep disruption in women undergoing menopause. In a cohort study of 255 women, moderate to severe hot flushes were significantly and independently associated with poor sleep.[58] The impact of VMS on sleep also increases with VMS severity. For example, a large population-based study (N=2,703) found that postmenopausal women with severe VMS (defined in this study as ≥7 hot flushes or night sweats per day of at least moderate severity) were more than three times as likely to have moderate to extreme sleep problems than women with mild or moderate VMS (adjusted odds ratio [OR]: 3.16, 95% confidence interval [CI]: 1.73, 5.80).[59] Additionally, a cross-sectional survey demonstrated a significant association (n=359; $p<0.05$) between the number of severe daytime VMS episodes and impairment in general activity domain of the Work Productivity and Activity Impairment (WPAI) questionnaire.[60] Considering the impact that VMS may have on sleep, treatments that lower VMS frequency and severity could in turn improve sleep quality, which may have associated benefits, such as increased energy, concentration and work productivity.

The menopausal transition is also associated with an impact on mental health and has been reported to be linked to symptoms of both anxiety and depression. In a survey of 4,014 women in the UK, 69% said they had experienced difficulties with anxiety or depression due to menopause.[30] Furthermore, an analysis of data from the Penn Ovarian Aging Study showed a significant 5.44-fold increase in the likelihood of experiencing depressive symptoms during menopausal transition in women with no history of depression.[61] Although the relationship between the menopause and depression has been widely debated, there is evidence to suggest that VMS plays a key role.[28, 29] Among 1,358 women in a US survey, VMS, particularly more severe VMS, were significantly associated with depressive symptoms as assessed using the Patient Health Questionnaire depression scale (PHQ-8).[62]

Socioeconomic impact

With VMS significantly affecting sleep, depression, and QoL, as well as other areas of daily life, they are likely to also have an impact on work productivity. As menopause hits many people at the peak of their careers, the societal costs of VMS can be considerable. Among 4,014 women in a UK-based survey, 44% reported that their ability to work had been affected by menopausal symptoms, and 61% reported that they had lost motivation at work.[30] In the same UK-based survey, 10% of women who had been employed during the menopause reported that they had left work due to menopausal symptoms, with a further 13% of women reporting that they had considered leaving their job.[30] Furthermore, a large cross-sectional survey of women across France, Germany, Italy, Spain, and the UK (n=3,801) reported that the severity of VMS is significantly associated with the extent to which workplace productivity is affected.[63] Despite the considerable productivity losses and associated costs due to VMS, menopause-related challenges are rarely discussed in the workplace.[64-66] This may further exacerbate the impact of VMS, as the challenges of coping with symptoms at work may lead to reduced performance, greater perceived workplace stress, and intention to leave the labour force.[67]

Furthermore, there are considerable healthcare resource costs associated with VMS, particularly considering the substantial number of menopausal people who are affected. A large cross-sectional survey (N=3,081) that included women from the UK reported a substantial burden on postmenopausal women with severe VMS, which is associated with increased healthcare resource use and costs.[63] The mean health utilities observed among these women (0.63) were lower than those observed for women in the UK general population (0.85),[68] but also among patients with other conditions including obesity, hypertension and depression, and were similar to those observed with asthma and diabetes.[69] A significant effect of VMS was also observed with respect to the number of physician visits, and the number of visits further increased with severity of VMS.[63]

Social impact

Beyond the daily discomfort caused by hot flushes, living with VMS can impact other areas of life. Problems with sleep, mood, concentration, energy and sexual activity can in turn affect relationships; women with VMS and their partners report less intimacy, loss of closeness, arguments, and frustration, which can create considerable tension in relationships.[31] In a cross-sectional survey of women experiencing moderate to severe VMS, half of all women reported that relationships with a spouse or partner had been negatively impacted by VMS and 38% reported that VMS affected their social activities.[70] Thus, not only do VMS have an immediate impact on the women experiencing them, they also have a broader impact on relationships, family, and friends.

B.1.3.3 Clinical pathway of care

Treatment pathway for VMS in the UK

The effective management of people with VMS is pivotal for their overall health and well-being during the menopause. The NICE clinical guideline on the menopause (NG23) remains the main source of recommendations on the management of VMS in the UK, which are largely based on available low-to-medium quality evidence and on expert opinion.[3, 32] Additionally, the NICE guideline on early and locally advanced breast cancer (NG101) provides recommendations on VMS management in people with breast cancer.

In NICE NG23, HRT represents the mainstay treatment of VMS, offering effective symptomatic relief that also extends beyond VMS alone.[3] HRT encompasses oestrogens and, when applicable, progestogens, effectively addressing a wide spectrum of menopausal challenges, including vaginal atrophy, mood disturbances, sleep disturbances, cognitive changes, and bone health. This comprehensive treatment effect ensures that people with VMS experience holistic symptom alleviation, enhancing their overall well-being and QoL. The effectiveness of HRT in treating a range of menopausal symptoms underscores its positioning as the first-line treatment choice for the management of short- and long-term menopausal symptoms in people with VMS for whom HRT is suitable.

Beyond HRT, NICE NG23 does not recommend the use of selective serotonin reuptake inhibitors (SSRIs; fluoxetine, paroxetine and citalopram), serotonin and norepinephrine reuptake inhibitors (SNRIs; venlafaxine), clonidine (currently the only other licensed non-hormonal pharmacological treatment for VMS), gabapentin or pregabalin as first-line treatment options for VMS alone.[3, 33] This is primarily due to the fact that these non-hormonal treatments do not specifically target VMS, have limited or conflicting evidence of efficacy in relieving VMS and are associated with high rates of discontinuation resulting from their unpleasant adverse effects.

Given the limited clinical evidence base and the associated adverse effects of SSRIs, SNRIs, clonidine, gabapentin and pregabalin, there remains an unmet need for licensed, well-tolerated non-hormonal pharmacological treatments for VMS in menopausal people for whom HRT is deemed unsuitable.

It should be noted that the NICE guidelines on the management of the menopause are currently in the process of being updated [GID-NG10241].[71] A key change between the published guidance (NG23) and the proposed guidance is the consideration of psychological interventions i.e., CBT, as a treatment option for menopausal people with vasomotor-predominant symptoms. The proposed guidance notes the evidence of CBT as a method that reduces how 'bothered' a person feels during VMS events. However, citing quality issues and uncertainty in the evidence, the Committee caveated that CBT should not be considered a routine treatment for all, but rather as an added option.[3, 71]

Positioning of fezolinetant in the treatment pathway for VMS in the UK

Fezolinetant is an oral, first-in-class, non-hormonal therapy that specifically targets and blocks NKB binding at the KNDy neuron, moderating its activity on the thermoregulatory centre of the brain to reduce the frequency and severity of moderate to severe VMS.[1] Phase 3 RCTs, detailed in Section B.2, have demonstrated the efficacy of fezolinetant in statistically significantly reducing the frequency and severity of moderate to severe VMS while achieving a safety profile comparable to placebo.[2] These treatment benefits translated to clinically meaningful improvements in sleep, daily functioning, mental health and quality of life.

Given the availability and clinical effectiveness of HRT, UK clinical experts therefore expect fezolinetant to be positioned for use in menopausal people with VMS for whom HRT is deemed unsuitable for medical reasons, due to the current unmet need for licensed, effective, and tolerable non-hormonal treatment options in UK clinical practice (Figure 4).[7, 9] This is also in line with the feedback received from members of the BMS during the draft scope consultation.[6]

With the introduction of fezolinetant, which directly targets VMS owing to its unique non-hormonal selective NK3 receptor antagonism, UK clinical experts strongly support the use of fezolinetant

ahead of SSRIs, SNRIs, clonidine, gabapentin and pregabalin in the treatment pathway for menopausal people with VMS who are deemed HRT-unsuitable for medical reasons.[7]

Clinical experts also discussed the prescription of fezolinetant following reimbursement. Notably, clinical experts stressed the importance of making fezolinetant available for prescription at the primary care level. UK clinical experts have noted that the availability of fezolinetant within the primary care setting will reduce disparities in access and may alleviate some of the pressure on specialist menopause clinics, by providing an effective treatment for people with VMS who lack any treatment options (Section B.1.4). Access to specialist menopause clinics is currently location-dependent, with some areas lacking clinics entirely, and other areas facing waiting lists of up to a year. Additionally, many specialist menopause clinics are in reality run in the primary care setting by GPs with a special interest in menopause. If the availability of fezolinetant is restricted to secondary care only, the GPs running these specialist menopause clinics will not be able to prescribe fezolinetant. These views were corroborated by members of the BMS, who cited concerns with equity in access given that fezolinetant is currently only available privately, and noted that these equity concerns would remain were fezolinetant to be recommended for prescription in secondary care only.[6]

To clarify who should be prescribed fezolinetant in UK clinical practice, Astellas supported the establishment of an independent panel of UK clinical experts with expertise in menopause (including six consultant gynaecologists and/or endocrinologists, a nurse menopause specialist, a GP menopause specialist, a consultant oncologist and one consultant pharmacist) to create a consensus statement defining the populations suitable for fezolinetant within the NHS where HRT may be deemed unsuitable. During the creation of this statement, an Astellas employee was available for troubleshooting, answering questions and providing editorial support but did not actively participate in discussions and creation of the consensus statement. The statement was drafted by the Chairs under the guidance and input of all participants. All healthcare professionals involved in this activity have confirmed they agree with the final consensus statement.[9]

The consensus statement recommends that fezolinetant is generally an option where VMS are the main concern and HRT is deemed unsuitable, in those whose comorbidities may be worsened by HRT, and where there are no contraindications to fezolinetant.[9] Notably, the consensus statement identified menopausal people who do not want to take hormones, for whom HRT does not provide adequate VMS relief, those where HRT use is contraindicated, those with a previous history of oestrogen dependent tumours (not on active treatment), those with intolerance/allergy to HRT or those with certain medical conditions where HRT would be cautioned or contraindicated (e.g., venous thromboembolism, cardiovascular disease, severe hypertension, uncontrolled/complex diabetes mellitus, porphyria etc.) as key patients eligible for fezolinetant.[9] A full list of considerations for fezolinetant eligibility is provided in Table 4.

Table 4: Considerations for fezolinetant eligibility

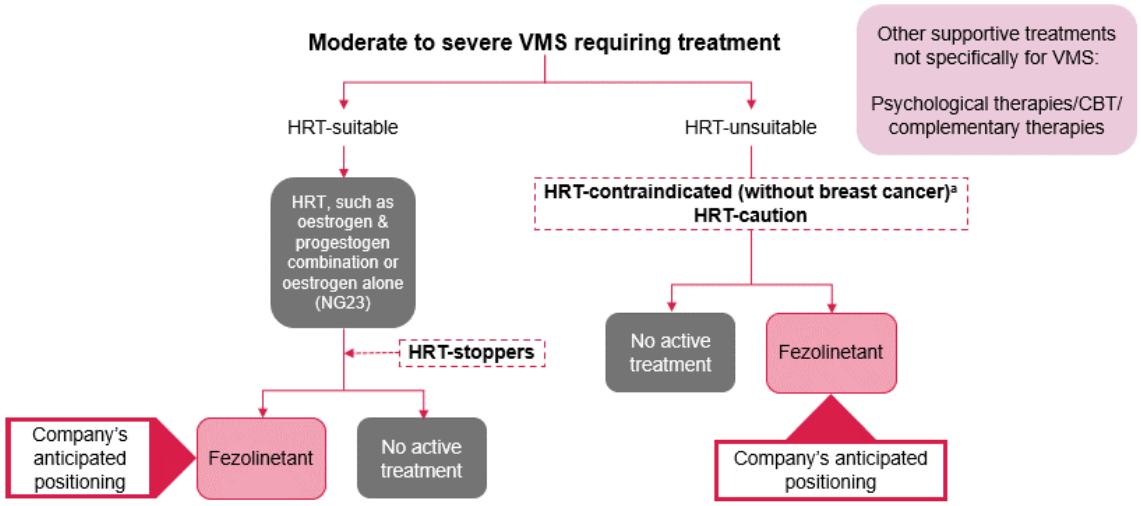
Populations eligible for fezolinetant within its product license and/or published data available	Populations eligible for fezolinetant where no data available and careful risk/benefit assessment is advised*
• People who do not want to take hormones • People for whom HRT does not provide adequate VMS relief, or who have tried different regimens with no effect, or for whom HRT has caused side effects e.g.	• Perimenopausal people who do not want HRT or who have had problems with HRT • People >65 years of age • People with very high BMI; no published

bleeding problems • People who have benign gynaecological or breast disease that may be aggravated by oestrogen e.g. endometriosis, adenomyosis, fibroids, endometrial abnormalities. • People where HRT use is contraindicated. • People who are intolerant or with known allergy to HRT	safety data exists if BMI > 38 • People who have migraines or epilepsy that may be made worse by oestrogen • People with certain medical conditions where HRT would be cautioned or contraindicated e.g. Venous thromboembolism, cardiovascular disease, severe hypertension, uncontrolled/complex diabetes mellitus, porphyria etc • People with a strong family history of breast cancer or other oestrogen sensitive malignancies or other risk factors for breast cancer e.g. premalignant lesions, atypia, prior ductal carcinoma in situ, prior chest radiotherapy • People who have had oestrogen dependent tumours, e.g. breast, ovarian, uterine • People with prior liver disease - should have liver function tests at baseline and at 3 months, 6 months and yearly follow up visits thereafter

* Shared management with secondary/tertiary/specialist care is advised. Further research recommended by panel for all these categories
Abbreviations: BMI: Body Mass Index; HRT: hormone replacement therapy.
Source: Astellas (Data on File). UK HRT-Unsuitable Consensus Statement.[9]

Given the insights from the UK clinical expert community, fezolinetant is not expected to displace HRT as the first-line therapy for treating moderate to severe VMS in menopausal people who are suitable for HRT in UK clinical practice. Rather, fezolinetant is expected to represent the first-line treatment choice in a relatively small number of menopausal people with moderate to severe VMS for whom HRT is deemed unsuitable for medical reasons, ahead of other non-hormonal treatments e.g., SSRIs/SNRIs or gabapentin. Considering the positioning of fezolinetant in this subpopulation, and the lack of recommendations for and limited use of non-hormonal therapies in this population, no active treatment is considered to be the only relevant comparator for this appraisal.

Figure 4: Anticipated treatment pathway including the proposed positioning of fezolinetant for people with VMS for whom HRT is deemed unsuitable for medical reasons



[a] Fezolinetant is not licensed for use in women with breast cancer, or those being treated for breast cancer. For those no longer receiving active oncologic treatment, a benefit/risk assessment should be made prior to considering the use of fezolinetant to treat VMS associated with menopause.[2]
Abbreviations: CBT: cognitive behaviour therapy; HRT: hormone replacement therapy; VMS: vasomotor symptoms.

B.1.4 Equality considerations

As noted in the draft scope consultation comments and the Equality Impact Assessment, there is a substantial and urgent unmet need for treatment options for menopausal people with moderate to severe VMS, as many people struggle to access effective non-hormonal treatments.[6, 72] Historically there has been a lack of innovation in this area of women's health, particularly for non-hormonal treatment options. New non-hormonal treatment options would therefore be welcome to this patient group.[72]

VMS variations between ethnicities

There are differences in the prevalence and severity of VMS between ethnicities which should be considered. A higher prevalence and longer duration of VMS has been reported in women of African and Caribbean descent, and symptoms are also more severe.[46] These women are also more likely to report problems with sleep disturbance. Findings from the SWAN study further support that VMS are more prevalent in black and Hispanic women than other ethnicities.[73, 74]

Socioeconomic differences

Differences in VMS severity between socioeconomic levels also exist. Women in a household where the main earner works in semi-skilled or unskilled manual work have been reported to be more likely to have difficulty with physical and psychological postmenopausal symptoms than those where the main earner works in a managerial role.[30] Severe symptoms were also reported to be worse for working class women, with 44% of women describing three or more symptoms as 'very difficult'.[30]

Location-dependent disparity in access

UK clinical experts highlighted location-dependent disparities in access to specialist menopause clinics and appointment waiting lists.[7] Such disparities in access would be exacerbated should fezolinetant not be available for prescription in a primary care setting.[6] In addition to this, many of these specialist menopause clinics are in reality run in the primary care setting by GPs with a special interest in menopause. If the availability of fezolinetant is restricted to secondary care only, many of these specialist menopause clinics will not be able to provide their patients access to fezolinetant.

B.2 Clinical effectiveness

Summary of clinical effectiveness of fezolinetant

Evidence for fezolinetant 45 mg in moderate to severe VMS

- The efficacy and safety of fezolinetant for the treatment of moderate to severe VMS has been evaluated in three multicentre, randomised, double-blind, placebo-controlled phase 3 trials: DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2. Supporting evidence on the long-term safety and tolerability of fezolinetant 45 mg has also been evaluated in a large, randomised placebo-controlled 52-week SKYLIGHT 4 study, results of which are presented in Appendix F.
- The DAYLIGHT trial was a phase 3b, randomised, double-blind, placebo-controlled 24-week study that assessed the efficacy and safety of fezolinetant 45 mg versus placebo, in the treatment of postmenopausal women aged ≥40 years and ≤65 years who had moderate to severe VMS and were deemed unsuitable for HRT.
- SKYLIGHT 1 and SKYLIGHT 2 were identically designed, concurrent, phase 3, randomised, placebo-controlled, 12-week double-blind, parallel group, multicentre trials investigating the efficacy and safety of fezolinetant 30 mg and 45 mg once daily in 12-week placebo-controlled period followed by a 40-week active treatment extension period.

Efficacy

- Across all three trials fezolinetant 45 mg met its primary and most of the key secondary endpoints.
  - In DAYLIGHT, treatment with fezolinetant 45 mg demonstrated a statistically significant least squares (LS) mean (standard error [SE]) reduction in daily frequency of moderate to severe VMS versus placebo at Week 24 (−1.93 [0.36], p<0.001). Additionally, a statistically significant LS mean (SE) reduction in daily severity of moderate to severe VMS was observed with fezolinetant 45 mg as compared with placebo at Week 24 (−0.39 [0.09], p<0.001).
  - In individual SKYLIGHT 1 and SKYLIGHT 2 trials, fezolinetant 45 mg demonstrated a statistically significant LS mean reduction in both frequency and severity of moderate to severe VMS versus placebo at Week 4 (SKYLIGHT 1: −2.07 [0.42], adjusted p=0.007; SKYLIGHT 2: −2.55 [0.46], adjusted p<0.001) and Week 12 (SKYLIGHT 1: −2.55 [0.43], adjusted p=0.007; SKYLIGHT 2: −2.53 [0.55], adjusted p<0.001). The efficacy of fezolinetant 45 mg in alleviating patient-reported sleep disturbance (PROMIS SD SF 8b) was demonstrated in both individual and pooled trials at Week 12.
- The benefit of fezolinetant 45 mg was observed as early as one week after the start of treatment, with continued improvement to Week 4 and sustained benefit through Week 24 for DAYLIGHT and Week 52 for SKYLIGHT 1 and SKYLIGHT 2.
- Improvements in the daily frequency of moderate to severe VMS were broadly consistent across subgroup analyses of the fezolinetant 45 mg group in all three trials. The only exception to this general observation was the subgroup analysis of current smokers in DAYLIGHT, the results of which should be interpreted with caution due to the very small sample size.
- The efficacy of fezolinetant 45 mg was also demonstrated in exploratory endpoints measuring quality of life, daily functioning and psychological symptoms (low mood and anxiety), as measured by MENQOL, EQ-5D-5L, WPAI-VMS and PHQ-4.
  - In DAYLIGHT, patients treated with fezolinetant 45 mg reported greater improvements in LS mean (SE) MENQOL (vasomotor domain) (–0.96 [0.22]), EQ-5D-5L VAS (+1.6 [1.5]), WPAI-VMS domain (overall productivity loss: −6.77 [3.21]; activity impairment: −8.38 [2.46]) and PHQ-4 (−0.44 [0.27]) scores versus placebo at Week 24.
  - In individual SKYLIGHT 1 and SKYLIGHT 2 trials, fezolinetant 45 mg demonstrated improved LS mean (SE) MENQOL (vasomotor domain) (–1.00 [0.20] and –1.05 [0.21]) and EQ-5D-5L VAS (+3.2 [1.4] and +4.2 [1.7]) scores versus placebo at Week 4. At Week 12, similar results were observed for EQ-5D-5L VAS (3.4 [1.8]) in SKYLIGHT 2 only and

MENQOL (vasomotor domain) (–0.99 [0.22] and –0.76 [0.22]) scores for both trials, respectively

- Additionally, in SKYLIGHT 1 and SKYLIGHT 2, improvements were observed in the WPAI-VMS domain scores (overall productivity loss: −10.68 [3.88] and −9.99 [3.95]; activity impairment: −9.52 [2.52] and −6.51 [2.79]) versus placebo at Week 4. Similar trends were observed at Week 12, with improvements demonstrated in WPAI-VMS domain scores for both trials (overall productivity loss: −18.08 [3.97] and −13.92 [4.16]; activity impairment: −8.97 [2.73] and −8.23 [2.85]), respectively.

Safety

- In the DAYLIGHT trial, patients receiving fezolinetant 45mg showed comparable frequencies of treatment-emergent adverse events (TEAEs) to those receiving placebo. Both groups had a low incidence of serious TEAEs, with only one serious TEAE attributed to the study drug in the fezolinetant 45 mg group. No deaths occurred in the study. Additionally, no notable differences were observed in the occurrence of specific TEAEs of special interest, including liver test elevations, uterine bleeding, endometrial hyperplasia, cancer or disordered proliferative endometrium, and thrombocytopenia. Notably, the incidence of liver enzyme elevations was low across both groups, and no patients had bilirubin levels >2 ULN. Furthermore, there were no instances of Hy's law observed in the trial.
- In both SKYLIGHT 1 and SKYLIGHT 2 trials, during the 12-week double-blind period, the incidence of TEAEs for patients on fezolinetant 45 mg was comparable to the placebo group. No deaths were reported, and serious TEAEs were infrequent and comparable between the fezolinetant 45 mg and placebo groups. None of the serious TEAEs were deemed related to the study treatment by the investigators. TEAEs leading to withdrawal of treatment were low and similar between the treatment groups in both trials. No notable differences were observed in the frequency of any TEAEs of special interest across the treatment groups in both trials. This included factors such as liver enzyme elevations, uterine bleeding, endometrial hyperplasia, cancer or disordered proliferative endometrium and thrombocytopenia. Furthermore, the incidence of liver enzyme elevations (ALT or AST levels >3 x ULN) was low, with these events generally being asymptomatic, transient, and resolving either during treatment or after treatment discontinuation. None of the patients in either trial exhibited bilirubin levels >2 ULN, and there were no cases of Hy's law observed.
- Overall, the safety profile of fezolinetant 45 mg demonstrates a low incidence of adverse events, with no concerning patterns or severe adverse effects detected.

Indirect treatment comparisons

- Given that no active treatment represents the sole comparator to fezolinetant in this submission in line with UK clinical expert opinion, it was not deemed necessary to conduct an indirect treatment comparison given that clinical evidence on the efficacy of no active treatment was sourced from the placebo arm of the DAYLIGHT trial.

Conclusions

- Fezolinetant is a first-in-class nonhormonal therapy that has been clinically demonstrated to significantly reduce both the frequency and severity of moderate to severe VMS, which translate to improved outcomes in terms of sleep, quality of life, daily functioning and mental health.
- The mechanism of fezolinetant is targeted, acting specifically on the thermoregulatory pathways that underpin VMS. Therefore, the use of fezolinetant in menopausal people with moderate to severe VMS for whom HRT is deemed unsuitable for medical reasons would represent a step change in treatment care pathway, given the current unmet need for non-hormonal alternatives to HRT and the current lack of robust, licensed and evidenced-based therapies for treating the VMS.

B.2.1 Identification and selection of relevant studies

A systematic literature review (SLR) of the published literature was conducted to identify relevant clinical evidence on the comparative efficacy and safety of active prescription treatments for postmenopausal women with moderate to severe VMS.

The original SLR, conducted on 25th June 2021, used broad inclusion criteria that included both hormonal and non-hormonal pharmaceutical therapies; however, the SLR updates conducted on 14th June 2023 (SLR Update 1) and 9th April 2024 (SLR Update 2) used narrow inclusion criteria, focussing mainly on non-hormonal therapies for moderate to severe VMS. As such, any publications with non-hormonal interventions from the original SLR were re-extracted in line with the eligibility criteria for the SLR Update 1. In total, 10 relevant publications from the original SLR were re-extracted, bringing the total number of included publications in the SLR update 1 to 27 reporting on 14 unique studies (all randomised controlled trials [RCTs]).

In SLR Update 2 (2024), 624 records were retrieved by electronic database searches, of which 524 were duplicated against the hits from SLR Update 1 (2023), resulting in 100 unique records that were screened at the title/abstract review stage. Twenty-seven publications were then screened at the full-text review stage and six publications were found to be eligible for inclusion. Two additional publications were identified through supplementary searches; one each from the congress searches and bibliography searches.

In total, the clinical SLR includes 35 publications reporting on 14 unique studies.

Full details of the SLR search strategy, study selection process and results are presented in Appendix D.

B.2.2 List of relevant clinical effectiveness evidence

Of the studies identified in the clinical SLR, the trial of direct relevance to the decision problem of this submission is DAYLIGHT (NCT05033886)[75], with supporting evidence sourced from SKYLIGHT 1 (NCT04003155)[76] and SKYLIGHT 2 (NCT04003142)[77]. Additional clinical evidence on the long-term safety and tolerability of fezolinetant 45 mg is provided by SKYLIGHT 4 for completeness, although not of direct relevance to the decision problem (see Appendix F).

DAYLIGHT was a parallel-group, double-blind, multicentre, phase 3b randomised controlled trial (RCT), investigating the efficacy and safety of fezolinetant 45 mg in women with moderate to severe VMS and deemed unsuitable for HRT – the population of relevance to the decision problem addressed in this submission.

SKYLIGHT 1 and SKYLIGHT 2 were two concurrent, identically designed, parallel-group, double-blind, multicentre, phase 3 RCTs, investigating the efficacy and safety of fezolinetant 30 mg and 45 mg once daily in 12-week placebo-controlled period followed by a 40-week active treatment extension period. Both trials were the pivotal registration trials presented to the Medicines and Healthcare products Regulatory Agency (MHRA) in support of the marketing authorisation for fezolinetant 45 mg as a treatment for moderate to severe VMS. It should be noted that only the 45 mg dose of fezolinetant is licensed in the UK; data relating to the 30 mg dose are therefore not relevant for health-economic decision-making. Baseline characteristics for patients who received fezolinetant 30 mg in the SKYLIGHT trials are included for completeness only, and any

discussions of clinical effectiveness results are presented for the patients who were treated with the licensed 45 mg dose only (see Section B.2.6.2).

An overview of DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 trials is presented in Table 5, and the methodology and results are presented in Section B.2.3 onwards. Alongside the data presented for the individual SKYLIGHT 1 and SKYLIGHT 2 trials, additional data are also presented for pooled SKYLIGHT 1 and SKYLIGHT 2, where available. The pooled analysis was pre-specified and conducted to assess the efficacy of fezolinetant during the first 12 weeks of treatment in key subgroups of women based on their baseline demography, VMS history and HRT history. Further details on the pooled analyses are reported in the pre-specified analysis report included in the reference pack with this submission.[78]

Table 5: Clinical effectiveness evidence

Study	DAYLIGHT (NCT05033886)[75]	SKYLIGHT 1 (NCT04003155)[76]	SKYLIGHT 2 (NCT04003142)[77]
Study design	Phase 3b, randomised, placebo-controlled, 24-week double-blind, parallel group, multicentre study	Phase 3, randomised, placebo-controlled, 12-week double-blind, parallel group, multicentre study, followed by a 40-week double-blind, non-controlled extension treatment period	
Population	Menopausal women aged 40 to 65 years of age with moderate to severe VMS and deemed **unsuitable for HRT**	Menopausal women aged 40 to 65 years of age with moderate to severe VMS	
Intervention(s)	Fezolinetant 45 mg QD (n=227)	• Fezolinetant 45 mg QD (n=176) • Fezolinetant 30 mg QD (n=176)	• Fezolinetant 45 mg QD (n=167) • Fezolinetant 30 mg QD (n=166)
Comparator(s)	Placebo (n=226)	Placebo (n=175)	Placebo (n=168)
Indicate if study supports application for marketing authorisation	No	Yes – marketing authorisation for fezolinetant in moderate to severe VMS was informed by both pivotal trials, with supporting evidence from SKYLIGHT 4 (see Appendix F)	
Indicate if study used in the economic model	Yes – primary source of clinical data for the model base case	Yes – supporting clinical data from the HRT-unsuitable subpopulation from pooled SKYLIGHT 1 and SKYLIGHT 2 informed the model base case (see Section B.2.7 for more details)	
Rationale if study not used in model	N/A	N/A	

Reported outcomes specified in the decision problem	• **Mean change in the frequency of moderate to severe VMS**
	• Mean change in the severity of moderate to severe VMS
	• Sleep disturbance
	  ○ Mean change in PROMIS SD SF 8b
	• Psychological symptoms and health-related quality of life
	  ○ Mean change in the MENQOL total and domain scores
		• **Mean change in the frequency of moderate to severe VMS**
		• Mean change in the severity of moderate to severe VMS
		• Sleep disturbance
		  ○ Mean change in PROMIS SD SF 8b
		• Psychological symptoms and health-related quality of life
		  ○ **Mean change on the EQ-5D-5L total score and VAS**
		  ○ Mean change in the MENQOL total and domain scores
		• Adverse effects of treatment
		  ○ Frequency and severity of adverse events
		  ○ Transvaginal ultrasound and endometrial biopsy findings	

	○ Mean change in Patient Health Questionnaire for Anxiety and Depression (PHQ-4) total and domain scores ○ **Mean change on the EQ-5D-5L total score and VAS** • Adverse effects of treatment ○ Frequency and severity of adverse events ○ Transvaginal ultrasound and endometrial biopsy findings ○ Change from baseline to each time point in vital signs: sitting systolic and diastolic blood pressure and pulse rate ○ Change from baseline to each time point in electrocardiogram parameters ○ Change from baseline to each time point in laboratory tests: haematology, biochemistry and urinalysis	○ Change from baseline to each time point in vital signs: sitting systolic and diastolic blood pressure and pulse rate ○ Change from baseline to each time point in electrocardiogram parameters ○ Change from baseline to each time point in laboratory tests: haematology, biochemistry and urinalysis
All other reported outcomes	• Mean change in WPAI-VMS domain scores	• Mean change on the WPAI-VMS domain scores • Mean change in PROMIS SRI SF 8a

Endpoints in bold are those that are used to inform the cost-effectiveness model.
Abbreviations: EQ-5D-5L: EuroQoL five dimensions five level version; HRT: hormone replacement therapy; MENQOL: Menopause Quality of Life Questionnaire; N/A: not applicable; OD: once daily; PHQ-4: Patient Health Questionnaire for Anxiety and Depression; PROMIS SD SF 8b: Patient-reported Outcomes Measurement Information System Sleep Disturbance – Short Form 8b; PROMIS SRI SF 8a: Patient-reported Outcomes Measurement Information System Sleep Related Impairment – Short Form 8a; VAS: visual analogue scale; VMS: vasomotor symptoms; WPAI-VMS: Work Productivity and Activity Impairment questionnaire specific to VMS.
Source: Astellas (Data on File). DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 Study Protocols and CSRs.[79-85]

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

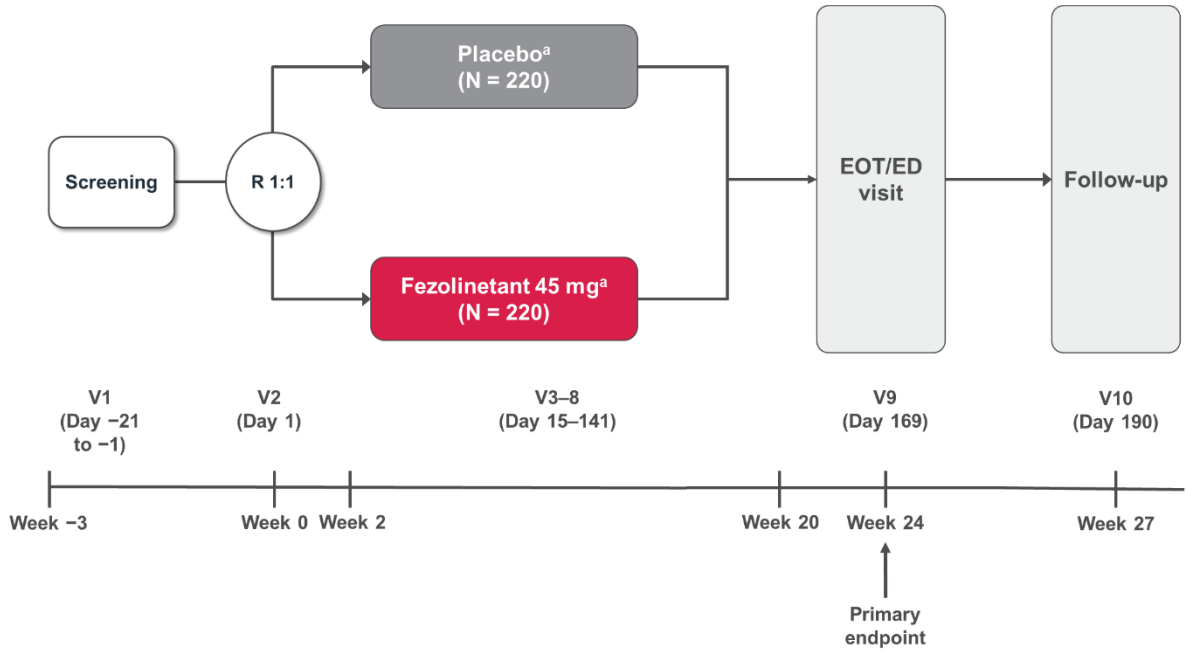
B.2.3.1 Trial design

B.2.3.1.1 DAYLIGHT

The DAYLIGHT trial was a phase 3b, randomised, double-blind, placebo-controlled 24-week study that assessed the efficacy and safety of fezolinetant 45 mg versus placebo, in the treatment of postmenopausal women aged ≥40 years and ≤65 years who had moderate to severe VMS and were deemed unsuitable for HRT.

A schematic of the DAYLIGHT trial is presented in Figure 5. The trial comprised a 21-day screening period, followed by a 24-week treatment period.

Figure 5: Study design of DAYLIGHT



For visits prior to the EOT visit, patients returned to the study site for study assessment and procedures within ±5 days of the scheduled day. At EOT visit, the visit window was ±3 days. Signed informed consent was obtained from all patients before any study-related procedures were conducted. Patients who discontinued the study treatment early completed an ED visit at the time of discontinuation and then a safety follow-up visit three weeks thereafter. Patients who discontinued study treatment early remained in the study and continued to complete the electronic daily VMS diary and ePRO assessments as scheduled through Week 24. The safety follow-up visit assessments could have been postponed to align with the next scheduled visit. Visits during the 24-week treatment period were conducted at Weeks 2, 4, 8, 12, 16, 20 and 24. Visits 3 through 8 (inclusive) could have been conducted as home healthcare visits. [a]Treatment was administered orally once daily until EOT or ED of the 24-week treatment period.
Abbreviations: ED: early discontinuation; EOT: end of treatment; ePRO: electronic patient-reported outcome; V: visit; VMS: vasomotor symptoms.
Source: Astellas (Data on File). DAYLIGHT Study Protocol.[79]

Screening occurred up to 21 days prior to randomisation. Eligibility was determined via medical history and HRT questionnaire, physical examination, clinical laboratory testing and vital signs. Within the last 10 days prior to randomisation, patients were required to have a minimum average of seven moderate to severe VMS per day. Throughout the screening period, patients

were required to document their VMS daily using an electronic diary. Study site staff reviewed the electronic diary during Visit 2 (Day 1) to ascertain study eligibility. Results of screening assessments that did not meet the parameters required by the eligibility criteria (for example, clinical laboratory tests, vital signs, physical examination and electrocardiogram) were allowed to be repeated once within the 21-day screening period, without the need to register the patient as a screen failure. In order to rescreen after prior screen failure, a new informed consent from was signed by the patient and they were entered into screening with a new patient identification number. Rescreening was only allowed once for an individual patient. Patients were not permitted to be rescreened if the e-diary data had been assessed for eligibility via e-diary vendor report.

Approximately 440 patients were planned to be enrolled and randomised in a 1:1 ratio across two treatment arms and stratified by smoking status (smoker or non-smoker) through an interactive response technology (IRT). A total of 453 patients were enrolled and randomised as follows:

- Fezolinetant 45 mg tablets once daily (n=227)
- Placebo tablets once daily (n=226)

Patients who completed the 24 weeks of treatment attended an end of treatment (EOT) visit and a safety follow-up visit three weeks thereafter. Patients who discontinued the study treatment early completed an early discontinuation (ED) visit at the time of discontinuation and then a safety follow-up visit three weeks from the last dose of study treatment. Patients who discontinued early remained in the study and continued to complete the electronic daily VMS diary and electronic patient-reported outcome (ePRO) assessments as scheduled through Week 24.

During the 24-week treatment period, patients returned to the study site for Visits 2 through 9 for scheduled assessments as per study protocol. Site-based patient-reported outcome (PRO) measures were self-administered using an electronic device. Assessments at Visit 2 (Day 1) had to occur before randomisation/first dosing; assessments at Weeks 4, 12, 16 and 24 had to occur before dosing. All self-administered assessments were performed first upon arrival at the site and before all other procedures. The VMS diary was kept by the patients. As at the screening period, patients were required to record the VMS frequency and severity of each VMS in the electronic diary each day from Week 0 through to Week 24. Night sweats were recorded no later than in the morning upon awakening to start a new day. For Visits 3 through 8, visits could have been conducted as home healthcare visits.

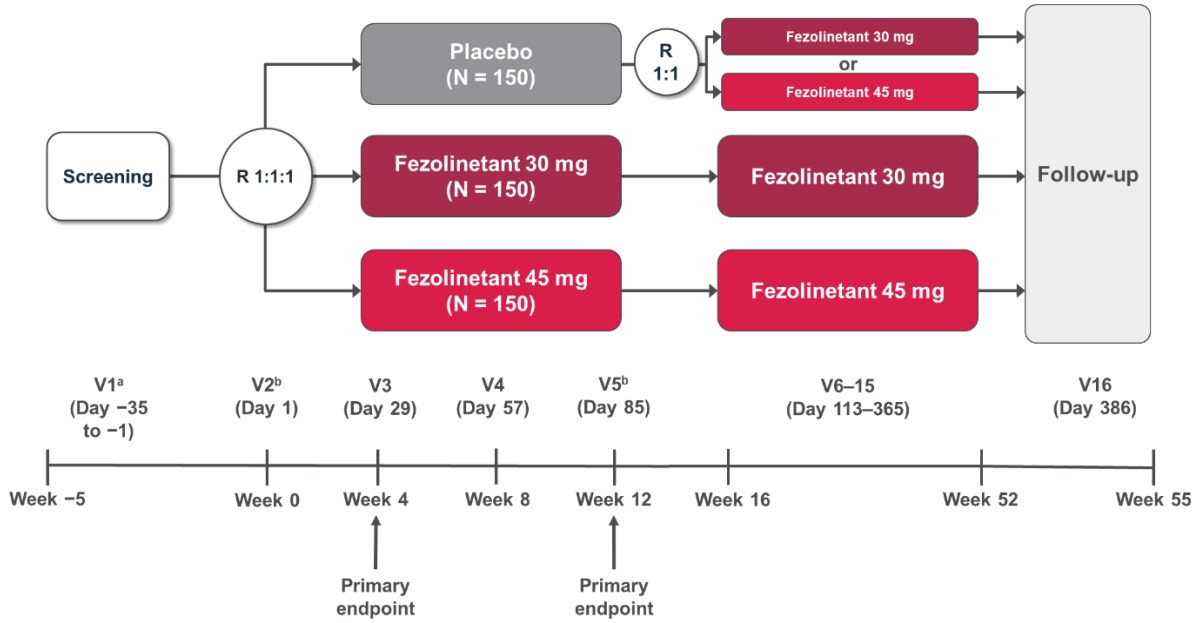
Patients were assigned study treatment as a kit in blister packaging. The first dose of the study treatment took place at the study site on Visit 2 (Day 1) under the supervision of the study staff. On study visit days, the daily dose of the study treatment was taken at the study site, under the supervision of the study staff or at home under the supervision of the home healthcare personnel, after collection of pre-dose blood samples. On all other days throughout the treatment period, patients were instructed to take their dose of the study drug at home in the morning with water. Patients were asked to return all unused study drug. Compliance with study drug intake was assessed by counting returned study drug and recorded in the source documents and the IRT. Adverse events (AEs) and intake of concomitant medication(s) were continuously monitored from informed consent until the last study-related activity.

B.2.3.1.2 SKYLIGHT 1 and SKYLIGHT 2

The SKYLIGHT 1 and SKYLIGHT 2 trials were randomised, 12-week double-blind, placebo-controlled, parallel group, multicentre phase 3 clinical studies that assessed the efficacy and safety of two fezolinetant dose regimens (30 mg or 45 mg once daily) versus placebo in women (≥40 years and ≤65 years) with moderate to severe VMS. Although these were two separate pivotal trials, they both had an identical study design and methodology and are therefore summarised together below.

A schematic of the design of the SKYLIGHT 1 and SKYLIGHT 2 trials is present in Figure 6. Both trials comprised a screening period (Days −35 to −1, including the screening visit [Visit 1] and collection of VMS frequency and severity assessments) and a 52-week treatment period (Day 1 [Visit 2] to Week 52 [Visit 15]).

Figure 6: Study design of the SKYLIGHT 1 AND SKYLIGHT 2 trials



The study was kept double-blinded throughout the 52 weeks and was conducted on an outpatient basis. Patients returned to the study site for visits and procedures that occurred within ±3 days of the scheduled day; unscheduled visits could be arranged outside the planned visits. Signed informed consent was obtained from all patients before any study-related procedures were conducted. Patients who chose not to consent to pharmacogenomic sampling were not excluded from study participation. Screening (Visit 1), randomisation, Week 0 (visit 2), and EOT/ED week 52 (visit 15) visits had to be conducted in the clinic and not with home healthcare, telemedicine conferences, or local laboratory. [a]Screening was performed up to 35 days before randomisation, with a minimum of 10 days to allow for baseline data collection of VMS frequency and severity. [b]Refer to the schedule of assessments for Visits 2b and 5b in the study protocol.
Abbreviations: ED: early discontinuation; EOT: end of treatment; V: visit.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 Study Protocols.[80, 81]

The screening visit (Visit 1), conducted up to 35 days before randomisation (Visit 2), allowed a minimum of 10 days for collecting baseline data on VMS frequency and severity. Patients were allowed one retest under the same screening number within the 35-day screening window upon approval from the medical monitor. A single instance of rescreening was also permitted, resulting in a new screening number and 35-day screening window, subject to medical monitor approval. Patients who had already undergone the VMS eligibility report were not rescreened. To be eligible, patients had to have a minimum average of 7 to 8 moderate to severe VMS a day, or 50 to 60 moderate to severe VMS per week within the 10 days prior to randomisation. Throughout

the screening period, patients were required to document their VMS daily using an electronic diary. Study site staff reviewed the electronic diary during Visit 2 to ascertain study eligibility.

The first 12 weeks of the trials were randomised, double-blinded and placebo-controlled. Approximately 450 patients were planned to be enrolled and randomised in a 1:1:1 ratio across three treatment arms according to randomisation schedules and stratified by smoking status (active smoker or non-smoker) through the IRT.

A total of 527 and 501 patients for SKYLIGHT 1 and SKYLIGHT 2 were enrolled and randomised as follows:

- Fezolinetant 30 mg tablets once daily (SKYLIGHT 1: n=176; SKYLIGHT 2: n=166)
- Fezolinetant 45 mg tablets once daily (SKYLIGHT 1: n=176; SKYLIGHT 2: n=167)
- Placebo tablets once daily (SKYLIGHT 1: n=175; SKYLIGHT 2: n=168)

After completing 12 weeks of treatment, patients on placebo were re-randomised in a 1:1 ratio to either fezolinetant 30 mg or 45 mg once daily for a further 40 weeks (that is, 52-week treatment period; hereafter referred to as the “52-week period”) in the active treatment-extension period (that is, extension period without placebo control) using the IRT. Patients who completed treatment with fezolinetant 30 mg or 45 mg in the 12-week double-blind period continued to receive the same allocated treatment for the remaining 40 weeks during the active treatment-extension period. Following the completion (or ED) of the 52-week treatment period (Week 52), patients completed an EOT or ED visit and final safety follow-up visit three weeks after the last dose of study treatment was administered (Week 55).

During the 52-week treatment period, patients returned to the study site for Visits 2 through 15 for scheduled assessments as per study protocol. Site-based patient-reported outcome (PRO) measures were self-administered using an electronic device. Assessments at Visit 2 had to occur before randomisation/first dosing; assessments at Weeks 4, 12, 24 and 52 had to occur before dosing. All self-administered assessments were performed first upon arrival at the site and before all other procedures. In the event a patient withdrew from the study, efforts to collect information on the site-based PRO measures were made before or shortly after discontinuation. The VMS diary was kept by patients. As at the screening period, patients were required to record the VMS frequency and severity of each VMS in the electronic diary each day from Week 0 through to the follow-up visit (Week 55 [Visit 16]). Night sweats were recorded no later than in the morning upon awakening to start a new day.

Patients were assigned study treatment as a kit in blister packaging. The first dose of the study treatment took place at the study site on Visit 2 (Day 1) under the supervision of the study staff. On study visit days, the daily dose of the study treatment was taken at the study site, under the supervision of the study staff, after collection of pre-dose blood samples. On all other days throughout the treatment period, patients were instructed to take their dose of the study drug at home in the morning with water. Patients were asked to return all unused study drug. Compliance with study drug intake was assessed by counting returned study drug and recorded in the source documents and the interactive response technology. Adverse events (AEs) and intake of concomitant medication(s) were continuously monitored from informed consent until the last study-related activity.

Modifications to the study conduct due to COVID-19

All attempts were to be made to conduct the protocol-defined scheduled visits. In cases where a patient was unable to visit the clinic due to site closure related to the COVID-19 pandemic, alternative measures could be implemented to ensure patient safety and continuity of care while participating in the study:

- Telemedicine Conferences (telephone visits) to evaluate changes in a patient's medical condition or medications and completion of electronic PRO questionnaires
- Safety laboratory tests collected at a local lab to include biochemistry, haematology, liver biochemistry and coagulation panel testing
- Patients were instructed to continue entering their VMS in the electronic diary as instructed regardless of site closure
- A COVID-19 login workaround for electronic clinical outcomes assessment device tablets was implemented to be available in the event of site closure, patients were unable to come in for their visit, or the inability of the site to share the device with the patient
- Home healthcare services could be available in cases where arrangements were made in advance by the site upon request from the patient(s)

Due to the nature of the screening (Visit 1), randomisation, Week 0 (Visit 2) and EOT/ED Week 52 (Visit 15) visits, which included important study procedures, these visits had to be conducted in the clinic. All other visits during the COVID-19 pandemic could be conducted using all or some of the following services: home healthcare, telemedicine conferences (telephone visits) and safety laboratory tests collected at a local laboratory.

Guidelines for remote monitoring during the COVID-19 pandemic included enhanced site management activities that ensured that patient safety and data quality were being maintained when on-site visits were not possible or occurred less frequently.

Patients who failed their screening due to COVID-19 pandemic study suspension and had a documented evaluable endometrial biopsy from the original screening period did not have to undergo a repeat biopsy if they decided to rescreen.

B.2.3.2 Trial methodology

B.2.3.2.1 DAYLIGHT

A summary of the methodology of the DAYLIGHT trial is presented in Table 6.

Potential DAYLIGHT trial patients were requested to complete a questionnaire of their HRT treatment history at the screening visit (up to 21 days prior to randomisation) in order to assess if they are HRT-unsuitable, and therefore eligible for the trial. The responses to the HRT questionnaire were assessed by an independent investigator and HRT-unsuitability categories were allocated based on the eligibility criteria presented in Table 6. The HRT history questionnaire used in DAYLIGHT is presented in Appendix M.

Additionally, to ensure that the four HRT-unsuitable categories (HRT-contraindicated, HRT-caution, HRT-stoppers and HRT-averse) were mutually exclusive, the following hierarchy was used: HRT-contraindicated → HRT-caution → HRT-stoppers → HRT-averse. Patients were

assessed for each group in order and could only be categorised in the first group they were assigned to. However, it is important to note that in real-world practice, patients could belong to more than one group, such as being both HRT-stoppers and HRT-caution, or both HRT-contraindicated and HRT-averse.

For patients who were HRT-contraindicated and HRT-caution, written documentation regarding the conditions listed had to have been present in the medical files of the patients to qualify under these definitions. For HRT-stoppers who stopped treatment due to a lack of efficacy and HRT-related side effects, accurate and comprehensive documentation had to have been provided. This includes, where applicable, the length of HRT treatment, the reasons for determining inefficacy, and the type and duration of HRT-related side effects.

Finally, for patients who were HRT-averse, documentation was provided regarding the nature and extent of the patients' consultation with their healthcare provider, and their reason not to take HRT.

Table 6: Summary of DAYLIGHT trial methodology

Trial name	DAYLIGHT (NCT05033886)[75]
Location	International: DAYLIGHT enrolled patients in approximately 69 sites across 16 countries (Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Hungary, Italy, Netherlands, Norway, Poland, Spain, Sweden, Turkey and the UK). Five UK sites enrolled patients as follows per treatment arm: • Fezolinetant 45 mg (n=26) • Placebo (n=21)
Trial design	Phase 3b, two-arm, randomised, 24-week double-blind, placebo-controlled, parallel group, multicentre study.
Eligibility criteria for patients	**Key Inclusion Criteria** • Institutional Review Board/Independent Ethics Committee approved written informed consent and privacy language as per national regulations obtained from the patient prior to any study-related procedures • Patient is born female, aged ≥40 years and ≤65 years of age at the screening visit • Patients had to be seeking treatment or relief for VMS associated with menopause and confirmed as menopausal per one of the following criteria at the screening visit: ○ Spontaneous amenorrhoea for ≥12 consecutive months ○ Spontaneous amenorrhoea for ≥6 months with biochemical criteria of menopause (follicle-stimulating hormone >40 IU/L) ○ Had bilateral oophorectomy ≥6 weeks prior to the screening visit (with or without hysterectomy) • Patient had VMS and was unsuitable to receive HRT: HRT-contraindicated, HRT-caution, HRT-stoppers and HRT-averse patients. The definitions for HRT-unsuitable categories are provided below: ○ HRT-contraindicated: patients with: ▪ Undiagnosed vaginal bleeding ▪ History of: • Breast cancer or oestrogen dependent tumours • Arterial thromboembolic disease (for example, angina, myocardial infarction, cerebrovascular accident, transient ischemic attack)

	• Venous thrombophilic disorder (for example, deep vein thrombosis, pulmonary embolism) • Hypersensitivity to oestrogen and progesterone therapy or any of the excipients • Porphyria ○ HRT-caution: patients who had a history of: ▪ Diabetes mellitus, hyperlipidaemia ▪ Smoking (current) ▪ Migraine, obesity (body mass index >29.9 kg/m^2) ▪ Systemic lupus erythematosus, epilepsy ▪ Family history of breast cancer in the first degree relative or mutation of breast cancer gene (*BRCA1* and *BRCA2*) ○ HRT-stoppers: patients who had discontinued HRT due to: ▪ Lack of efficacy ▪ HRT-related side effects ▪ Advised by healthcare provider to stop due to length of time on HRT or due to patient's age ≥60 years old ○ HRT-averse: patients who made an informed choice to not take HRT after a consultation about the benefit risks of HRT • Patient had a minimum average of 7 moderate to severe events of VMS per day as recorded in the electronic diary during the last 10 days prior to randomisation • Patient was in good general health as determined on the basis of medical history, general physical examination, laboratory and other medical assessments in the opinion of the investigator • Patient had a negative serology panel (including hepatitis B surface antigen, hepatitis C virus antibody and human immunodeficiency virus antibody screens) **Key exclusion criteria** • Patient used a prohibited therapy for VMS (for example, prescription, over-the-counter or herbal) prior to screening and for the duration of treatment with investigational product • Patient had known documented substance abuse or alcohol addiction within six months of screening • Patient had a history of a malignant tumour within the last five years, except for basal cell carcinoma • Patient had endometrial thickness >8 mm on the locally read screening transvaginal ultrasound or any clinically significant findings that that made the patient ineligible in the opinion of the investigator • Patient had a history of severe allergy, hypersensitivity or intolerance to the investigational product and/or any of its excipients • Patient had a history of seizures or other convulsive disorders, unless well controlled • Patient had a medical condition or chronic disease (including history of neurological [including cognitive], renal, cardiovascular, gastrointestinal, pulmonary [for example, moderate asthma], endocrine or gynaecological disease) or malignancy that could confound interpretation of the study outcome in the opinion of the investigator • Patient had active liver disease, jaundice, elevated liver aminotransferases at screening (alanine aminotransferase or aspartate aminotransferase), elevated total bilirubin or direct bilirubin, elevated International Normalised Ratio or elevated alkaline phosphatase

	A full list of the eligibility criteria for patients in the DAYLIGHT trial are presented in Appendix M.
Settings and locations where the data were collected	Patients were assigned the study drug as a kit containing blister wallets of fezolinetant 15 mg or fezolinetant 30 mg (for a total dose of fezolinetant 45 mg), or matching placebo. The first dose of the study drug took place at the study site on Day 1 (Visit 2) under the supervision of the study staff. On study visit days, the daily dose of the study drug was taken at the study site under the supervision of the study staff or at home under the supervision of the home healthcare personnel, after collection of pre-dose blood samples. On all other days throughout the treatment period, patients were instructed to take the study drug at home, in the morning with a glass of room temperature water.
Study drugs and method of administration	Patients took two tablets orally once daily, with placebo and active tablets being indistinguishable in appearance and shape. Patients on: • Fezolinetant 45 mg received one 15 mg tablet and one 30 mg tablet • Placebo received two placebo tablets (one 30 mg placebo tablet and one 15 mg placebo tablet) to match In the case of a missed dose, patients were instructed to take the dose as soon as they realise they missed the dose. However, where more than six hours had passed since the planned dosing time, patients were instructed to skip the missed dose and continue with the next scheduled dose. Dose modifications were not permitted in the DAYLIGHT trial.
Permitted and disallowed concomitant medication	**Previous and concomitant treatment (medication and non-medication therapy)** • Medications for VMS taken during the 12 months prior to screening and other medication taken 90 days prior to the screening visit and up to the first dose of the study drug (treatment period) were documented in the appropriate eCRF as prior medication • Medications taken after the first dose of study medication through the last study-related activity were documented on the appropriate electronic case report form as concomitant medications. Prior and concomitant medications documented included, but were not limited to, vitamins, herbal remedies (for example, St. John's wort, valerian) and over-the-counter and prescription medication • Patients were instructed not to take any concomitant medications without first consulting the investigator or study coordinator throughout the duration of the study **Previous medication (drugs and therapies):** • For women who had recently discontinued HRT, the therapy must have had been discontinued for at least the following durations prior to the screening visit: • One week or five half-lives of the prior vaginal hormonal products (for example, rings, creams, gels and inserts), whichever was longer; • Four weeks or longer for prior transdermal oestrogen alone or oestrogen/progestin products; • Eight weeks or longer for prior oral oestrogen and/or progestin therapy; • Eight weeks or longer for prior intrauterine progestin therapy; • Three months or longer for prior progestin implants and oestrogen alone injectable drug therapy; or • Six months or longer for prior oestrogen pellet therapy or progestin injectable drug therapy.

	Prohibited concomitant medications: The following medications and therapies were prohibited during treatment with the study drug (from signing of informed consent through the last dose of study IP): • Use of hormonal medications such as HRT or hormonal contraception or any treatment for menopausal symptoms (for example, prescription, over-the-counter or herbal) • Investigational research products that have not been approved for any indication in the country where the patient is enrolled • Strong or moderate cytochrome 1A2 inhibitors (for example, cimetidine, ciprofloxacin, enoxacin and fluvoxamine) • Tamoxifen or aromatase inhibitors (for example, anastrozole and letrozole) • Luteinising hormone-releasing hormone agonists and antagonists (for example, leuprolide, goserelin and degarelix) Further details on prohibited concomitant medications in the DAYLIGHT trial are presented in Appendix M.
Primary endpoint	Mean change in the frequency of moderate to severe VMS from baseline to Week 24
Secondary endpoints	**Key secondary endpoint:** • Mean change in the severity of moderate to severe VMS from baseline to Week 24 **Other secondary endpoints:** • Mean change in the patient-reported sleep disturbance by the PROMIS SD SF 8b total score from baseline to Week 24 • Mean change in the frequency of moderate to severe VMS from baseline to each visit, up to Week 20 • Mean change in the severity of moderate to severe VMS from baseline to each visit, up to Week 20 • Mean percent change in the frequency of moderate and severe VMS from baseline to each visit, up to Week 24 • Responder of percent reduction ≥ 50%, ≥ 75% and at 100% in the frequency of moderate and severe VMS from baseline to each visit, up to Week 24 • Frequency and severity of treatment-emergent AEs, clinical laboratory assessments, vital signs and electrocardiogram
Exploratory endpoints	Exploratory endpoints included: • Mean change in the EQ-5D-5L total score and VAS from baseline to Weeks 4, 12, 16 and 24 • Mean change in the WPAI-VMS domain scores from baseline to each to Weeks 4, 12, 16 and 24 • Mean change in PHQ-4 total and domain scores from baseline to Weeks 4, 12, 16 and 24 • Mean change in the MENQOL total score from baseline to Weeks 4, 12, 16 and 24 • Mean change in the MENQOL domain scores from baseline to Weeks 4, 12, 16 and 24 • Score on the PGI-C VMS at Weeks 4, 12, 16 and 24
Pre-specified subgroups	Subgroup variables included: • Smoking status (active smoker or non-smoker) • Race (white, other)

	• BMI (<25 kg/m^2, ≥25 kg/m^2)
	• Age:
	  ○ Category 1: <55 years, ≥55 years
	  ○ Category 2: ≥40 to <46 years, ≥46 to <51 years, ≥51 to <56 years, ≥56 to <61 years and ≥61 to <66 years

Abbreviations: AE: adverse effect; EQ-5D-5L: EuroQoL five dimensions five level version; HRT: hormone replacement therapy; MENQOL: Menopause Quality of Life Questionnaire; PGI-C: Patient Global Impressions scale – Clinician; PHQ-4: Patient Health Questionnaire for Anxiety and Depression; PROMIS SD SF 8b: Patient-reported Outcomes Measurement Information System Sleep Disturbance – Short Form 8b; VMS: vasomotor symptoms; WPAI-VMS: Work Productivity and Activity Impairment for Vasomotor Symptoms.
Source: Astellas (Data on File). DAYLIGHT Study Protocol and CSR.[79, 82]

B.2.3.2.2 SKYLIGHT 1 and SKYLIGHT 2

A summary of the study methodology of the SKYLIGHT 1 and SKYLIGHT 2 trials is present in Table 8.

Whilst the DAYLIGHT trial exclusively enrolled patients who were HRT-unsuitable, patients in the SKYLIGHT 1 and SKYLIGHT 2 trials were not required to be HRT-unsuitable to be considered for inclusion in the trials. However, HRT treatment history was collected at the screening visit (Visit 1; Days −35 to −1); these informed pre-specified subgroup analyses in the pooled HRT-unsuitable population. Patients were categorised into the HRT subgroups listed below (Table 7) based on their responses to the HRT questionnaire incorporated in the eCRF (see Appendix M for HRT unsuitability questionnaire used in SKYLIGHT 1 and SKYLIGHT 2). The conditions and medical history terms listed below, were used to classify female patients who were between 40 and 65 years of age with moderate to severe VMS associated with menopause into the different HRT subgroups.

It should be noted that the approach to defining HRT-stoppers slightly differed between DAYLIGHT and the SKYLIGHT trials. A single definition of HRT-stoppers was used in DAYLIGHT compared with two definitions used in the SKYLIGHT trials (see Table 7). The definition used in DAYLIGHT i.e., patients who had discontinued HRT due to lack of efficacy, HRT-related side effects, those advised by their healthcare provider to stop HRT due to the length of time on HRT or due to patient's age ≥60 years old, aligned more closely with the "HRT Stopped for Medical Reasons" category in the SKYLIGHT trials.

Table 7: HRT-unsuitable subgroups as defined in the SKYLIGHT trials

HRT-unsuitable subgroup	Definition
HRT-contraindicated	Patients with any of the following underlying conditions: • Undiagnosed abnormal genital (vaginal) bleeding • Known, suspected or history of breast cancer • Known or suspected oestrogen dependent tumours (e.g. breast cancer, endometrial cancer, leiomyosarcoma, ovarian cancers – granulosa cell, endometroid and low grade serous type) • Arterial thromboembolic disease • Venous thromboembolic disease • Any (other) thrombophilic disorder • Acute liver disease, or a history of liver disease where liver function tests have failed to return to normal • Known or suspected pregnancy

	• Known hypersensitivity to oestrogen and progesterone therapy or any of the excipients
HRT-caution	Patients with the following underlying medical condition(s) that warranted a cardiovascular or breast cancer risk assessment before prescribing HRT: • Diabetes mellitus - with or without vascular involvement; Hyperlipidaemia (high cholesterol or triglycerides) • Smoker • Obesity • Migraine • Family history of breast cancer in first degree relative or patient with mutation of BRCA 1 and 2 • Lupus • Epilepsy This subgroup also included patients who had been advised by their healthcare professional not to take HRT due to an (unspecified) underlying medical condition.
HRT-stoppers	Patients who stopped HRT.
HRT stopped for medical concerns	Patients who stopped HRT due to medical concerns.
HRT-naïve/willing	Patients who had not previously been treated with HRT and who were willing to take HRT.
HRT-averse	Patients who were not being treated with HRT at the time of randomisation and who were not willing to take HRT.
HRT-contraindicated or -caution	Combined patients who fell into either the HRT-contraindicated or HRT-caution subgroups.
HRT-unsuitable	Patients who were in any of the following subgroups: • HRT-contraindicated • HRT-caution • HRT stopped for medical concerns • HRT-averse

Abbreviations: HRT: hormone replacement therapy

Table 8: Summary of methodology for SKYLIGHT 1 and SKYLIGHT 2 trials

Trial name	SKYLIGHT 1 (NCT04003155) and SKYLIGHT 2 (NCT04003142)
Location	International: SKYLIGHT 1 enrolled patients in 97 sites across 7 countries (US, Canada, Czech Republic, Hungary, Poland, Spain, and the UK); SKYLIGHT 2 in 91 sites across 7 countries (US, Canada, Czech Republic, Latvia, Poland, Spain, and the UK). Across both trials 5 UK sites enrolled patients as follows per treatment arm: • Fezolinetant 30 mg (n=3) • Fezolinetant 45 mg (n=4) • Placebo (n=5)
Trial design	Identically designed, concurrent, phase 3, randomised, placebo-controlled, 12-week double-blind, parallel group, multicentre trials, followed by a 40-week double-blind, non-controlled extension treatment period
Eligibility criteria for patients[a]	**Key inclusion criteria** • Signed informed consent must be obtained before any study-related assessments are performed

	• Female patients aged ≥40 years and ≤65 years at screening with moderate to severe VMS
	• BMI ≥18 kg/m^2 and ≤38 kg/m^2
	• Patient had to be seeking treatment or relief for VMS and confirmed as menopausal as per one of the following criteria:
	  ○ Spontaneous amenorrhoea for ≥12 consecutive months
	  ○ Spontaneous amenorrhoea for ≥6 months with biochemical criteria of menopause (follicle stimulating hormone >40 IU/L); or
	  ○ Having had bilateral oophorectomy ≥6 weeks prior to the screening visit
	• Within the 10 days prior to randomisation, patient must have a minimum average of 7 to 8 moderate to severe VMS a day, or 50 to 60 moderate to severe VMS per week
	• Normal, negative, or no clinically significant findings on mammogram within the previous 12 months or at screening
	• Normal or not clinically significant Papanicolaou test (or equivalent cervical cytology) result within the previous 12 months or at screening
	• *For women with uterus*: willing to undergo a transvaginal ultrasound to evaluate the uterus and ovaries at screening and at Week 52 (end of treatment), and at early discontinuation for women who withdraw from the study before completion
	• *For women with uterus*: willing to undergo an endometrial biopsy at screening and at Week 52 (end of treatment) unless has had a supracervical or full hysterectomy; the endometrial biopsy obtained at screening should be considered evaluable and they should be willing to undergo endometrial biopsy in case of uterine bleeding or early discontinuation of the study or study drug
	Key exclusion criteria
	• Receiving strong or moderate cytochrome P450 1A2 inhibitors, hormone replacement therapy, hormonal contraceptive, or any treatment for vasomotor symptoms (prescription, over the counter, or herbal)
	• Previous or existing history of a malignant tumour, except for basal cell carcinoma
	• Systolic blood pressure ≥130 mm Hg or diastolic blood pressure ≥80 mm Hg based on an average of two or three readings on at least two different occasions within the screening period; women who did not meet these criteria might, at the discretion of the investigator, be reassessed after initiation or review of antihypertensive measures; women with a medical history of hypertension could be enrolled at the discretion of the investigator once they were medically clear (stable and compliant)
	• History within the past 6 months of undiagnosed uterine bleeding
	• A medical condition or chronic disease (for example, history of neurological [for example, cognitive], hepatic, renal, cardiovascular, gastrointestinal, pulmonary [for example, moderate asthma], endocrine, or gynaecological disease) or malignancy that could confound interpretation of the study
	• Active liver disease, jaundice, or elevated liver aminotransferases (alanine aminotransferase or aspartate aminotransferase), elevated total or direct bilirubin, elevated international normalised ratio, or elevated alkaline phosphatase
	• Creatinine more than 1.5 times upper limit of normal; or estimated glomerular filtration rate ≤59 mL/min per 1.73 m^2 at screening
	
	For a full list of inclusion and exclusion criteria, can be found in Appendix M.

Settings and locations	Patients were assigned study drug as a kit in blister packaging. The first intake of study drug took place at the study site on Day 1 (Visit 2) under the

where the data were collected	supervision of the study staff. On study visit days, the daily dose of study drug will be taken at the study site, under the supervision of the study staff, after collection of pre-dose blood samples. On all other days throughout the treatment period, patients were instructed to take their dose of study drug at home, in the morning with water
Study drugs and method of administration	Patients took two tablets orally once daily with placebo and active tablets being indistinguishable in appearance and shape. Patients on: • Fezolinetant 30 mg received one 30 mg tablet and one 15 mg placebo tablet • Fezolinetant 45 mg received one 15 mg tablet and one 30 mg tablet • Placebo received two placebo tablets (one 30 mg placebo tablet and one 15 mg placebo tablet) to match In the case of a missed dose, the patient should skip the missed dose and continue with the next scheduled dose. Dose increases and decreases were not allowed
Permitted and disallowed concomitant therapy	**Previous and concomitant treatment (medication and non-medication therapy)** • Medication for VMS taken during the 12 months prior to screening and other medication taken 90 days prior to the screening visit and up to the first dose of study medication (treatment period) was documented in the appropriate electronic case report form (eCRF) as prior medication • Medications taken after the first dose of study medication through to the last study-related activity was documented on the appropriate eCRF as concomitant medication. Prior and concomitant medications documented included, but were not limited to, vitamins, herbal remedies (for example, St. John's wort, valerian) and over the counter and prescription medication • Patients were instructed not to take any concomitant medication without first consulting the investigator or study coordinator throughout the study duration **Previous medication (drugs and therapies)** • For women who recently discontinued hormone therapy, the therapy had to be discontinued for at least the following durations prior to the screening visit: ○ One week or longer for prior vaginal hormonal products (rings, creams, gels and inserts) ○ Four weeks or longer for prior transdermal oestrogen alone or oestrogen/progestin products ○ Eight weeks or longer for prior oral oestrogen and/or progestin therapy ○ Eight weeks or longer for prior intrauterine progestin therapy ○ Three months or longer for prior progestin implants and oestrogen alone injectable drug therapy, or ○ Six months or longer for prior oestrogen pellet therapy or progestin injectable drug therapy **Concomitant medications (drugs and therapies)** • All concomitant medications and therapies (prescriptions, over the counter and herbal), other than the study drug, administered from informed consent through 30 days post the last dose of study drug were collected in the eCRF **Prohibited Concomitant Medications** The following medications and therapies were prohibited throughout the study (from signing of informed consent form through to the last study-related activity): • Use of hormonal medications, such as hormone therapy, HRT or hormonal contraception or any treatment for menopausal symptoms (prescription,

	over the counter or herbal) was not allowed during the study • Investigational research products that had not been approved for any indication in the country where the patient was enrolled • Strong or moderate cytochrome P450 1A2 inhibitors A full list of excluded concomitant medications can be found in Appendix M.
Primary endpoints	**Co-primary efficacy endpoints:** • Mean change in the frequency of moderate to severe VMS from baseline to Week 4 • Mean change in the frequency of moderate to severe VMS from baseline to Week 12 • Mean change in the severity of moderate to severe VMS from baseline to Week 4 • Mean change in the severity of moderate to severe VMS from baseline to Week 12
Secondary endpoints	**Key secondary efficacy endpoint:** • Mean change in the PROMIS SD SF 8b total score from baseline to Week 12 **Other secondary efficacy endpoints:** • Mean change in the frequency of moderate to severe VMS from baseline to each week up to Week 12 • Mean change in the severity of moderate to severe VMS from baseline to each week up to week 12 • Mean percent reduction in the frequency of moderate to severe VMS from baseline to each week up to Week 12 • Percent reduction ≥50% and at 100% in the frequency of moderate to severe VMS from baseline to each week up to Week 12 • Score on the PGI-C VMS at each visit • Mean change in the frequency of moderate to severe VMS from baseline to Week 24 (descriptive) • Mean change in the severity of moderate to severe VMS from baseline to week 24 (descriptive) **Other secondary safety endpoints:** • Frequency and severity of adverse events • Transvaginal ultrasound and endometrial biopsy findings • Change from baseline to each time point in vital signs: sitting systolic and diastolic blood pressure and pulse rate • Change from baseline to each time point in electrocardiogram parameters • Change from baseline to each time point in laboratory tests: haematology, biochemistry and urinalysis
Exploratory endpoints	**Exploratory endpoints included:** • Mean change in the PROMIS SRI SF 8a total score from baseline to each visit • Mean change on the EQ-5D-5L total score and VAS from baseline to each visit • Mean change on the WPAI-VMS domain scores from baseline to each visit • Mean change on the MENQOL total score from baseline to each visit • Mean change on the MENQOL domain scores from baseline to each visit
Pre-specified subgroups	The consistency of treatment effect on the co-primary endpoints was evaluated in the pooled SKYLIGHT 1 and SKYLIGHT 2 data as supportive efficacy data. The following prespecified subgroup analyses were performed:

	• Intrinsic factors:
	  ○ Age category 1 (<55 years versus ≥55 years)
	  ○ Race: category 1 (white versus non-white), category 2 (black versus non-black) and category 3 (Asian versus Non-Asian)
	  ○ Ethnicity: Hispanic or Latino versus Not Hispanic or Latino
	  ○ BMI: ≥18.5 kg/m^2 to <25.0 kg/m^2 versus ≥ 25.0 kg/m^2 to <30.0 kg/m^2 versus ≥30.0 kg/m^2
	• Extrinsic factors:
	  ○ Smoking status (current versus former/never)
	  ○ Region: Europe versus North America
	
	Additional *post hoc* subgroup analyses were conducted for the pooled data:
	• Intrinsic factors:
	  ○ Age category 2 (<50 years versus 50 to 60 years versus >60 years)
	  ○ Time since amenorrhoea: <2 years versus 2 to 10 years versus >10 years
	  ○ VMS duration: <1 year versus 1 to 10 years versus >10 years
	  ○ VMS frequency at baseline: ≥median versus <median (median=9.78)
	  ○ VMS severity at baseline: ≥median versus <median (median=2.36)
	• Extrinsic factors:
	  ○ Sleep disturbance (PROMIS SD SF 8b): <1st quartile (=23) versus 1st quartile (=23) to <median (=27) versus median (=27) to < 3rd quartile (=32) versus ≥3rd quartile
	  ○ Medical history of hypertension: yes versus no
	  ○ Oophorectomy history: yes versus no
	  ○ Hysterectomy history: yes versus no
	  ○ Hysterectomy/oophorectomy history: hysterectomy without oophorectomy versus hysterectomy with oophorectomy
	  ○ Previous HRT treatment: yes versus no
	  ○ Reason for stopping HRT: lack of improvement in symptoms versus other
	  ○ Concurrent use of SSRIs: yes versus no
	  ○ Alcohol history: current versus former/never

Abbreviations: BMI: Body mass index; eCRF: electronic case report form; EQ-5D-5L: EuroQoL five dimensions five level version; HRT: hormone replacement therapy; MENQOL: Menopause Quality of Life Questionnaire; PGI-C: Patient Global Impression of Change; PROMIS SD SF 8b: Patient-reported Outcomes Measurement Information System Sleep Disturbance – Short Form 8b; PROMIS SRI SF 8a: Patient-reported Outcomes Measurement Information System Sleep-Related Impairment – Short Form 8a; SSRI: selective serotonin reuptake inhibitor; VAS: visual analogue scale; VMS: vasomotor symptoms; WPAI-VMS: Work Productivity and Activity Impairment questionnaire specific to VMS.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 Study Protocols, CSRs and ISE.[80, 81, 83-85]

B.2.3.3 Outcome definitions

Outcome definitions in DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 are presented in Table 9.

Table 9: Definitions of outcomes

Outcome	Definition

Vasomotor Symptoms	• Consistent with both EMA[48] and FDA[86] guidance, the following definitions of VMS were used for the conduct of clinical evaluations in moderate to severe VMS:
	  ○ Mild: sensation of heat without sweating
	  ○ Moderate: sensation of heat with sweating, able to continue activity

	○ Severe: sensation of heat with sweating, causing cessation of activity • Frequency of moderate or severe VMS events were calculated as the sum of moderate or severe VMS events per day • VMS severity was calculated using the following weighted average approach: ○ For post-baseline endpoints (this calculation allows for capturing of all VMS on treatment, including mild): $\frac{([\text{number of mild VMS/day} \times 1] + [\text{number of moderate VMS/day} \times 2] + [\text{number of severe VMS/day} \times 3])}{\text{Total number of daily (or weekly) mild/moderate/severe VMS}}$ ○ For baseline (the calculation for severity at baseline used a similar formula, but does not include mild VMS events in the numerator or denominator): $\frac{([\text{number of moderate VMS/day} \times 2] + [\text{number of severe VMS/day} \times 3])}{\text{Total number of daily (or weekly) moderate/severe VMS}}$ • Daily mean frequency and severity per week were calculated as the average frequency or severity over seven days. Patients had to provide data for 50% of any given week (≥4 days) for their data to be included in the analysis
PROMIS SD SF 8b	The PROMIS SD SF 8b assessed self-reported sleep disturbance over the past seven days and included perceptions of restless sleep; satisfaction with sleep; refreshing sleep; difficulties sleeping, getting to sleep or staying asleep; amount of sleep; and sleep quality. Because it assessed the patient's experience of sleep disturbance, the measure did not focus on specific sleep-disorder symptoms or ask patients to report objective measures of sleep (for example, total amount of sleep, time to fall asleep, and amount of wakefulness during sleep). Responses to each of the eight items ranged from one to five, and the range of possible summed raw scores was eight to 40. Higher scores on the PROMIS SD SF 8b indicated increased sleep disturbance.
PROMIS SRI SF 8a[a]	The PROMIS SRI SF 8a is an 8-item PRO measure that evaluated self-reported perceptions of alertness, sleepiness and tiredness during usual waking hours and the perceived functional impairments during wakefulness associated with sleep problems or impaired alertness. Though this measure did not directly assess cognitive, affective or performance impairment, it did measure waking alertness, sleepiness and function within the context of overall sleep-wake function. The PROMIS SRI SF 8a is a universal rather than disease-specific instrument, and had a 7-day recall period. Responses to each of the 8 items on the PROMIS SRI SF 8a ranged from 1 to 5, and the range of possible summed raw scores was 8 to 40. Higher scores indicated more sleep-related impairment.
MENQOL	The MENQOL is a 29-item PRO measure that assessed the impact of 4 domains of menopausal symptoms, as experienced over the last week: vasomotor (items 1 to 3), psychosocial (items 4 to 10), physical (items 11 to 26) and sexual (items 27 to 29). Items pertaining to a specific symptom were rated as present or not present, and if present, how bothersome on a zero (not bothersome) to 6 (extremely bothersome) scale. Each item score ranged from 1 to 8, and each domain was scored separately; each domain mean ranged from 1 to 8. The overall questionnaire score was the mean of the domain means. Higher scores represented more bothersome menopausal symptoms.
EQ-5D-5L with Visual Analogue Scale (VAS)	The EQ-5D-5L is a five-item standardised measure of health status that provides a simple, generic measure of health for clinical and economic appraisal. This measure comprises five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has five levels: no problems, slight problems, moderate problems, severe problems and extreme problems. The patient is asked to indicate their health state by selecting the most

	appropriate statement in each of the five dimensions. This decision results in a one-digit number that expresses the level selected for that dimension. The digits for the five dimensions can be combined into a five-digit number that describes the patient's health state. The EQ-5D VAS is a patient-reported measure that records the respondent's self-rated health on a vertical VAS where the endpoint is labelled 'Best imaginable health state' and 'Worst imaginable health state.' The scale ranges from zero to 100, where 100 indicates the patient is in their best possible health state and zero indicates the patient is in their worst possible health state. Patients mark an 'X' on the scale to rate their health status that day.
WPAI-VMS	The WPAI-VMS is a six-item patient-reported outcome measure that examines VMS-related work productivity and activity in the preceding seven day. It comprises four domains: 1) absenteeism (the percentage of work time missed because of VMS in the past seven days); 2) presenteeism (the percentage of impairment experienced while at work in the past seven days because of VMS); 3) overall work productivity loss (overall work impairment measured by combining absenteeism and presenteeism to determine the total percentage of missed time); and 4) activity impairment (the percentage of impairment in daily activities because of VMS in the past seven days). If the patient is unemployed, only two of the six items on this questionnaire require completion. WPAI-VMS outcomes are expressed as impairment percentages, with higher numbers indicating greater impairment and less productivity (that is, worse outcomes).
PHQ-4[b]	The PHQ-4 is a four-item patient-reported outcome measure that assesses two domains of mental health (anxiety and depression) during the past two weeks. Items are presented with a four-point Likert scale response option. The score range is from zero to three and associate with responses of "not at all" to "nearly every day." Items are scored at the individual level and can be added to a total score that can be categorically interpreted as: normal (zero to two), mild (three to five), moderate (six to eight) and severe (nine to 12). The interpretation of higher scores is a greater likelihood of underlying anxiety or depression.
Safety	**Definition of treatment-emergent adverse event** Treatment emergent adverse events was defined as an adverse event (AE) observed after starting administration of the study drug and 21 days after the last dose of study drug. **Definition of serious adverse events** An AE was considered "serious" if, in the view of either the investigator or sponsor, it resulted in any of the following outcomes: • Resulted in death • Was life-threatening (an AE was considered "life-threatening" if, in the view of either the investigator or sponsor, its occurrence places the patient at immediate risk of death. It does not include an AE that, had it occurred in a more severe form, might have caused death) • Resulted in persistent or significant disability/incapacity or substantial disruption of the ability to conduct normal life functions • Resulted in congenital anomaly or birth defect • Required inpatient hospitalisation (except for planned procedures as allowed per study) or leads to prolongation of hospitalisation (except if prolongation of planned hospitalization is not caused by an AE). Hospitalisation for treatment/observation/examination caused by AE was considered as serious. • Discontinuation due to increases in liver enzymes • Other medically important events: ○ Medical and scientific judgement was exercised in deciding whether expedited reporting was appropriate in other situations, such as important

medical events that may not have been immediately life-threatening or resulted in death or hospitalisation but may have jeopardised the subject or may have required intervention to prevent one of the other outcomes listed in the definition above. These events, including those that may have resulted in disability/incapacity, were usually considered serious. Examples of such events were intensive treatment in an emergency room or at home for allergic bronchospasm; blood dyscrasias or convulsions that did not result in hospitalisation; or development of drug dependency or drug abuse

Criteria for defining the severity of an adverse event

The investigator will use the following definitions to rate the severity of each AE:

- Mild: No disruption of normal daily activities
- Moderate: Affect normal daily activities
- Severe: Inability to perform daily activities

Definition of Liver Abnormalities

Confirmed abnormalities was characterised as moderate and severe where upper limit of normal (ULN):

	ALT or AST		TBL
Moderate	>3 × ULN	or	>2 × ULN
Severe	>3 × ULN	and	>2 × ULN

In addition, a patient was considered to have severe hepatic test abnormalities for any of the following:

- ALT or AST >8 × ULN
- ALT or AST >5 × ULN for more than 2 weeks
- ALT or AST >3 × ULN AND TBL >2 × ULN or INR >1.5, or
- ALT or AST >3 × ULN with the appearance of fatigue, nausea, vomiting, right upper quadrant pain or tenderness, fever, rash and/or eosinophilia (>5% increase above baseline)

The investigator could have determined that abnormal liver test results, other than as described above, qualified as moderate or severe abnormalities, necessitating additional monitoring and follow-up.

Hy's Law Definition: drug-induced jaundice caused by hepatocellular injury, without a significant obstructive component, has a high rate of negative outcomes, from 10 to 50% mortality (or transplant).

The 2 "requirements" for Hy's Law are:[87]

1. Evidence that a drug can cause hepatocellular-type injury, generally shown by an increase in transaminase elevations higher 3 × ULN ("2 × ULN elevations are too common in treated and untreated patients to be discriminating").
2. Cases of increased bilirubin (at least 2 × ULN) with concurrent transaminase elevations at least 3 × ULN and no evidence of intra- or extra-hepatic bilirubin obstruction (elevated ALP) or Gilbert's syndrome.

[a] Outcome was only measured in SKYLIGHT 1 and SKYLIGHT 2. [b] Outcome was only measured in DAYLIGHT.

Abbreviations: ALP: alkaline phosphatase; ALT: alanine aminotransferase; AST: aspartate aminotransferase; EMA: European Medicines Agency; EQ-5D-5L: EuroQoL five dimensions five level version; FDA: Food and Drug Administration; MENQOL: Menopause Quality of Life Questionnaire; PHQ-4: Patient Health Questionnaire for Anxiety and Depression; PROMIS SD SF 8b: Patient-reported Outcomes Measurement Information System Sleep Disturbance – Short Form 8b; PROMIS SD SRI 8a: Patient-reported Outcomes Measurement Information System Sleep Related Impairment – Short Form 8a; TBL: total bilirubin; VMS: vasomotor symptoms.
Source: Astellas (Data on File). DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 Study Protocols.[79-81]

B.2.3.4 Baseline characteristics

B.2.3.4.1 DAYLIGHT

A summary of the baseline characteristics of patients in the DAYLIGHT trial are presented in Table 10.

Overall, demographics and baseline characteristics were similar between treatment groups in DAYLIGHT. Median age was similar across the fezolinetant 45 mg (55.0 years) and placebo (54.0 years) groups. Most patients were white (fezolinetant 45mg: 96.0%; placebo: 97.3%) and were either former smokers or had never smoked (fezolinetant 45 mg: 84.1%; placebo: 84.5%).

Approximately 72% of patients received prior treatment with HRT (fezolinetant 45 mg: 73.5%; placebo: 71.2%), approximately 9% of patents had a history of oophorectomy (fezolinetant 45 mg: 10.2%; placebo: 8.4%), and approximately 14.2% of patients had a history of hysterectomy (fezolinetant 45 mg: 9.7%; placebo: 18.6%). The proportion of these baseline characteristics and medical history was similar across treatment groups.

Table 10: Demographics, baseline characteristics, HRT history and targeted medical history in DAYLIGHT (Full Analysis Set)

Parameter	DAYLIGHT	
	Placebo (n=226)	Fezolinetant 45 mg (n=226)
Demographics		
Age		
Mean (SD)	54.1 (4.6)	54.9 (4.8)
Median	54.0	55.0
Min–Max	43–65	40–65
Age group (Category 1), n (%)		
< 55 years	125 (55.3)	108 (47.8)
≥ 55 years	101 (44.7)	118 (52.2)
Age group (Category 2), n (%)		
≥ 40 to <46	6 (2.7)	3 (1.3)
≥ 46 to <51	40 (17.7)	36 (15.9)
≥ 51 to <56	93 (41.2)	83 (36.7)
≥ 56 to <61	64 (28.3)	72 (31.9)
≥ 61 to <66	23 (10.2)	32 (14.2)
Race (Category 1), n (%)		
White	218 (97.3)	217 (96.0)
Black or African American	0	4 (1.8)

Asian	5 (2.2)	1 (0.4)
Other	1 (0.4)	1 (0.4)
More Than One Race	0	3 (1.3)
Missing	2	0
Race (Category 2), n (%)		
White	218 (97.3)	217 (96.0)
Other[a]	6 (2.7)	9 (4.0)
Missing	2	0
Weight (kg)		
Mean (SD)	72.56 (13.18)	74.23 (12.60)
Median	71.45	72.55
Min–Max	45.0–118.0	49.0–110.0
BMI (kg/m^2)		
Mean (SD)	26.98 (4.52)	27.42 (4.33)
Median	26.43	26.81
Min–Max	17.6–42.4	19.0–39.8
BMI (kg/m^2), n (%) – Category 1		
<18.5	3 (1.3)	0
≥18.5 to <25	82 (36.3)	67 (29.6)
≥25 to <30	90 (39.8)	104 (46.0)
≥30	51 (22.6)	55 (24.3)
BMI (kg/m^2), n (%) – Category 2		
<25	85 (37.6)	67 (29.6)
≥25	141 (62.4)	159 (70.4)
Smoking status stratification factor, n (%)[b]		
Current	35 (15.5)	36 (15.9)
Former/never	191 (84.5)	190 (84.1)
HRT history		
Prior HRT, n (%)		
Yes	65 (28.8)	60 (26.5)
No	161 (71.2)	166 (73.5)
HRT-unsuitable group, n (%)		
Contraindicated	23 (10.2)	27 (11.9)
Caution	91 (40.3)	74 (32.7)
Stoppers	37 (16.4)	32 (14.2)
Averse	75 (33.2)	93 (41.2)
Patients previously treated, reason for stopping HRT, n (%)[c]		
Lack of improvement in symptoms	24 (36.9)	18 (30.0)
Side effects	19 (29.2)	21 (35.0)
Worried about possible long-term risks	14 (21.5)	15 (25.0)
Family history of breast cancer	5 (7.7)	7 (11.7)

HCP advised due to length of time on HRT	10 (15.4)	8 (13.3)
HCP advised due to patient age	5 (7.7)	5 (8.3)
HCP advised due to medical reasons	9 (13.8)	9 (15.0)
Other personal reason	7 (10.8)	6 (10.0)
Unknown	0	1 (1.7)
Willing to take HRT for VMS, n (%)[d]		
Yes	11 (6.8)	5 (3.0)
No	150 (93.2)	161 (97.0)
Advised by healthcare professional not to take HRT, n (%)[d]		
Yes	56 (34.8)	44 (26.5)
Underlying medical condition[d]	27 (65.9)	28 (73.7)
Family history of breast cancer[d]	18 (43.9)	15 (39.5)
Missing	15	6
No	103 (64.0)	121 (72.9)
Unknown	2 (1.2)	1 (0.6)
Targeted medical history		
Time since onset of VMS (months)		
Mean (SD)	62.43 (54.23)	64.38 (53.41)
Median	49.95	49.05
Min–Max	0.0–266.2	0.0–327.4
Time since onset of amenorrhoea (months)		
n	211	214
Mean (SD)	58.60 (49.46)	71.85 (60.72)
Median	42.28	54.32
Min–Max	0.0–290.2	3.1–314.4
History of oophorectomy, n (%)		
Yes	19 (8.4)	23 (10.2)
No	207 (91.6)	203 (89.8)
History of hysterectomy, n (%)		
Yes	22 (9.7)	42 (18.6)
No	204 (90.3)	184 (81.4)

The full analysis set (FAS) comprised all patients who were randomised and received one or more doses of study drug; the randomised treatment for each patient was used for summaries by treatment group based on the FAS, even if a patient erroneously received a different treatment. [a]For any participant who reported more than one race, the participant was included in "other" under race category 2 [b]Current versus former/never smoking status was a stratification factor for randomisation. [c]The denominator is number of patients who have previously been treated with HRT; a patient can have more than one reason for stopping HRT. [d]The denominator is number of patients who have not been previously treated with HRT.
Abbreviations: BMI: body mass index; HRT: hormone replacement therapy; Max: maximum; Min: minimum; VMS: vasomotor symptoms; SD: standard deviation.
Source: Astellas (Data on File). DAYLIGHT CSR.[82]

B.2.3.4.2 SKYLIGHT 1 and SKYLIGHT 2

A summary of the demographics and baseline characteristics of patients in SKYLIGHT 1 and SKYLIGHT 2 is presented in Table 11.

Overall, demographics and baseline characteristics were similar between treatment groups in each pivotal study. In the pooled analysis, patients in each treatment group had a median age of 54.0 years and were either former smokers or had never smoked (fezolinetant 30 mg: 83.8%; fezolinetant 45 mg: 83.3%; placebo: 83.3%). Most patients were white (fezolinetant 30 mg: 82.2%; fezolinetant 45mg: 80.4%; placebo: 80.7%), with the remaining 17.8%, 19.6% and 19.3%, respectively, being of an unspecified or other ethnic origins including black or African American, American Indians or Alaska natives, Asians or other Pacific islanders. Based on census data from 2021, these demographics are broadly aligned with that of the population of England and Wales (82% white; 18% black, Asian, mixed or other ethnic group).[88]

Most patients (fezolinetant 30 mg: 85.8%; fezolinetant 45 mg: 84.2%; placebo: 86.8%) in the pooled analysis of SKYLIGHT 1 and SKYLIGHT 2 were categorised as HRT-unsuitable and approximately 20% of patients in each treatment arm received prior HRT (fezolinetant 30 mg: 20.3%; fezolinetant 45 mg: 20.5%; placebo: 19.0%). Approximately 22% had a history of oophorectomy (fezolinetant 30 mg: 20.6%; fezolinetant 45 mg: 22.3%; placebo: 21.9%), and approximately 32% of patients had a history of hysterectomy (fezolinetant 30 mg: 33.3%; fezolinetant 45 mg: 33.4%; placebo: 29.8%). The proportion of these baseline characteristics and medical history were similar across treatment groups.

Table 11: Demographics, baseline characteristics, HRT history and targeted medical history (Final Analysis Set)

Parameter	SKYLIGHT 1			SKYLIGHT 2			Pooled studies (SKYLIGHT 1 AND SKYLIGHT 2)		
	Placebo (n=175)	Fezolinetant 30 mg (n=173)	Fezolinetant 45 mg (n=174)	Placebo (n=167)	Fezolinetant 30 mg (n=166)	Fezolinetant 45 mg (n=167)	Placebo (n=342)	Fezolinetant 30 mg (n=339)	Fezolinetant 45 mg (n=341)
Demographics									
Age									
Mean (SD)	54.7 (4.8)	54.1 (4.8)	54.3 (5.1)	54.7 (4.8)	53.9 (4.9)	54.3 (5.4)	54.7 (4.7)	54.0 (4.9)	54.3 (5.3)
Median	54.0	53.0	54.0	54.0	54.0	54.0	54.0	54.0	54.0
Min–Max	41–65	42–65	40–65	44–65	42–65	40–65	41–65	42–65	40–65
Age category, n (%)									
< 55 years	91 (52.0)	93 (53.8)	97 (55.7)	89 (53.3)	93 (56.0)	82 (49.1)	180 (52.6)	186 (54.9)	179 (52.5)
≥ 55 years	84 (48.0)	80 (46.2)	77 (44.3)	78 (46.7)	73 (44.0)	85 (50.9)	162 (47.4)	153 (45.1)	162 (47.5)
Race, n (%)									
White	142 (81.1)	147 (85.5)	142 (81.6)	134 (80.2)	131 (78.9)	132 (79.0)	276 (80.7)	278 (82.2)	274 (80.4)
Black or African American	28 (16.0)	21 (12.2)	26 (14.9)	31 (18.6)	35 (21.1)	33 (19.8)	59 (17.3)	56 (16.6)	59 (17.3)
American Indian or Alaska Native	2 (1.1)	0	1 (0.6)	0	0	1 (0.6)	2 (0.6)	0	2 (0.6)
Asian	3 (1.7)	3 (1.7)	3 (1.7)	1 (0.6)	0	0	4 (1.2)	3 (0.9)	3 (0.9)
Other Pacific Islander	0	0	1 (0.6)	0	0	0	0	0	1 (0.3)
More Than One Race	0	1 (0.6)	1 (0.6)	1 (0.6)	0	1 (0.6)	1 (0.3)	1 (0.3)	2 (0.6)
Missing	0	1	0	0	0	0	0	1	0
Ethnicity, n (%)									
Hispanic or Latino	46 (26.4)	42 (24.3)	48 (27.6)	32 (19.3)	34 (20.5)	41 (24.6)	78 (22.9)	76 (22.4)	89 (26.1)
Not Hispanic or Latino	128 (73.6)	131 (75.7)	126 (72.4)	134 (80.7)	132 (79.5)	126 (75.4)	262 (77.1)	263 (77.6)	252 (73.9)
Missing	1	0	0	1	0	0	2	0	0

Weight (kg)									
Mean (SD)	74.4 (12.14)	75.04 (13.86)	75.70 (12.89)	74.57 (14.68)	75.33 (14.09)	74.62 (12.45)	74.49 (13.42)	75.18 (13.95)	75.17 (12.67)
Median	73.60	73.80	74.35	71.70	74.05	73.00	72.60	74.00	73.50
Min–Max	47.7–111.0	42.0–121.2	50.6–110.6	46.2–125.0	48.0–108.4	45.0–107.4	46.2–125.0	42.0–121.2	45.0–110.6
BMI (kg/m^2)									
n	175	173	173	167	166	167	342	339	340
Mean (SD)	28.19 (4.28)	28.09 (4.80)	28.33 (4.39)	28.16 (4.99)	27.94 (4.69)	27.91 (4.35)	28.17 (4.64)	28.02 (4.74)	28.12 (4.37)
Median	28.23	27.85	27.89	27.78	27.76	27.28	28.10	27.76	27.64
Min–Max	18.8–37.7	18.0–37.8	18.4–37.9	18.6–38.0	18.1–37.6	18.0–37.5	18.6–38.0	18.0–37.8	18.0–37.9
BMI category (kg/m^2), n (%)									
<18.5	0	1 (0.6)	1 (0.6)	0	1 (0.6)	1 (0.6)	0	2 (0.6)	2 (0.6)
≥18.5 to <25	44 (25.1)	50 (28.9)	40 (23.1)	53 (31.7)	54 (32.5)	45 (26.9)	97 (28.4)	104 (30.7)	85 (25.0)
≥25 to <30	71 (40.6)	60 (34.7)	79 (45.7)	62 (37.1)	58 (34.9)	73 (43.7)	133 (38.9)	118 (34.8)	152 (44.7)
≥30	60 (34.3)	62 (35.8)	53 (30.6)	52 (31.1)	53 (31.9)	48 (28.7)	112 (32.7)	115 (33.9)	101 (29.6)
Missing	0	0	1	0	0	0	0	0	1
Smoking status stratification factor, n (%)[a]									
Current	22 (12.6)	21 (12.1)	23 (13.2)	35 (21.0)	34 (20.5)	34 (20.4)	57 (16.7)	55 (16.2)	57 (16.7)
Former/never	153 (87.4)	152 (87.9)	151 (86.8)	132 (79.0)	132 (79.5)	133 (79.6)	285 (83.3)	284 (83.8)	284 (83.3)
HRT history									
Suitable for HRT, n (%)[b]									
Yes	-	-	-	-	-	-	[redacted]	[redacted]	[redacted]
No	-	-	-	-	-	-	[redacted]	[redacted]	[redacted]
HRT-unsuitable subgroup, n (%)[b]									
Contraindicated or Caution	-	-	-	-	-	-	[redacted]	[redacted]	[redacted]
Contraindicated	-	-	-	-	-	-	[redacted]	[redacted]	[redacted]
Caution	-	-	-	-	-	-	[redacted]	[redacted]	[redacted]
Stopper	-	-	-	-	-	-	[redacted]	[redacted]	[redacted]

Stopped for Medical Concerns	-	-	-	-	-	-	[redacted]	[redacted]	[redacted]
Averse	-	-	-	-	-	-	[redacted]	[redacted]	[redacted]
Naïve/Willing	-	-	-	-	-	-	[redacted]	[redacted]	[redacted]
Prior HRT, n (%)									
Yes	33 (19.4)	31 (18.1)	30 (17.8)	31 (18.6)	37 (22.6)	38 (23.3)	64 (19.0)	68 (20.3)	68 (20.5)
No	137 (80.6)	140 (81.9)	139 (82.2)	136 (81.4)	127 (77.4)	125 (76.7)	273 (81.0)	267 (79.7)	264 (79.5)
Missing	5	2	5	0	2	4	5	4	9
Patients previously treated, reason for stopping HRT, n (%)[c]									
Lack of improvement in Symptoms	13 (39.4)	7 (22.6)	9 (30.0)	10 (32.3)	17 (45.9)	8 (21.1)	23 (35.9)	24 (35.3)	17 (25.0)
Side effects	6 (18.2)	7 (22.6)	5 (16.7)	4 (12.9)	5 (13.5)	11 (28.9)	10 (15.6)	12 (17.6)	16 (23.5)
Worried about possible long-term risks	11 (33.3)	10 (32.3)	5 (16.7)	9 (29.0)	8 (21.6)	12 (31.6)	20 (31.3)	18 (26.5)	17 (25.0)
Family history of breast cancer	0	1 (3.2)	1 (3.3)	3 (9.7)	2 (5.4)	1 (2.6)	3 (4.7)	3 (4.4)	2 (2.9)
HCP advised due to length of time on HRT	3 (9.1)	1 (3.2)	2 (6.7)	3 (9.7)	2 (5.4)	3 (7.9)	6 (9.4)	3 (4.4)	5 (7.4)
HCP advised due to patient age	0	1 (3.2)	1 (3.3)	0	0	0	0	1 (1.5)	1 (1.5)
HCP advised due to medical reasons	0	2 (6.5)	2 (6.7)	2 (6.5)	4 (10.8)	2 (5.3)	2 (3.1)	6 (8.8)	4 (5.9)
Other personal reason	2 (6.1)	13 (41.9)	6 (20.0)	5 (16.1)	7 (18.9)	10 (26.3)	7 (10.9)	20 (29.4)	16 (23.5)
Unknown	2 (6.1)	0	3 (10.0)	0	2 (5.4)	0	2 (3.1)	2 (2.9)	3 (4.4)
Willing to take HRT for VMS, n (%)[d]									
Yes	60 (43.8)	73 (52.1)	63 (45.3)	47 (34.6)	41 (32.3)	49 (39.2)	107 (39.2)	114 (42.7)	112 (42.4)
No	77 (56.2)	67 (47.9)	76 (54.7)	89 (65.4)	86 (67.7)	76 (60.8)	166 (60.8)	153 (57.3)	152 (57.6)
Advised by healthcare professional not to take HRT, n (%)[d]									
Yes	10 (7.3)	8 (5.7)	12 (8.6)	18 (13.2)	21 (16.5)	17 (13.6)	28 (10.3)	29 (10.9)	29 (11.0)

Underlying medical condition[e]	5 (62.5)	6 (85.7)	10 (83.3)	11 (61.1)	15 (83.3)	15 (93.8)	16 (61.5)	21 (84.00)	25 (89.3)
Family history of breast cancer[e]	5 (62.5)	3 (42.9)	5 (41.7)	7 (38.9)	5 (27.8)	1 (6.3)	12 (46.2)	8 (32.0)	6 (21.4)
Missing	2	1	0	0	3	1	2	4	1
No	91 (66.4)	94 (67.1)	88 (63.3)	110 (80.9)	100 (78.7)	103 (82.4)	201 (73.6)	194 (72.7)	191 (72.3)
Unknown	36 (26.3)	38 (27.1)	39 (28.1)	8 (5.9)	6 (4.7)	5 (4.0)	44 (16.1)	44 (16.5)	44 (16.7)
Targeted medical history									
Time since onset of VMS (months)									
Mean (SD)	81.9 (73.6)	77.1 (66.3)	72.2 (59.3)	81.9 (73.3)	76.2 (74.4)	81.7 (74.1)	81.9 (73.4)	76.7 (70.3)	76.9 (67.0)
Median	57.9	53.5	53.9	59.2	47.3	56.3	58.8	49.1	56.0
Min–Max	2–422	5–355	1–293	3–364	3–370	2–396	2–422	3–370	1–396
Time since onset of amenorrhoea (months)									
n	170	169	172	159	163	162	329	332	334
Mean (SD)	81.2 (80.5)	92.3 (83.2)	86.7 (79.7)	85.6 (80.2)	85.7 (79.2)	88.9 (79.5)	83.3 (80.2)	89.0 (81.2)	87.8 (79.5)
Median	50.8	70.1	63.2	57.3	53.4	69.0	54.3	60.6	65.2
Min–Max	2–443	6–404	2–417	7–491	3–370	6–422	2–491	3–404	2–422
History of oophorectomy, n (%)									
Yes	38 (21.7)	36 (20.8)	38 (21.8)	37 (22.2)	34 (20.5)	38 (22.8)	75 (21.9)	70 (20.6)	76 (22.3)
No	137 (78.3)	137 (79.2)	136 (78.2)	130 (77.8)	132 (79.5)	129 (77.2)	267 (78.1)	269 (79.4)	265 (77.7)
History of hysterectomy, n (%)									
Yes	51 (29.1)	60 (34.7)	58 (33.3)	51 (30.5)	53 (31.9)	56 (33.5)	102 (29.8)	113 (33.3)	114 (33.4)
No	124 (70.9)	113 (65.3)	116 (66.7)	116 (69.5)	113 (68.1)	111 (66.5)	240 (70.2)	226 (66.7)	227 (66.6)

The full analysis set (FAS) comprised all patients who were randomised and received one or more doses of study drug; the randomised treatment for each patient was used for summaries by treatment group based on the FAS, even if a patient erroneously received a different treatment. [a]Current versus former/never smoking status was a stratification factor for randomisation. [b]Data on the suitability of HRT and the proportion of patients in each HRT-unsuitable category were only reported for the pooled analysis of SKYLIGHT 1 and SKYLIGHT 2. [c]The denominator is number of patients who have previously been treated with HRT; a patient can have more than one reason for stopping HRT. [d]The denominator is number of patients who have not been previously treated with HRT. [e]The denominator is number of patients who have been advised not to take HRT and the reason is not missing; patients can have an underlying medical condition and a family history of breast cancer.

Abbreviations: BMI: body mass index; HRT: hormone replacement therapy; Max: maximum; Min: minimum; VMS: vasomotor symptoms; SD: standard deviation.

Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs.[83, 84]

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

B.2.4.1 Trial populations

B.2.4.1.1 DAYLIGHT

The Full Analysis Set (FAS) was used to assess efficacy. The Safety Analysis Set (SAF) was used for safety-related analyses. In both analysis sets, the key population analysed included all patients who were randomised and received at least one dose of study treatment.

The description and number of patients in each analysis population as reported in the DAYLIGHT trial are presented in Table 12.

Of the 453 patients randomised in this study, 1 patient in the fezolinetant 45 mg group did not take the study treatment and was therefore not included in the FAS or SAF. A total of 452 patients took at least 1 dose of study treatment and were included in the FAS and SAF.

Table 12: Trial populations used for the analysis of outcomes in the DAYLIGHT trial

Analysis set		DAYLIGHT
Enrolled[a]	Description	• All patients who signed the informed consent form.
	N	N=821
Randomised	Description	• All patients who were randomised. Patients were analysed according to the study treatment group to which they were randomised.
	N	N=453
FAS	Description	• All patients who were randomised and received at least one dose of study drug • Patients were analysed according to the study treatment group to which they were randomised
	N	N=452
SAF	Description	• All patients who were randomised and received at least one dose of the study drug • Patients were analysed according to the study drug they actually received. A patient who erroneously received a treatment different from their randomised treatment were assigned to the treatment group that the patient received as their first dose
	N	N=452

[a]Patients who were screened multiple times were counted for each screening attempt; 785 unique patients signed the informed consent form.
Abbreviations: FAS: full analysis set; SAF: Safety Analysis Set.
Source: Astellas (Data on File). DAYLIGHT Study Protocol and CSR.[79, 82]

B.2.4.1.2 SKYLIGHT 1 and SKYLIGHT 2

The definitions of the study populations analysed in the SKYLIGHT 1 and SKYLIGHT 2 trials are presented in Table 13.

In SKYLIGHT 1, out of the 527 randomised patients, 522 took at least one dose of study treatment and were included in both the FAS and SAF. Five patients did not take the study treatment. Eighty-seven patients were excluded from the PPS at Week 4 and 158 at Week 12. In total, 499 patients received at least one dose of fezolinetant during the 12-week double-blind or active treatment extension periods and were included in the FAS and SAF.

A total of 501 patients were randomised in SKYLIGHT 2. Of these, 500 received study treatment and were included in the FAS and SAF. One patient did not take the study treatment. Among the 500 patients, 81 were excluded from the PPS at Week 4 and 139 at Week 12. A total of 484 patients took at least one dose of fezolinetant during either the 12-week double-blind or active treatment extension periods and were included in the FAS and SAF.

The number of patients in each analysis set stratified by treatment arm during the 12-week double-blind period and the 52-week period of SKYLIGHT 1 and SKYLIGHT 2 are presented in Appendix M.

Table 13: Trial populations used for the analysis of endpoints in SKYLIGHT 1 and SKYLIGHT 2

Analysis set		SKYLIGHT 1	SKYLIGHT 2
Enrolled[a]	Description	• All patients who signed the informed consent form	
	N	N=2,404	N=1,676
12-week double-blind period			
FAS	Description	• The FAS comprised all patients who were randomised and received at least one dose of study drug, even if a patient erroneously received a different treatment • This was the primary analysis set for efficacy analyses	
	N	N=522[b]	N=500
PPS	Description	• The PPS excluded patients from the FAS that had no measurement of the primary efficacy endpoint at Week 4 or Week 12, <85% interactive diary compliance or <85% treatment compliance • There were two PPS populations: one specific to the co-primary analyses at Week 4 and another specific to Week 12	
	N	Week 4: N=440 Week 12: N=369	Week 4: N=419 Week 12: N=361
SAF	Description	• The SAF comprised all randomised patients who received at least one dose of study drug • A patient erroneously receiving a treatment different from their randomised treatment will be assigned to the treatment group that the patient received as their first dose	
	N	N=522[a]	N=500
52-week period			
FAS	Description	• As above	
	N	N=499	N=484
SAF	Description	• As above	
	N	N=499	N=484

[a]Patients who were screened multiple times were counted for each screening attempt; 2,205 and 1,662 unique patients signed the informed consent form for SKYLIGHT 1 and SKYLIGHT 2, respectively. [b]In SKYLIGHT 1, one patient randomised to the fezolinetant 45 mg group was incorrectly given fezolinetant 30 mg, which occurred during the first four-week treatment dispense. This patient was considered as part of the fezolinetant 45 mg group in the efficacy analyses and as part of the fezolinetant 30 mg group in the safety analyses.
Abbreviations: FAS: full analysis set, PPS: Per Protocol Set; SAF: Safety Analysis Set.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 Study Protocols and CSRs.[80, 81, 83, 84]

B.2.4.2 Statistical methods

B.2.4.2.1 DAYLIGHT

A summary of the statistical analyses used to analyse the study endpoints at Week 24, alongside sample size calculations and methods for handling missing data, is presented in Table 14.

Table 14: Statistical methods for the analysis of the DAYLIGHT trial

Hypothesis objective	All null hypotheses were of no treatment difference, and all alternative hypotheses were two-sided
Statistical analysis	**Analysis of efficacy** • The efficacy analysis was conducted on the FAS and the interpretation of results from statistical tests was based on the FAS • All statistical comparisons were conducted using two-sided tests at an alpha level of 0.05 unless specifically stated otherwise • The key secondary endpoint of mean change in the severity of moderate to severe VMS from baseline to week 24 will be considered statistically significant if the primary endpoint of mean change in the frequency of moderate to severe VMS from baseline to week 24 from the primary analysis is statistically significant • No additional adjustments for multiplicity were made for this study **Primary estimand** The estimand of the primary objective used a treatment policy strategy and was defined by the following five attributes: • **Treatment:** Fezolinetant 45 mg and placebo • **Population:** Women patients ≥40 years and ≤65 years of age with moderate to severe VMS associated with menopause and deemed unsuitable for HRT, as defined by the inclusion/exclusion criteria of the study • **Primary Endpoint:** Change in the average daily frequency of moderate to severe VMS from baseline to Week 24 • **Intercurrent events and their corresponding strategies:** All data from the 24-week treatment period were included, that is, treatment policy strategy that considered data regardless of patient discontinuation from the treatment prior to Week 24 • **Population level summary:** Difference in means of change from baseline **Primary endpoint analysis** • The primary objective was to evaluate the efficacy of fezolinetant 45 mg versus placebo on the frequency of moderate to severe VMS. The primary objective was investigated through the mean change observed in the frequency of moderate to severe VMS from baseline to Week 24. • For the primary endpoint, the primary statistical analysis was a

mixed models repeated measures analysis (MMRM) assuming missing data mechanism is missing at random (MAR). The data that was used for the treatment policy estimand was based on the 24-week study period, that is, a treatment policy that considered data regardless of patient discontinuation from the treatment prior to Week 24

- An MMRM analysis was used with change in the average daily frequency of moderate to severe VMS from baseline up to Week 24, treatment group, week and smoking status as factors, with baseline weight and baseline value as covariates, as well as an interaction of treatment group by week and baseline value by week. An unstructured covariance structure shared across treatment groups was used to model the within-patient errors. If the algorithm did not converge:
  - If the repeated measures were equally spaced, then Toeplitz was tried first and then compound symmetry as a covariance structure.
  - If the repeated measures were not equally spaced, then spatial power was tried first and then compound symmetry as a covariance structure. This only applied to non-diary secondary and exploratory endpoints.
- The Kenward-Roger approximation was used to estimate denominator degrees of freedom and adjust standard errors (SEs). This analysis used a restricted maximum likelihood–based repeated-measures approach. Least squares (LS) means were estimated using weights proportional to the percentage of women who were randomised as current smokers. From this analysis, the following results were presented:
  - LS mean estimate, SE, and two-sided 95% CI for change from baseline to each treatment week within a treatment group,
  - For comparisons between the fezolinetant treatment group versus placebo:
  - Difference in LS mean estimates (fezolinetant – placebo), SE and corresponding two-sided 95% CI for the change from baseline at each week.
  - The differences in the LS mean estimates were used to obtain two-sided p-value for the difference.

Sensitivity analysis

- Multiple imputation (MI) approach was used to examine the sensitivity of the primary analysis results to departures from the underlying assumption of MAR.
- The analysis of the primary endpoint will be conducted using a simplified MMRM model with treatment group and week as factors, with baseline measurement as a covariate, as well as an interaction of treatment by week and baseline by week.

Key Secondary endpoint analysis

- The key secondary objective was “to evaluate the efficacy of fezolinetant 45 mg versus placebo on the severity of moderate to severe VMS” and was investigated by the mean change from baseline to Week 24, in the severity of moderate to severe VMS.
- The key secondary endpoint was considered to be statistically significant if the primary endpoint of mean change in the frequency of moderate to severe VMS from baseline to Week 24 from the primary analysis was statistically significant.
- VMS Severity was calculated using a weighted average as shown

below:

- For post baseline endpoints:

$$\frac{([\text{number of mild HFs/day} \times 1] + [\text{number of moderate HFs/day} \times 2] + [\text{number of severe HFs/day} \times 3])}{\text{Total number of daily mild/moderate/severe HFs}}$$

- For baseline:

$$\frac{([\text{number of moderate HFs/day} \times 2] + [\text{number of severe HFs/day} \times 3])}{\text{Total number of daily moderate/severe HFs}}$$

- VMS severity at baseline did not include mild VMS events in the numerator or denominator
- An MMRM analysis was performed on the change in the severity of moderate to severe VMS from baseline up to Week 24
- The treatment policy estimand, as discussed earlier, was used for all key secondary endpoint analyses

Other Secondary endpoints (continuous)

- Secondary continuous endpoints included:
  - Mean change in the patient-reported sleep disturbance by the PROMIS SD SF 8b total score from baseline to Week 24
  - Mean change in the frequency of moderate to severe VMS from baseline to Weeks 1, 4, 8, 12, 16 and 20
  - Mean change in the severity of moderate to severe VMS from baseline to Weeks 1, 4, 8, 12, 16 and 20
  - Mean percent change in the frequency of moderate and severe VMS from baseline to Weeks 1, 4, 8, 12, 16, 20 and 24
- An MMRM analysis was performed on the above endpoints with change from baseline in these endpoints up to Week 24 as response.
- The treatment policy estimand, as discussed earlier, was used for all key secondary endpoint analyses

Other Secondary endpoints (responder)

- Responder endpoint: Responder of percent reduction ≥ 50%, ≥ 75% and at 100% in the frequency of moderate and severe VMS from baseline to Weeks 1, 4, 8, 12, 16, 20 and 24
- For each of the responder endpoints at each week, logistic regression was used for the analysis with treatment group and smoking status (current versus. former/never) as factors and baseline measurement (mean frequency of VMS) as a covariate. Comparisons between the fezolinetant and placebo was calculated based on the odds ratio using a 2 sided 95% CI and p-values

Exploratory endpoints

- MENQOL total and domain scores, EQ-5D domain scores, WPAI-VMS and PHQ-4 total score was assessed using an MMRM model as described for the primary endpoints above
- PGIC-C in VMS was analysed using Cochran-Mantel-Haenszel tests
- The treatment policy estimand, as discussed earlier, was used for all exploratory endpoint analyses

Sample size, power calculation	• Approximately 440 patients in total were planned to be randomised, 220 patients in the fezolinetant 45 mg treatment group and 220 patients in the placebo group at a 1:1 ratio • For a pairwise comparison of the primary endpoint using a two-sample t-test at a two-sided 5% alpha, it was assumed that 220 patients in each group would provide at least 80% power to detect a difference from placebo of −1.8, assuming a SD of 5.6. This size was calculated assuming approximately 30% of patients may discontinue the study prematurely • For a pairwise comparison of the key secondary endpoint using a 2-sample t-test at a two-sided 5% alpha, estimated power was assumed to 31% and 59% with 220 patients in each group to detect a difference from placebo of −0.2 and −0.3 respectively, assuming a SD of 1.2. The alpha was only propagated to the key secondary endpoint if the primary endpoint was statistically significant at a 5% significance level
Data management, patient withdrawals	**Missing data** • Generally, the mechanism of missing data was assumed to be MAR. There was no explicit imputation of missing data for the primary analysis • To evaluate the robustness of the primary analysis results versus departure from the underlying MAR assumption, a sensitivity analysis was conducted using a discontinuation-reason based multiple imputation method. Specifically, a jump-to-reference (that is, placebo) method was used to impute the missing data from patients who discontinued active treatment due to treatment-emergent AEs by assuming that the treatment benefits are diminished after discontinuation of the treatment • The start and stop dates of AEs and concomitant medication were imputed. The imputed dates were used to allocate the concomitant medication and AEs to a treatment group, in addition to determining whether an AE was or was not treatment emergent

Abbreviations: AE: adverse event; EQ-5D: EuroQoL five dimensions; FAS: Final Analysis Set; MAR: missing at random; MENQOL: Menopause Quality of Life Questionnaire; MMRM: mixed model for repeated measures; PGI-C VMS: Patient Global Impression of Change in Vasomotor Symptoms; PHQ-4: Patient Health Questionnaire for Anxiety and Depression; PPS: Per Protocol Set; PROMIS SD SF 8b: Patient-reported Outcomes Measurement Information System Sleep Disturbance – Short Form 8b; SAF: Safety Analysis Set; SD: standard deviation; VMS: vasomotor symptoms; WPAI-VMS: Work Productivity and Activity Impairment questionnaire specific to VMS.
Source: Astellas (Data on File). DAYLIGHT Study Protocol.[79]

B.2.4.2.2 SKYLIGHT 1 and SKYLIGHT 2

The statistical analyses used to analyse the four co-primary endpoints at Week 4 and Week 12, alongside sample size calculations and methods for handling missing data, are presented in Table 15.

Table 15: Statistical methods for the analysis of the SKYLIGHT 1 and SKYLIGHT 2 trials

Hypothesis objective	**Co-primary endpoints** The hypothesis for each pairwise comparison is given as follows: • H_0: The change from baseline at Week 4 (or 12) for fezolinetant and placebo are the same • H_1: The change from baseline at Week 4 (or 12) for fezolinetant and placebo are not the same
Statistical analysis	**Analysis of efficacy** • The efficacy analysis was conducted on the FAS and PPS. The

	interpretation of results from statistical tests was based on the FAS. The PPS was used to assess the robustness of the results from the statistical tests based on the FAS (sensitivity analysis) • All statistical comparisons were conducted using two-sided tests at an alpha level of 0.05 unless specifically stated otherwise. Additional testing details are provided below regarding the four co-primary endpoints. • All null hypotheses were of no treatment difference, all alternative hypotheses were two-sided and controlled for two comparisons with placebo using the Hochberg approach, unless specifically stated otherwise • Assessments after the 12-week placebo-controlled period are descriptive only, because there is no placebo control **Estimand** • The estimand of the co-primary objective used a hypothetical strategy and compared patients as though they had continued on the assigned treatment **Co-primary endpoint analysis** • The four co-primary efficacy endpoints are the mean change in the frequency of moderate to severe VMS and the mean change in severity of VMS (per 24 hours) from baseline to week 4 and week 12 • For each of the four co-primary efficacy endpoints, a mixed models repeated measures analysis of covariance (MMRM) was used with treatment group, week and smoking status (current versus former/never) as factors, with baseline weight and baseline measurement as covariates, as well as an interaction of treatment by week and an interaction of baseline measurement by week. An unstructured covariance structure shared across treatment groups was used to model the within-patient errors. The Kenward-Roger approximation was used to estimate denominator degrees of freedom and adjust standard errors. This analysis used a restricted maximum likelihood-based repeated-measures approach. The treatment difference was estimated at all study weeks • The MMRM analysis incorporated all available on-treatment data to calculate mean treatment effect estimates without requiring explicit imputation for missing data (that is, for discontinued patients). This approach was consistent with the hypothetical strategy used for the estimand of the primary objective, which aims to estimate a treatment effect as if post-randomisation events that may preclude observation of the primary endpoints had not occurred. In essence, this means comparing patients as though they had continued on their assigned treatment • Comparisons between the fezolinetant and placebo was calculated based on least-squares mean contrasts using a two-tailed 95% confidence interval • The family-wise type I error rate for the two active dose groups compared with placebo for the four co-primary efficacy endpoints were controlled using a Hochberg approach. All four co-primary endpoints must be statistically significant for a given dose to be considered successful. Given that, the largest p-value from the four co-primary endpoints in each dose group will be used to test that dose because it represents the least significant of the co-primary endpoints. If the larger of the two maximum p values, one from each dose, is less than 0.05, then both dose groups are considered

	statistically significant. If not, then if the smaller of the p values is less than 0.025, the second compared dose is considered statistically significant. Otherwise, neither dose is considered statistically significant. The trial will be considered successful if at least one dose is statistically significant • Subgroup analyses of the co-primary endpoints used the MMRM model used for the studies, with the addition of study and subgroup as factors and subgroup-by-treatment, subgroup-by-week and subgroup-by-treatment-by-week as interaction terms. The treatment by subgroup interaction was tested using two-sided alpha level of 0.1 **Secondary analysis** • The PROMIS SD SF 8b scale was analysed using MMRM, similar to the primary analysis of the co-primary endpoints. If all the co-primary endpoints are statistically significant between fezolinetant and placebo at both doses, then the 5% alpha from the analysis of the co-primary endpoints will be passed to the key secondary endpoint as part of the family-wise error rate • For percent reduction in co-primary endpoints at Weeks 4 and 12, an MMRM model was used • For each of the secondary responder endpoints, logistic regression was used for the analysis for each week and endpoint. A missing value will be imputed as a non-responder • PGI-C VMS was analysed using Cochran Mantel Haenszel test **Exploratory endpoints** • EQ-5D domain scores, MENQOL (domain and total scores), WPAI-VMS total score and PROMIS SRI total scores were assessed using an MMRM model as described for the co-primary endpoints above
Sample size, power calculation	• A total of 450 randomised patients was planned: 150 patients in each treatment arm • In a pairwise comparison using a two-sample t-test at a two-sided 2.5% alpha (adjusting for two active doses), 120 patients would provide approximately 79% power to detect a difference from placebo in mean daily frequency change from baseline of −2 assuming a standard deviation (SD) of 5 • In a pairwise comparison using a two-sample t-test at a tw0-sided 2.5% alpha (adjusting for two active doses), 120 patients would provide approximately 79% power to detect a difference from placebo in mean daily severity change from baseline of −0.40 assuming a SD of 1 • Assuming approximately 20% of patients discontinue prematurely, the number of patients was increased from 120 to 150 patients per treatment arm • This sample size would also provide over 95% power to detect a difference of 4.3 from placebo on the key secondary endpoint of the PROMIS SD SF 8b questionnaire, using a two-sample t-test at a two-sided 2.5% alpha (adjusting for two active doses) assuming a SD of 7[89] • The assumptions for this sample size calculation were based on the data from the phase 2b dose ranging study that included seven active doses. Although this study did not include the fezolinetant 45mg once daily dose, the ranges of the treatment effects on the frequency and severity from this study provided the basis for the

	assumptions of the sample size calculation • The combined power for testing all four co-primary endpoints will be lower than the power for each considered individually; however, the co-primary endpoints are correlated (especially the same endpoint at different time points) and this will serve to limit the potential reduction of power
Data management, patient withdrawals	**Data management** • Both trials had two planned database locks, one at the end of 12-week double-blind period, for primary placebo-controlled efficacy and safety, and another lock at 52 weeks. During the conduct of the 40-week active treatment extension period, some data fields from the 12-week period were updated • Pre-planned, sensitivity analyses of selected 12-week endpoints were conducted on the final locked 52-week database, including the co-primary endpoints and key secondary endpoint. These sensitivity analyses were to ensure the robustness of the results at 12 weeks. Therefore, 12-week results include the data from the initial lock while the 52-week results include the data that was updated **Missing data** • Missing data may be the result of missing Week 4 or Week 12 VMS frequency or severity measurements or the result of patients discontinuing treatment prior to Week 4 or Week 12 • Generally, the mechanism of missing data was assumed to be missing at random (MAR). There was no explicit imputation of missing data for the primary analysis • To evaluate the robustness of the primary analysis results versus departure from the underlying MAR assumption, a sensitivity analysis was conducted using a discontinuation-reason based multiple imputation method. Specifically, a jump-to-reference (that is, placebo) method was used to impute the missing data from patients who discontinued active treatment due to treatment-emergent adverse events (AEs) by assuming that the treatment benefits are diminished after discontinuation of the treatment. • The start and stop dates of AEs and concomitant medication were imputed. The imputed dates were used to allocate the concomitant medication and AEs to a treatment group, in addition to determining whether an AE was or was not treatment emergent.

Abbreviations: AE: adverse event; EQ-5D: EuroQoL five dimensions; FAS: Final Analysis Set; MAR: missing at random; MENQOL: Menopause Quality of Life Questionnaire; MMRM: mixed model for repeated measures; PGI-C VMS: Patient Global Impression of Change in Vasomotor Symptoms; PPS: Per Protocol Set; PROMIS SD SF 8b: Patient-reported Outcomes Measurement Information System Sleep Disturbance – Short Form 8b; PROMIS SRI SF 8a: Patient-reported Outcomes Measurement Information System Sleep-Related Impairment – Short Form 8a; SAF: Safety Analysis Set; SD: standard deviation; VMS: vasomotor symptoms; WPAI-VMS: Work Productivity and Activity Impairment questionnaire specific to VMS.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 Study Protocols.[80, 81]

B.2.4.3 Patient flow in the relevant randomised controlled trials

Summaries of patient flow in the DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 trials are presented in Appendix D.

B.2.5 Critical appraisal of the relevant clinical effectiveness evidence

The quality of the DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 trials was assessed using the University of York Centre for Reviews and Dissemination assessment tool. A summary of these assessments is presented in Table 16.

Since the primary manuscript of the DAYLIGHT trial is yet to be published and is currently undergoing peer review, it was not captured in the clinical SLR conducted. Therefore, only the results of the quality assessments for the identified SKYLIGHT 1 and SKYLIGHT 2 trials are presented in Appendix D. An assessment of DAYLIGHT has been provided in Table 16 for completeness based on the DAYLIGHT trial protocol.

Table 16: Quality assessment of the DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 trials

Source of bias	Risk of bias		
	DAYLIGHT	SKYLIGHT 1	SKYLIGHT 2
Was randomisation carried out appropriately?	Low	Low	Low
Was the concealment of treatment allocation adequate?	Low	Low	Low
Were the groups similar at the outset of the study in terms of prognostic factors?	Low	Low	Low
Were the care providers, patients and outcome assessors blind to treatment allocation?	Low	Low	Low
Were there any unexpected imbalances in drop-outs between groups?	Low	Low	Low
Is there any evidence to suggest that the authors measured more outcomes than they reported?	Low	Low	Low
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	ITT	ITT	ITT

Abbreviations: ITT: intention-to-treat.

B.2.6 Clinical effectiveness results of the relevant studies

Summary of clinical effectiveness of fezolinetant

- In DAYLIGHT, treatment with fezolinetant 45 mg demonstrated a statistically significant LS mean (SE) reduction in daily frequency of moderate to severe VMS versus placebo at Week 24 (−1.93 [0.36], p<0.001). Additionally, a statistically significant LS mean (SE) reduction in daily severity of moderate to severe VMS was observed with fezolinetant 45 mg as compared with placebo at Week 24 (−0.39 [0.09], p<0.001).
- In individual SKYLIGHT 1 and SKYLIGHT 2 trials, fezolinetant 45 mg demonstrated a statistically significant LS mean reduction in both frequency and severity of moderate to severe VMS versus placebo at Week 4 (SKYLIGHT 1: −2.07 [0.42], adjusted p=0.007; SKYLIGHT 2: −2.55 [0.46], adjusted p<0.001) and Week 12 (SKYLIGHT 1: −2.55 [0.43], adjusted p=0.007; SKYLIGHT 2: −2.53 [0.55], adjusted p<0.001). The efficacy of fezolinetant 45 mg in alleviating patient-reported sleep disturbance (PROMIS SD SF 8b) was demonstrated in both individual and pooled trials at Week 12.

- The benefit of fezolinetant 45 mg was observed as early as one week after the start of treatment, with continued improvement to Week 4 and sustained benefit through Week 24 for DAYLIGHT and Week 52 for SKYLIGHT 1 and SKYLIGHT 2.
- Improvements in the daily frequency of moderate to severe VMS were broadly consistent across subgroups analyses of the fezolinetant 45 mg group in all three trials. The only exception to this general observation was the pre-specified subgroup analysis of current smokers in DAYLIGHT, the results of which should be interpreted with caution due to the very small sample size.
- The efficacy of fezolinetant 45 mg was also demonstrated in exploratory endpoints measuring quality of life, daily functioning and psychological symptoms (low mood and anxiety), as measured by the MENQOL, EQ-5D-5L, WPAI-VMS and PHQ-4.
- In DAYLIGHT, patients treated with fezolinetant 45 mg reported greater improvements in LS mean (SE) MENQOL (vasomotor domain) (–0.96 [0.22]), EQ-5D-5L VAS (+1.6 [1.5]), WPAI-VMS domain (overall productivity loss: −6.77 [3.21]; activity impairment: −8.38 [2.46]) and PHQ-4 (−0.44 [0.27]) scores versus placebo at Week 24.
- In individual SKYLIGHT 1 and SKYLIGHT 2 trials, fezolinetant 45 mg demonstrated improved LS mean (SE) MENQOL (vasomotor domain) (–1.00 [0.20] and –1.05 [0.21]) and EQ-5D-5L VAS (+3.2 [1.4] and +4.2 [1.7]) scores versus placebo at Week 4. At Week 12, similar results were observed for EQ-5D-5L VAS (3.4 [1.8]) in SKYLIGHT 2 only and MENQOL (vasomotor domain) (–0.99 [0.22] and –0.76 [0.22]) scores for both trials, respectively
- Additionally, in SKYLIGHT 1 and SKYLIGHT 2, improvements were observed in the WPAI-VMS domain scores (overall productivity loss: −10.68 [3.88] and −9.99 [3.95]; activity impairment: −9.52 [2.52] and −6.51 [2.79]) versus placebo at Week 4. Similar trends were observed at Week 12, with improvements demonstrated in WPAI-VMS domain scores for both trials (overall productivity loss: −18.08 [3.97] and −13.92 [4.16]; activity impairment: −8.97 [2.73] and −8.23 [2.85]), respectively.

Given the licensed posology for fezolinetant is 45 mg once daily and consistent with the decision problem for this submission, the clinical effectiveness results of fezolinetant are solely presented for the fezolinetant 45 mg treatment arm of DAYLIGHT, SKYLIGHT 1 and 2 trials.

B.2.6.1 DAYLIGHT

B.2.6.1.1 Primary efficacy endpoint

B.2.6.1.1.1 Frequency of moderate to severe VMS, change from baseline to Week 24

The primary efficacy endpoint in the DAYLIGHT trial was the change from baseline in the daily mean frequency of moderate to severe VMS experienced by patients at Week 24.

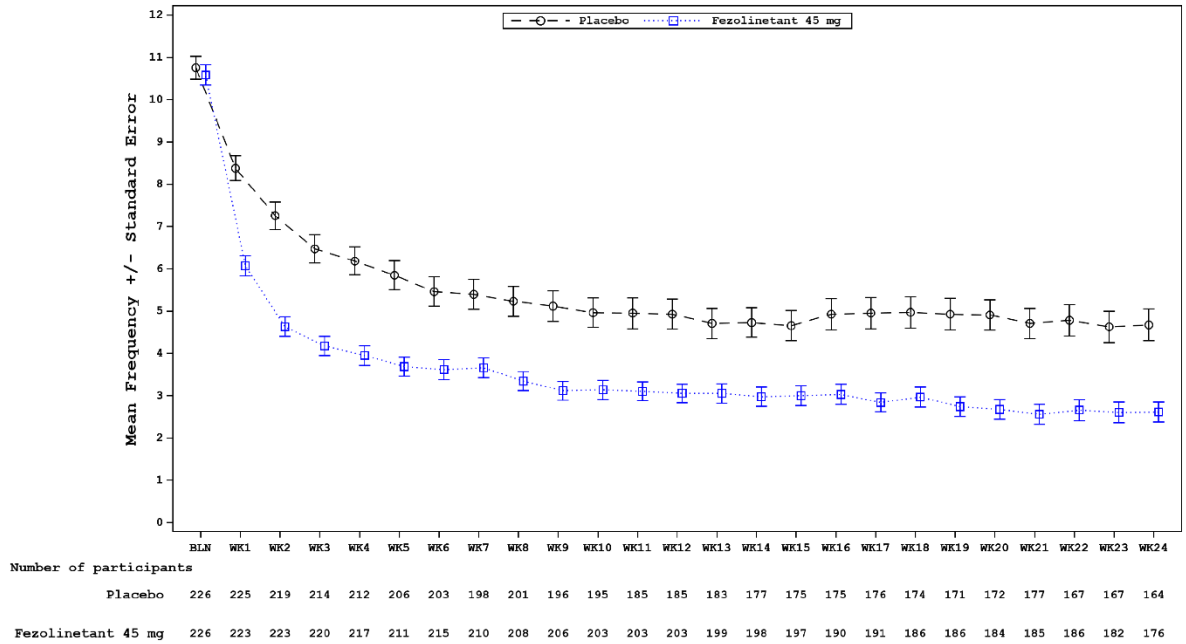
Compared with baseline, patients in the fezolinetant 45 mg arm experienced lower mean daily frequency of moderate to severe VMS versus placebo (−8.15 versus −6.09) at Week 24 (Table 17). The least squares (LS) mean difference between patients receiving fezolinetant 45 mg and placebo (−1.93, standard error [SE]: 0.36) was statistically significant (p<0.001). A reduction in daily moderate to severe VMS frequency was observed as early as Week 1 after treatment onset (unadjusted p<0.001; with observable improvements after the first dose) with continued improvement to Week 4 and a sustained benefit throughout the 24-week double-blind period (Figure 7).

Table 17: Change from baseline to Week 24 in mean frequency of moderate to severe VMS per 24 hours – Final Analysis Set

Analysis visit	Statistics	Placebo (n=226)	Fezolinetant 45 mg (n=226)
Baseline	n	226	226
	Mean (SD)	10.75 (4.08)	10.58 (3.57)
Week 24	n	164	176
	mean (SD)	4.67 (4.80)	2.61 (3.14)
	Change from baseline		
	n	164	176
	Mean (SD)	−6.09 (4.19)	−8.15 (4.43)
	LS mean[a] (SE)	−6.20 (0.26)	−8.13 (0.25)
	Difference in LS Means: fezolinetant versus placebo		
	LS mean[b] (SE)		−1.93 (0.36)
	two-sided unadjusted p value		<0.001

[a]A negative change indicated a reduction from baseline (that is, a favourable outcome). [b]Difference= fezolinetant−placebo.
Abbreviations: LS: least squares; MMRM: mixed models repeated measures; SD: standard deviation; SE: standard error; VMS: vasomotor symptoms.
Source: Astellas (Data on File). DAYLIGHT CSR.[82]

Figure 7: Mean frequency of moderate and severe VMS per 24 hours, from baseline to Week 24 – Final Analysis Set[a,b]



[a]Baseline: arithmetic mean of non-missing daily frequencies of moderate to severe VMS from the last 10 days immediately prior to randomisation as derived by the ePRO device. [b]Each post-baseline timepoint: arithmetic mean of non-missing daily frequencies of moderate to severe VMS, collected within the seven-day analysis visit time window. Mean daily frequency was only derived when information was reported on at least four days within each analysis visit time window.
Abbreviations: BLN: baseline; ePRO: electronic patient reported outcome; VMS: vasomotor symptoms; WK: Week.
Source: Astellas (Data on File). DAYLIGHT CSR.[82]

B.2.6.1.2 Key secondary efficacy endpoint

B.2.6.1.2.1 Severity of moderate to severe VMS, change from baseline to Week 24

The key secondary efficacy endpoint in the DAYLIGHT trial was the change from baseline in the mean severity of moderate to severe VMS experienced by patients at Week 24.

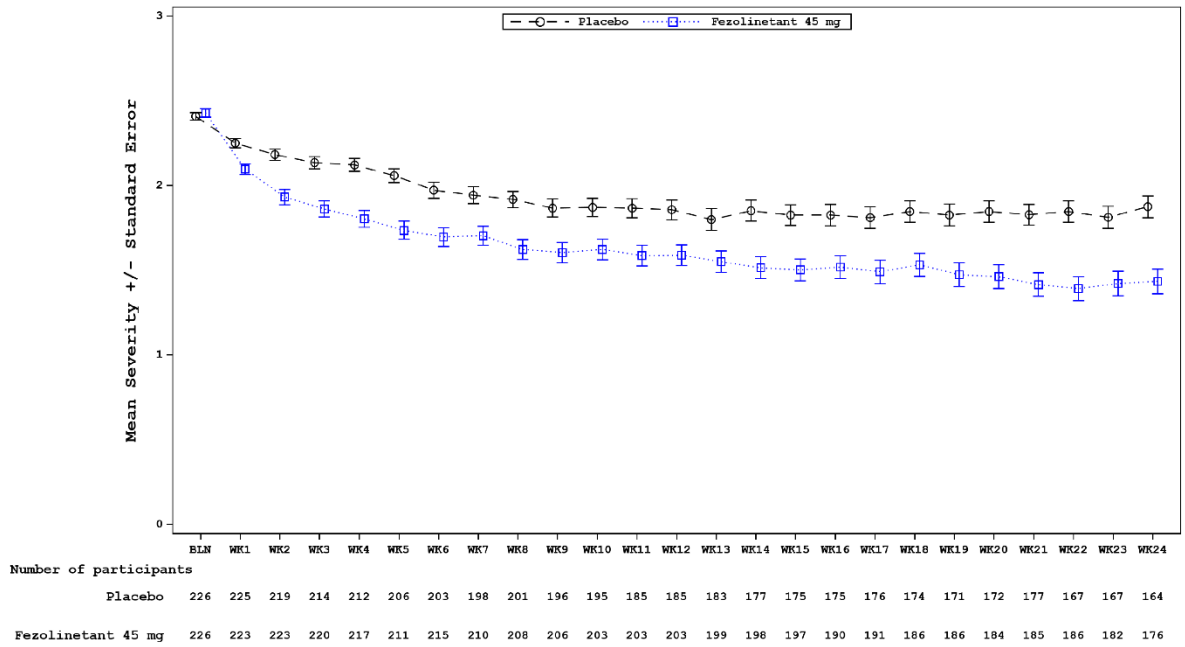
Compared with baseline, patients in the fezolinetant 45 mg arm experienced a greater improvement in their mean daily severity of moderate to severe VMS versus placebo (−0.99 versus −0.54) at Week 24 (Table 18). The LS mean difference between patients receiving fezolinetant 45 mg and placebo (−0.39, SE: 0.09) was statistically significant (p<0.001). A reduction in daily severity of moderate to severe VMS was observed as early as Week 1 after treatment onset (unadjusted p<0.001) with continued improvement to Week 4 and a sustained benefit throughout the 24-week double-blind period (Figure 8).

Table 18: Change from baseline to Week 24 in mean severity of moderate to severe VMS per 24 hours – Final Analysis Set

Analysis visit	Statistics	Placebo (n=226)	Fezolinetant 45 mg (n=226)
Baseline	n	226	226
	Mean (SD)	2.41 (0.34)	2.43 (0.36)
Week 24	n	164	176
	mean (SD)	1.87 (0.82)	1.43 (0.97)
	Change from baseline		
	n	164	176
	Mean (SD)	−0.54 (0.80)	−0.99 (0.97)
	LS mean[a] (SE)	−0.62 (0.06)	−1.01 (0.06)
	Difference in LS Means: fezolinetant versus placebo		
	LS mean[b] (SE)		−0.39 (0.09)
	two-sided unadjusted p value		<0.001

[a]A negative change indicated a reduction from baseline (that is, a favourable outcome). [b]Difference= fezolinetant−placebo.
Abbreviations: LS: least squares; MMRM: mixed models repeated measures; SD: standard deviation; SE: standard error; VMS: vasomotor symptoms.
Source: Astellas (Data on File). DAYLIGHT CSR.[82]

Figure 8: Mean weighted severity of VMS experienced per 24 hours, from baseline to Week 24 – Final Analysis Set[a,b,c]



[a]Baseline: arithmetic mean of non-missing daily weighted average severities of moderate to severe VMS from the last 10 days immediately prior to randomisation. [b]Each post-baseline timepoint: arithmetic mean of non-missing daily weighted average severities of mild, moderate or severe VMS, collected within the seven day analysis visit time window. Mean daily severity per analysis visit was only derived when information was reported on at least four days within each analysis visit time window. [c]Baseline includes moderate and severe incidences, while the post-baseline includes mild, moderate and severe incidences.
Abbreviations: BLN: baseline; VMS: vasomotor symptoms; WK: Week.
Source: Astellas (Data on File). DAYLIGHT CSR.[82]

B.2.6.1.3 Other secondary efficacy endpoints

B.2.6.1.3.1 Change from baseline to Week 24 in mean PROMIS SD SF 8b total score

PROMIS SD SF 8b is an outcome measure that assesses sleep disturbance. Compared with baseline, patients in the fezolinetant 45 mg arm self-reported a lower total score for the sleep deprivation domain of the PROMIS SD SF 8b questionnaire versus placebo (−7.3 versus −4.6) at Week 24 (Table 19), which was indicative of a greater improvement in the levels of sleep deprivation. The LS mean difference between fezolinetant 45 mg and placebo was −2.5 (SE: 0.7; p<0.001).

Table 19: Change from baseline to Week 24 in patient-reported sleep disturbance by the PROMIS SD SF 8b total – Final Analysis Set

Analysis visit	Statistics	Placebo (n=226)	Fezolinetant 45 mg (n=226)
Baseline	n	225	224
	Mean (SD)	27.6 (6.3)	28.3 (6.1)
Week 24	n	178	197
	mean (SD)	22.5 (7.8)	20.9 (7.3)

	Change from baseline		
	n	178	196
	Mean (SD)	−4.6 (8.1)	−7.3 (7.7)
	LS mean[a] (SE)	−4.5 (0.5)	−7.0 (0.5)
	Difference in LS Means: fezolinetant versus placebo		
	LS mean[b] (SE)		−2.5 (0.7)
	two-sided unadjusted p value		<0.001

[a]A negative change indicated a reduction from baseline (that is, a favourable outcome). [b]Difference= fezolinetant−placebo.
Abbreviations: LS: least squares; MMRM: mixed models repeated measures; PROMIS SD SF 8b: Patient-reported Outcomes Measurement Information System Sleep Disturbance – Short Form 8b; SD: standard deviation; SE: standard error; VMS: vasomotor symptoms.
Source: Astellas (Data on File). DAYLIGHT CSR.[82]

B.2.6.1.3.2 Frequency of moderate and severe VMS, percentage change from baseline to Week 24

Compared with baseline, a greater proportion of patients in the fezolinetant 45 mg arm experienced lower mean daily frequency of moderate and severe VMS versus placebo (−75.66% versus −58.16%) at Week 24 (Table 20). The LS mean difference between fezolinetant 45 mg and placebo was −16.55 (SE: 3.26; unadjusted p<0.001).

Table 20: Percentage change from baseline to Week 24 in the mean frequency of moderate and severe VMS per 24 hours – Final Analysis Set

Analysis visit	Statistics	Placebo (n=226)	Fezolinetant 45 mg (n=226)
Baseline	n	226	226
	Mean (SD)	10.75 (4.08)	10.58 (3.57)
Week 24	n	164	176
	mean (SD)	4.67 (4.80)	2.61 (3.14)
	Percentage change from baseline		
	n	164	176
	Mean (SD)	4.67 (4.80)	2.61 (3.14)
	LS mean (SE)	−6.20 (0.26)	−8.13 (0.25)
	95% CI (two-sided)	−6.71, −5.69)	(−8.63, −7.64)
	Difference in LS Means: fezolinetant versus placebo		
	LS mean[a] (SE)		−1.93 (0.36)
	95% CI (two-sided)		−2.64, −1.22
	two-sided unadjusted p value[b]		<0.001

The LS Means, SE, CI, and p values came from an MMRM analysis of covariance model with percent change from baseline as the dependent variable and treatment group, week and smoking status (current versus former/never) as factors, with baseline weight and baseline value as covariates, treatment group by week and baseline value by week as interaction terms. A negative change indicated a reduction from baseline (that is, a favourable outcome).
[a]LS mean differences were calculated by subtracting the LS mean of placebo group from the LS mean of the fezolinetant group. [b]p values presented for the comparison of fezolinetant with placebo from the above described MMRM model.
Abbreviations: CI: confidence intervals; LS: least squares; MMRM: mixed models repeated measures; SD: standard deviation; SE: standard error; VMS: vasomotor symptoms.
Source: Astellas (Data on File). DAYLIGHT CSR.[82]

B.2.6.1.4 Exploratory endpoints

B.2.6.1.4.1 Change from baseline to Week 24 in MENQOL (vasomotor domain)

MENQOL data were collected in the DAYLIGHT trial as an exploratory endpoint. Compared with baseline, patients in the fezolinetant 45 mg arm reported greater reduction (improvements) in MENQOL vasomotor domain score versus placebo (−3.20 versus −2.30) at Week 24 (Table 21). LS mean difference for fezolinetant 45 mg versus placebo was −0.96 (SE: 0.22; unadjusted p<0.001).

Table 21: Change from baseline in MENQOL vasomotor domain scores to Week 24 – Final Analysis Set

Analysis Visit	Statistics	Placebo (n=226)	Fezolinetant 45 mg (n=226)
Baseline	n	223	224
	Mean (SD)	6.52 (1.35)	6.54 (1.31)
Week 24	n	176	196
	Mean (SD)	4.25 (2.28)	3.34 (2.03)
	Change from baseline		
	n	176	196
	Mean (SD)	−2.30 (2.39)	−3.20 (2.24)
	LS mean (SE)	−2.20 (0.16)	−3.16 (0.15)
	95% CI (two−sided)	−2.51, −1.89)	−3.45, −2.87
	Difference in LS Means: fezolinetant versus placebo		
	LS mean[a] (SE)		−0.96 (0.22)
	95% CI (two−sided)		−1.39, −0.54
	P value (two−sided unadjusted)[b]		<0.001

The LS mean, SE, CI, and p values come from an MMRM analysis of covariance model with change from baseline as the dependent variable and treatment group, week and smoking status (current versus former/never) as factors, with baseline weight and baseline value as covariates, treatment group by week and baseline value by week as interaction terms. For the change from baseline in each of the domains, a negative value (decrease from baseline) indicates a better outcome. [a]Differences were calculated by subtracting the LS mean of placebo group from the LS mean of fezolinetant group. [b]p-value is for comparison of Fezolinetant with placebo from the above described MMRM model.
Abbreviations: CI: confidence intervals; LS: least squares; MENQOL: menopause-specific quality-of-life; MMRM: mixed models repeated measures; SD: standard deviation; SE: standard error.
Source: Astellas (Data on File). DAYLIGHT CSR.[82]

B.2.6.1.4.2 Change from baseline to Week 24 in EQ-5D-5L VAS

EQ−5D−5L VAS data were collected in the DAYLIGHT trial as an exploratory endpoint. Compared with baseline, patients in the fezolinetant 45 mg arm reported greater improvements in EQ−5D−5L VAS score versus placebo (+3.3 versus +0.9) at Week 24 (Table 22), with higher scores indicating improved health state. LS mean difference for fezolinetant 45 mg versus placebo was +1.6 (SE: 1.5; unadjusted p=0.307).

Table 22: Change from baseline in EQ−5D−5L VAS Score to Week 24 – Final Analysis Set

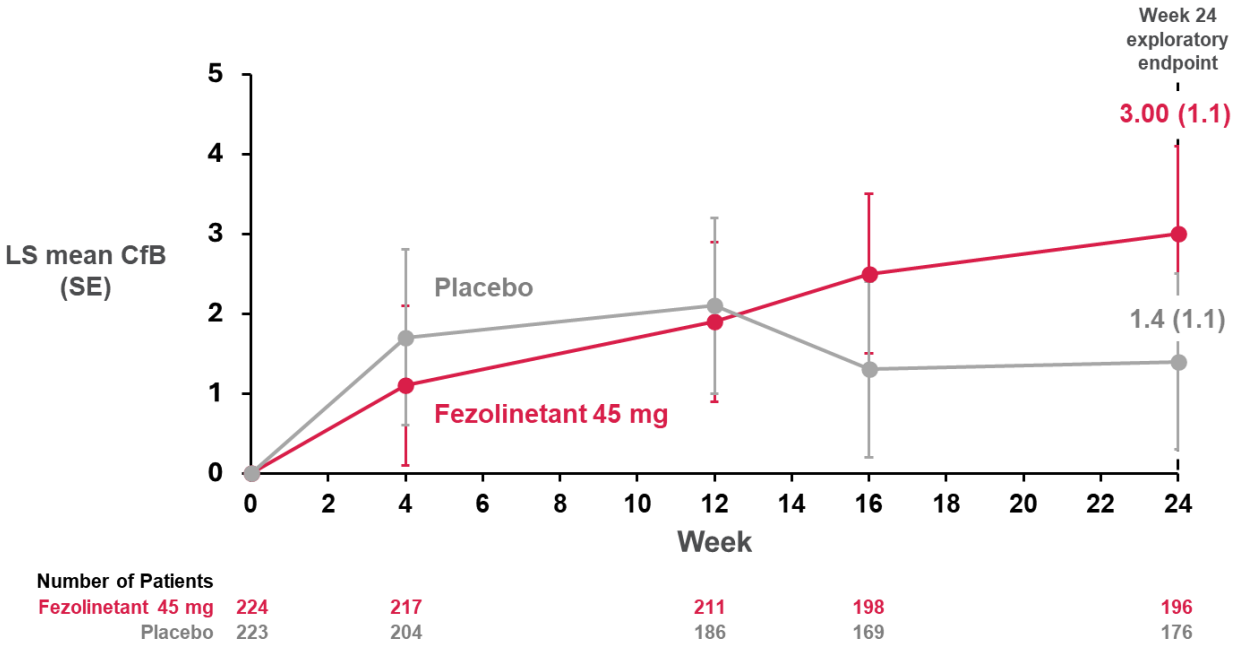
Analysis Visit	Statistic	Placebo (n=226)	Fezolinetant 45 mg (n=226)
Baseline	n	223	224
	Mean (SD)	76.0 (16.9)	74.7 (18.7)
Week 24	n	178	197
	Mean (SD)	78.1 (16.7)	78.7 (18.6)
	Change from baseline		
	n	176	196
	Mean (SD)	+0.9 (17.5)	+3.3 (16.2)
	LS mean (SE)	+1.4 (1.1)	+3.0 (1.1)
	95% CI (two−sided)	−0.8, +3.6	+0.9, +5.1
	Difference in LS Means: fezolinetant versus placebo		
	LS mean[a] (SE)		+1.6 (1.5)
	95% CI (two−sided)		−1.5, +4.6
	P value (two−sided unadjusted)[b]		0.307

The LS mean, SE, CI, and p values come from an MMRM analysis of covariance model with change from baseline as the dependent variable and treatment group, week and smoking status (current versus former/never) as factors, with baseline weight and baseline value as covariates, treatment group by week and baseline value by week as interaction terms. For the change from baseline in VAS, a positive value (increase from baseline) indicates a better outcome, that is, better possible health state. [a]Differences were calculated by subtracting the LS mean of placebo group from the LS mean of fezolinetant group. [b]p-value is for comparison of Fezolinetant with placebo from the above described MMRM model.

Abbreviations: CI: confidence intervals; EQ-5D-5L: EuroQoL five dimensions five level version; LS: least squares; MMRM: mixed models repeated measures; SD: standard deviation; SE: standard error; VAS: visual analogue scale.

Source: Astellas (Data on File). DAYLIGHT CSR.[82]

Figure 9: Change from baseline in EQ-5D-5L VAS Score to Week 24 – Final Analysis Set



Abbreviations: CfB: change from baseline; EQ-5D-5L: EuroQoL five dimensions five level version; SE: standard error; VAS: visual analogue scale.
Source: Astellas (Data on File). DAYLIGHT CSR.[82]

B.2.6.1.4.3 Change from baseline to Week 24 in WPAI-VMS

As presented in Table 23, patients who received fezolinetant 45 mg in the DAYLIGHT trial demonstrated a greater reduction in mean overall productivity loss at Week 24, over patients who received placebo (fezolinetant 45 mg: −23.35, SD: 32.86; placebo: −14.82, SD: 28.66; unadjusted p=0.036). Similar results were reported for the activity impairment domain. Of note, patients in the fezolinetant 45 mg arm reported greater reduced activity impairment compared with placebo (−30.82 versus −24.32) at Week 24 (Table 24).

Table 23: Change from baseline to Week 24 in WPAI−VMS domain (overall productivity loss) scores – Final Analysis Set

Analysis visit	Statistics	Placebo (n=226)	Fezolinetant 45 mg (n=226)
Baseline	n	164	155
	Mean (SD)	41.14 (26.12)	43.25 (27.43)
Week 24	n	127	127
	mean (SD)	23.30 (27.79)	18.02 (23.40)
	Change from baseline		
	n	112	112
	Mean (SD)	−14.82 (28.66)	−23.35 (32.86)
	LS mean (SE)	−15.97 (2.27)	−22.74 (2.26)
	95% CI (two−sided)	−20.44, −11.49	−27.18, −18.29
	Difference in LS Means: fezolinetant versus placebo		
	LS mean[a] (SE)		−6.77 (3.21)
	95% CI (two−sided)		−13.09, −0.45

	two−sided unadjusted p value[b]		0.036

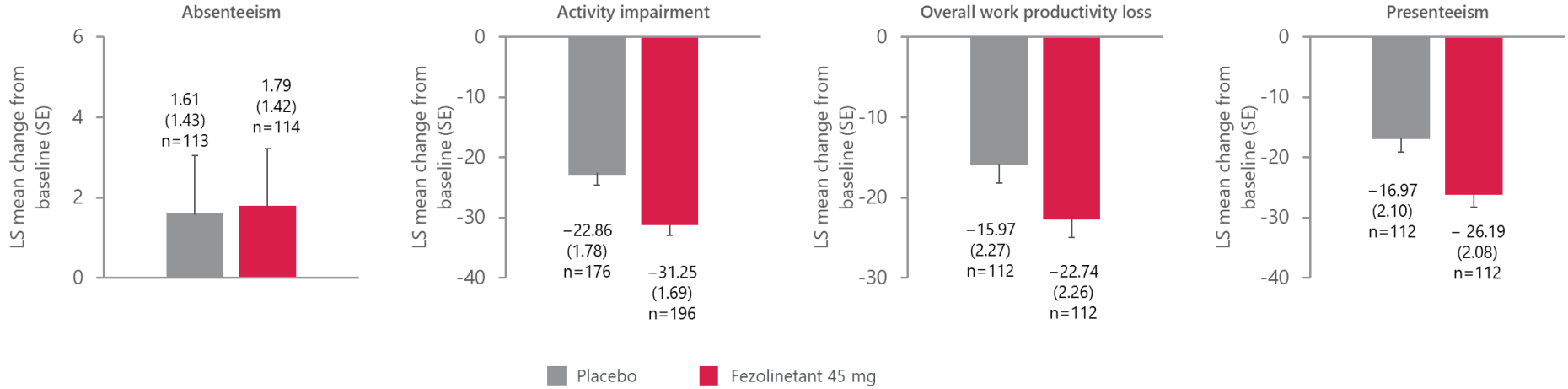
The LS mean, SEs, CIs, and p values came from an MMRM analysis of covariance model with change from baseline as the dependent variable and treatment group, week and smoking status (current versus former/never) as factors, with baseline weight and baseline value as covariates, treatment group by week and baseline value by week as interaction terms. For the change from baseline in each of the domains, a negative value (decrease from baseline) indicates a better outcome, that is, lesser impairment and greater productivity. [a]Differences were calculated by subtracting the LS mean of placebo group from the LS mean of the fezolinetant group. [b]p values were for the comparison of fezolinetant with placebo from the above described MMRM model.
Abbreviations: CI: confidence interval; FAS: full analysis set; LS: least squares; MMRM: mixed models repeated measures; SD: standard deviation; SE: standard error; WPAI−VMS: Work Productivity and Activity Impairment for VMS Questionnaire.
Source: Astellas (Data on File). DAYLIGHT CSR.[82]

Table 24: Change from baseline to Week 24 in WPAI−VMS domain (activity impairment) scores – Final Analysis Set

Analysis visit	Statistics	Placebo (n=226)	Fezolinetant 45 mg (n=226)
Baseline	n	223	224
	Mean (SD)	49.15 (28.83)	47.46 (29.04)
Week 24	n	178	197
	mean (SD)	25.06 (27.99)	15.99 (22.49)
	Change from baseline		
	n	176	196
	Mean (SD)	−24.32 (31.20)	−30.82 (32.10)
	LS mean (SE)	−22.86 (1.78)	−31.25 (1.69)
	95% CI (two−sided)	−26.37, −19.36	−34.57, −27.92
	Difference in LS Means: fezolinetant versus placebo		
	LS mean[a] (SE)		−8.38 (2.46)
	95% CI (two−sided)		−13.22, −3.55
	two−sided unadjusted p value[b]		<0.001

The LS mean, SEs, CIs, and p values came from an MMRM analysis of covariance model with change from baseline as the dependent variable and treatment group, week and smoking status (current versus former/never) as factors, with baseline weight and baseline value as covariates, treatment group by week and baseline value by week as interaction terms. For the change from baseline in each of the domains, a negative value (decrease from baseline) indicates a better outcome, that is, lesser impairment and greater productivity. [a]Differences were calculated by subtracting the LS mean of placebo group from the LS mean of the fezolinetant group. [b]p values were for the comparison of fezolinetant with placebo from the above described MMRM model.
Abbreviations: CI: confidence interval; LS: least squares; MMRM: mixed models repeated measures; SD: standard deviation; SE: standard error; WPAI−VMS: Work Productivity and Activity Impairment for VMS Questionnaire.
Source: Astellas (Data on File). DAYLIGHT CSR.[82]

Figure 10: Change from baseline to Week 24 in WPAI−VMS domain scores – Final Analysis Set



Abbreviations: LS: least square; SE: standard error; WPAI−VMS: Work Productivity and Activity Impairment questionnaire specific to VMS.
Source: Astellas (Data on File). DAYLIGHT CSR.[82]

B.2.6.1.4.4 Change from baseline to Week 24 in PHQ−4 total scores

Compared with baseline, patients in the fezolinetant 45 mg arm self−reported a lower total score for the anxiety and depression domains of the PHQ−4 questionnaire versus placebo (−1.14 versus −0.63) at Week 24 (Table 25), which was indicative of a lower likelihood of suffering from anxiety or depression. The LS mean difference between fezolinetant 45 mg and placebo was −0.44 (SE: 0.27; unadjusted p=0.099).

Table 25: Change from baseline to Week 24 in PHQ−4 total scores – Final Analysis Set

Analysis visit	Statistics	Placebo (n=226)	Fezolinetant 45 mg (n=226)
Baseline	n	223	224
	Mean (SD)	3.43 (3.01)	3.67 (3.27)
Week 24	n	178	197
	mean (SD)	2.72 (3.19)	2.43 (3.15)
	Change from baseline		
	n	176	196
	Mean (SD)	−0.63 (2.93)	−1.14 (3.00)
	LS mean (SE)	−0.67 (0.19)	−1.11 (0.18)
	95% CI (two−sided)	−1.05, −0.28)	−1.47, −0.75
	Difference in LS Means: fezolinetant versus placebo		
	LS mean[a] (SE)		−0.44 (0.27)
	95% CI (two−sided)		−0.97, 0.08)
	two−sided unadjusted p value[b]		0.099

The LS Mean, SE, CI, and p values come from an MMRM analysis of covariance model with change from baseline as the dependent variable and treatment group, week and smoking status (current versus former/never) as factors, with baseline weight and baseline value as covariates, treatment group by week and baseline value by week as interaction terms. For the change from baseline in total score, a negative value (decrease from baseline) indicates a better outcome, that is, less likelihood of anxiety or depression. [a]Differences are calculated by subtracting the LS mean of placebo group from the LS mean of fezolinetant group. [b]p−value is for comparison of fezolinetant with placebo from the above described MMRM model.
Abbreviations: CI: confidence interval; LS: least squares; MMRM: mixed models repeated measures; PHQ−4: Patient Health Questionnaire for Anxiety and Depression; SD: standard deviation; SE: standard error.
Source: Astellas (Data on File). DAYLIGHT CSR.[82]

B.2.6.2 SKYLIGHT 1 and SKYLIGHT 2

B.2.6.2.1 Co-primary efficacy endpoints

B.2.6.2.1.1 Frequency and severity of moderate to severe VMS, change from baseline to Weeks 4 and 12

Across SKYLIGHT 1 and SKYLIGHT 2, fezolinetant 45 mg met all four co−primary endpoints (Table 26, Figure 11, Table 27 and Figure 12). Patients treated with fezolinetant 45 mg had a statistically significant reduction from baseline to Weeks 4 and 12 in both the LS mean daily frequency and severity of moderate to severe VMS as compared with placebo.

Results for the pooled analysis of SKYLIGHT 1 and SKYLIGHT 2 supported the individual study analysis on the co−primary endpoints. Additionally, sensitivity analyses of co−primary endpoints supported the primary analysis of co−primary endpoints (Appendix M).

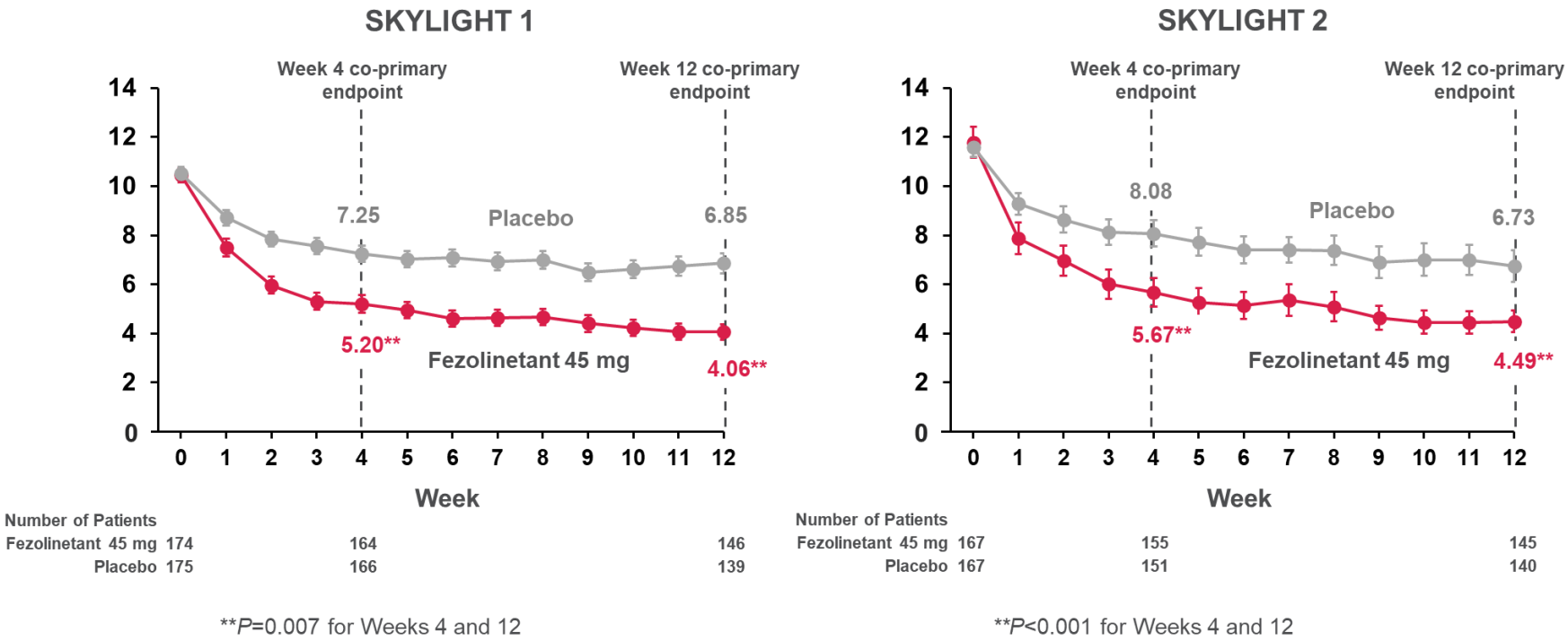
Table 26: Primary analysis of co−primary endpoints: change from baseline in mean frequency of moderate to severe VMS per 24 hours (Full Analysis Set); 12−week double−blind period

Analysis Visit	Statistic	SKYLIGHT 1		SKYLIGHT 2		Pooled studies (SKYLIGHT 1 AND SKYLIGHT 2)	
		Placebo (n=175)	Fezolinetant 45 mg (n=174)	Placebo (n=167)	Fezolinetant 45 mg (n=167)	Placebo (n=342)	Fezolinetant 45 mg (n=341)
Baseline	n	175	174	167	167	342	341
	Mean (SD)	10.51 (3.79)	10.44 (3.92)	11.59 (5.02)	11.79 (8.26)	11.04 (4.46)	11.10 (6.45)
Week 4	n	166	164	151	155	317	319
	Mean (SD)	7.25 (4.29)	5.20 (4.48)	8.08 (6.50)	5.67 (7.29)	7.64 (5.46)	5.43 (6.00)
	Change from baseline[a]						
	n	166	164	151	155	317	319
	Mean (SD)	−3.27 (4.18)	−5.20 (4.07)	−3.64 (4.15)	−6.24 (4.78)	−3.45 (4.16)	−5.70 (4.45)
	LS mean (SE)	−3.32 (0.29)	−5.39 (0.30)	−3.72 (0.33)	−6.26 (0.33)	−3.51 (0.22)	−5.79 (0.23)
	95% CI (two−sided)	−3.89, −2.74	−5.97, −4.81	−4.36, −3.07	−6.90, −5.62	−3.95, −3.07	−6.23, −5.35
	Difference in LS Means: fezolinetant versus placebo[b]						
	LS mean (SE)		−2.07 (0.42)		−2.55 (0.46)		−2.28 (0.32)
	95% CI (two−sided)		−2.89, −1.25		−3.45, −1.64		−2.91, −1.66
	P value (two−sided unadjusted)[c]		<0.001		<0.001		<0.001[d]
	P value (two−sided adjusted)[e]		0.007**		<0.001**		NA
Week 12	n	139	146	140	145	279	291
	Mean (SD)	6.85 (4.66)	4.06 (3.85)	6.73 (7.58)	4.49 (5.39)	6.79 (6.28)	4.27 (4.68)
	Change from baseline[a]						

	n	139	146	140	145	279	291
	Mean (SD)	−3.67 (4.18)	−6.38 (4.48)	−4.57 (5.14)	−7.43 (6.47)	−4.12 (4.70)	−6.90 (5.58)
	LS mean (SE)	−3.90 (0.31)	−6.44 (0.31)	−4.97 (0.39)	−7.50 (0.39)	−4.43 (0.25)	−6.94 (0.25)
	95% CI (two−sided)	−4.50, −3.29	−7.04, −5.84	−5.73, −4.20	−8.25, −6.74	−4.92, −3.94	−7.42, −6.45
	Difference in LS Means: fezolinetant versus placebo[b]						
	LS mean (SE)		−2.55 (0.43)		−2.53 (0.55)		−2.51 (0.35)
	95% CI (two−sided)		−3.40, −1.70		−3.60, −1.46		−3.20, −1.82
	P value (two−sided unadjusted)[c]		<0.001		<0.001		<0.001[d]
	P value (two−sided adjusted)[e]		0.007**		<0.001**		NA

All patients who were randomised and received at least one dose of study treatment (Full Analysis Set). The LS means, SE, CI and p values come from an MMRM analysis of covariance model with change from baseline as the dependent variable and treatment group, week (Week 1 to Week 12) and smoking status (current versus former/never) as factors, with baseline measurement and baseline weight as covariates, as well as an interaction of treatment by week and an interaction of baseline measurement by week. For the pooled data, the same MMRM analysis of covariance model used study protocol as an additional factor. [a]A negative change indicated a reduction/improvement from baseline (that is, a favourable outcome). [b]Differences were calculated by subtracting the LS mean of placebo group from the LS mean of fezolinetant group.[c]P value was for comparison of fezolinetant with placebo from the above described MMRM model. [d]For the integrated summary of efficacy, it was not necessary to conduct a multiplicity adjustment as the statistical significance of the primary endpoints were conducted in each pivotal studies with the multiplicity adjustment. [e]Largest p value within each dose compared to placebo.
**Statistically significant adjusting for multiplicity using the Hochberg procedure at the 5% significance level (statistical significance of the four co−primary endpoints).
Abbreviations: CI: confidence interval; LS: least squares; MMRM: mixed model for repeated measures; NA: not applicable; NR: not reported; VMS: vasomotor symptoms.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs and ISE.[83-85]

Figure 11: Primary analysis of co-primary endpoints: change from baseline in mean frequency of moderate to severe VMS per 24 hours (Full Analysis Set); 12-week double-blind period



Baseline: arithmetic mean of non-missing daily frequencies of moderate to severe vasomotor symptoms from the last 10 days immediately prior to randomisation as derived by the ePRO device. Each post-baseline timepoint: arithmetic mean of non-missing daily frequencies of moderate to severe vasomotor symptoms, collected within the seven-day analysis visit time window. Mean daily frequency has only been derived when information has been reported on at least four days within each analysis visit time window.
**Statistically significant adjusting for multiplicity using the Hochberg procedure at the 5% significance level (statistical significance of the four co-primary endpoints).
Abbreviations: ePRO: electronic patient-reported outcome; SE: standard error; VMS: vasomotor symptoms.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs.[83, 84]

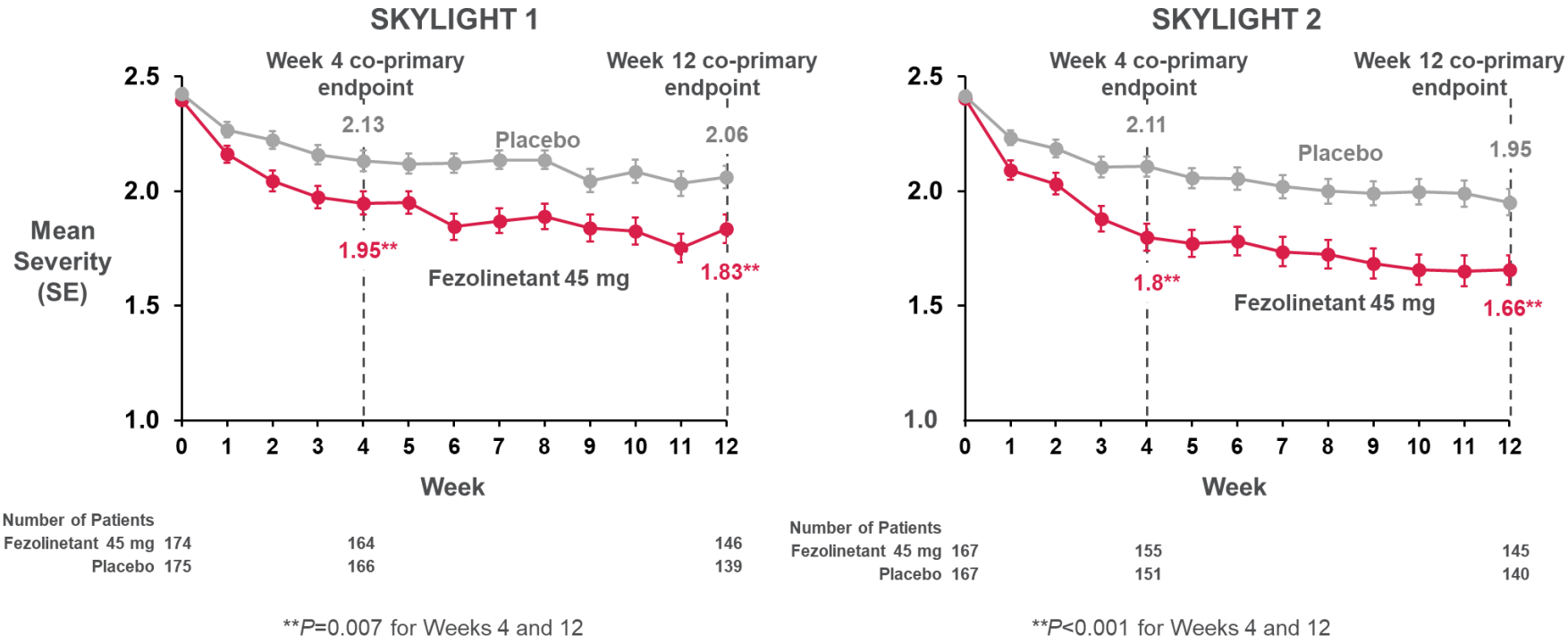
Table 27: Primary analysis of co−primary endpoints: change from baseline in mean severity of moderate to severe VMS per 24 hours (Full Analysis Set); 12−week double−blind period

Analysis Visit	Statistic	SKYLIGHT 1		SKYLIGHT 2		Pooled studies (SKYLIGHT 1 AND SKYLIGHT 2)	
		Placebo (n=175)	Fezolinetant 45 mg (n=174)	Placebo (n=167)	Fezolinetant 45 mg (n=167)	Placebo (n=342)	Fezolinetant 45 mg (n=341)
Baseline	n	175	174	167	167	342	341
	Mean (SD)	2.43 (0.35)	2.40 (0.35)	2.41 (0.32)	2.41 (0.34)	2.42 (0.34)	2.40 (0.35)
Week 4	n	166	164	151	155	317	319
	Mean (SD)	2.13 (0.58)	1.95 (0.64)	2.11 (0.56)	1.80 (0.74)	2.12 (0.57)	1.88 (0.70)
	Change from baseline[a]						
	n	166	164	151	155	317	319
	Mean (SD)	−0.28 (0.50)	−0.45 (0.61)	−0.31 (0.48)	−0.61 (0.63)	−0.29 (0.49)	−0.53 (0.62)
	LS mean (SE)	−0.27 (0.04)	−0.46 (0.04)	−0.32 (0.05)	−0.61 (0.05)	−0.30 (0.03)	−0.53 (0.03)
	95% CI (two−sided)	−0.35, −0.19	−0.54, −0.37	−0.41, −0.23	−0.70, −0.52	−0.36, −0.24	−0.59, −0.47
	Difference in LS Means: fezolinetant versus placebo[b]						
	LS mean (SE)		−0.19 (0.06)		−0.29 (0.06)		−0.24 (0.04)
	95% CI (two−sided)		−0.30, −0.07		−0.41, −0.16		−0.32, −0.15
	P value (two−sided unadjusted)[c]		0.002		<0.001		<0.001[d]
	P value (two−sided adjusted)[e]		0.007**		<0.001**		NA
Week 12	n	139	146	140	145	274	286
	Mean (SD)	2.06 (0.59)	1.83 (0.75)	1.95 (0.68)	1.66 (0.79)	2.01 (0.64)	1.75 (0.77)
	Change from baseline[a]						

	n	139	146	140	145	274	286
	Mean (SD)	−0.35 (0.58)	−0.58 (0.75)	−0.46 (0.65)	−0.74 (0.71)	−0.40 (0.61)	−0.65 (0.73)
	LS mean (SE)	−0.37 (0.05)	−0.57 (0.05)	−0.48 (0.06)	−0.77 (0.06)	−0.43 (0.04)	−0.67 (0.04)
	95% CI (two−sided)	−0.47, −0.26	−0.67, −0.47	−0.59, −0.36	−0.88, −0.65	−0.50, −0.35	−0.74, −0.59
	Difference in LS Means: fezolinetant versus placebo[b]						
	LS mean (SE)		−0.20 (0.08)		−0.29 (0.08)		−0.24 (0.06)
	95% CI (two−sided)		−0.35, −0.06		−0.45, −0.13		−0.35, −0.13
	P value (two−sided unadjusted)[c]		0.007		<0.001		<0.001[d]
	P value (two−sided adjusted)[e]		0.007**		<0.001**		NA

Baseline includes moderate to severe incidences. Post−baseline includes mild, moderate and severe incidences. All patients who were randomised and received at least one dose of study treatment (Full Analysis Set). The LS means, SE, CI and p values come from an MMRM analysis of covariance model with change from baseline as the dependent variable and treatment group, week (Week 1 to Week 12) and smoking status (current versus former/never) as factors, with baseline measurement and baseline weight as covariates, as well as an interaction of treatment by week and an interaction of baseline measurement by week. For the pooled data, the same MMRM analysis of covariance model used study protocol as additional factor. [a]A negative change indicated a reduction/improvement from baseline (that is, a favourable outcome). [b]Differences were calculated by subtracting the LS mean of placebo group from the LS mean of fezolinetant group.[c]P value was for comparison of fezolinetant with placebo from the above described MMRM model. [d]For the integrated summary of efficacy, it was not necessary to conduct multiplicity adjustment as the statistical significance of the primary endpoints were conducted in each pivotal studies with the multiplicity adjustment. [e]Largest p value within each dose compared with placebo.
**Statistically significant adjusting for multiplicity using the Hochberg procedure at the 5% significance level (statistical significance of the four co−primary endpoints).
Abbreviations: CI: confidence interval; LS: least squares; MMRM: mixed model for repeated measures; NA: not applicable; VMS: vasomotor symptoms.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs and ISE.[83-85]

Figure 12: Primary analysis of co−primary endpoints: change from baseline in mean severity of moderate to severe VMS per 24 hours (Full Analysis Set); 12−week double−blind period



Baseline: arithmetic mean of non−missing daily weighted average severities of moderate to severe vasomotor symptoms, from the last 10 days immediately prior to randomisation. Each post−baseline timepoint: arithmetic mean of non−missing weighted average severities of mild, moderate or severe vasomotor symptoms collected within the seven−day analysis visit time window. Mean daily severity per analysis visit has only been derived when information has been reported on at least four days within each analysis visit time window. Baseline includes moderate and severe while the post−baseline includes mild, moderate and severe incidences.
**Statistically significant adjusting for multiplicity using the Hochberg procedure at the 5% significance level (statistical significance of the four co−primary endpoints).
Abbreviations: SE: standard error; VMS: vasomotor symptoms.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs.[83, 84]

B.2.6.2.2 Key secondary efficacy endpoint

B.2.6.2.2.1 PROMIS SD SF 8b total score, change from baseline to Week 12

PROMIS SD SF 8b is an outcome measure that assesses sleep disturbance.

In SKYLIGHT 1 and SKYLIGHT 2, there was a numerically greater reduction from baseline to Week 12 in PROMIS SD SF 8b total score in the fezolinetant 45 mg groups relative to placebo (Table 28), indicating the beneficial effect of fezolinetant 45 mg on sleep disturbance. However, statistical significance of the LS mean difference between fezolinetant 45 mg and placebo was only met in SKYLIGHT 2. While statistical significance was not achieved in SKYLIGHT 1, this might be attributed, at least in part, to the fact that disturbed sleep was not a prerequisite for study inclusion, which could have influenced these results.[90] There were similar numerical decreases in PROMIS SD SF 8b total score in the pooled analysis, with an observed reduction from baseline in the fezolinetant 45 mg group relative to placebo.

B.2.6.2.3 Other secondary efficacy endpoints

B.2.6.2.3.1 Frequency and severity of moderate to severe VMS, change from baseline to each week up to Week 12

In the pivotal studies, the reduction in the frequency and severity of moderate to severe VMS was evident as early as one week after the start of treatment for the fezolinetant 45 mg groups and was maintained throughout the 12-week placebo-controlled period (Figure 13 and Figure 14). These improvements were greater for patients in the fezolinetant 45 mg groups compared with placebo.

The results of these analyses support the primary analysis conclusions for the co-primary endpoint analyses in the pivotal studies.

B.2.6.2.3.2 Frequency of moderate to severe VMS, percent change from baseline to each week up to Week 12

In the pivotal studies, patients treated with fezolinetant 45 mg had a greater percent reduction from baseline to Weeks 4 and 12 in the frequency of moderate to severe VMS relative to placebo (Table 29). These improvements in VMS frequency were observed as early as one week after the start of treatment and were maintained during the 12-week double-blind period. Patients in the fezolinetant 45 mg group had a LS mean decrease of −21.31% from baseline at Week 4 and a −21.48% decrease at Week 12 relative to placebo in the pooled analysis.

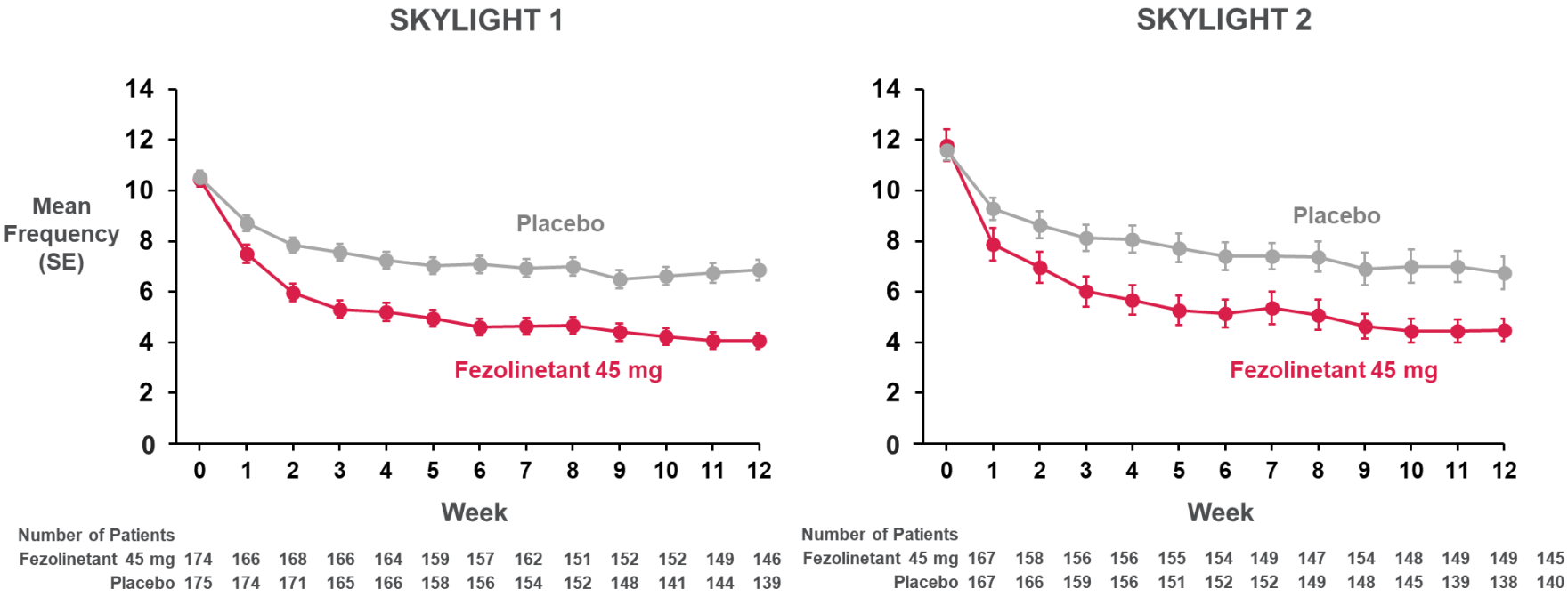
The percent change from baseline analysis over time for VMS frequency is concordant with the co-primary endpoint analyses, providing further evidence of a reduction in VMS frequency over the 12-week double-blind period.

Table 28: Change from baseline in PROMIS Sleep Disturbance – Short Form 8b (Full Analysis Set); 12−week double−blind period

Analysis Visit	Statistic	SKYLIGHT 1		SKYLIGHT 2		Pooled studies (SKYLIGHT 1 and SKYLIGHT 2)	
		Placebo (n=175)	Fezolinetant 45 mg (n=174)	Placebo (n=167)	Fezolinetant 45 mg (n=167)	Placebo (n=342)	Fezolinetant 45 mg (n=341)
Baseline	n	175	174	166	167	341	341
	Mean (SD)	26.4 (6.6)	27.1 (7.0)	27.4 (7.0)	26.2 (6.6)	26.9 (6.8)	26.7 (6.8)
Week 12	n	148	156	144	145	292	301
	Mean (SD)	23.1 (7.4)	22.3 (7.3)	23.8 (7.0)	21.2 (5.7)	23.4 (7.2)	21.8 (6.6)
	Change from Baseline[a]						
	n	148	156	143	145	291	301
	Mean (SD)	−3.2 (7.3)	−4.6 (7.3)	−3.6 (7.3)	−4.8 (6.8)	−3.4 (7.3)	−4.7 (7.0)
	LS mean (SE)	−3.2 (0.5)	−4.2 (0.5)	−3.4 (0.5)	−5.5 (0.5)	−3.3 (0.4)	−4.8 (0.4)
	95% CI (two−sided)	−4.2, −2.1	−5.3, −3.2	−4.5, −2.4	−6.5, −4.4	−4.0, −2.6	−5.5, −4.1
	Difference in LS Means: fezolinetant versus placebo[b]						
	LS mean (SE)		−1.1 (0.7)		−2.0 (0.7)		−1.5 (0.5)
	95% CI (two−sided)		−2.5, 0.4		−3.5, −0.6		−2.5, −0.5
	P value (two−sided unadjusted)[c]		0.155		0.007**		0.004

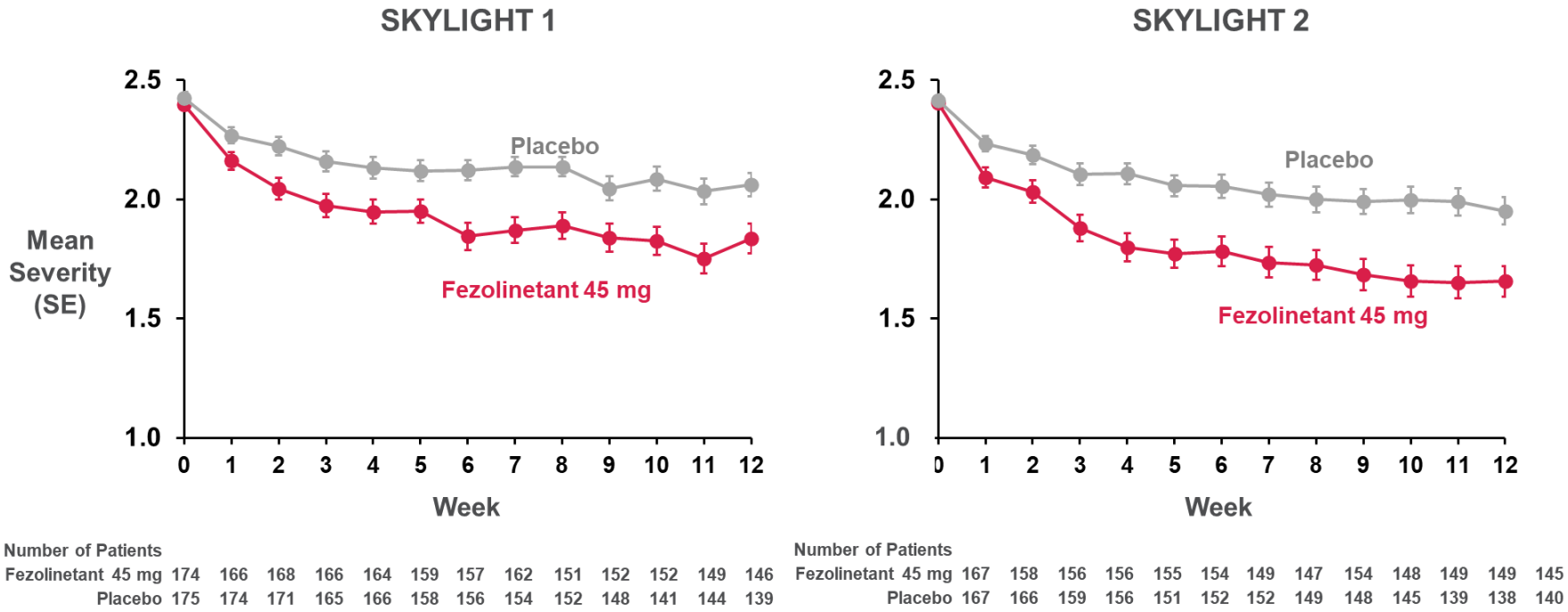
All patients who were randomised and received at least one dose of study treatment (Full Analysis Set). The LS means, SE, CI and P values come from a MMRM analysis of covariance model with change from baseline as the dependent variable and treatment group, week and smoking status (current versus former/never) as factors, with baseline measurement and baseline weight as covariates, as well as an interaction of treatment by week and an interaction of baseline measurement by week. [a]A negative change indicated a reduction/improvement from baseline (that is, a favourable outcome). [b]Differences were calculated by subtracting the LS mean of placebo group from the LS mean of fezolinetant group. [c]P value is for comparison of fezolinetant with placebo from the above described MMRM model.
**Statistically significant adjusting for the multiplicity using the Hochberg procedure.
Abbreviations: CI: confidence interval; LS: least squares; MMRM: mixed model for repeated measures; NA: not applicable; PROMIS SD SF 8b: Patient−Reported Outcome Measurement Information System Sleep Disturbance – Short Form 8b.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs and ISE.[83-85]

Figure 13: Mean frequency of moderate to severe VMS per 24 hours (Full Analysis Set); 12-week double-blind period



All patients who were randomised and received at least one dose of study treatment (Full Analysis Set). Summary statistics are based on non-missing observed data.
Abbreviations: SE: standard error; VMS: vasomotor symptoms.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs.[83, 84]

Figure 14: Mean severity of moderate to severe VMS per 24 hours (Full Analysis Set); 12-week double-blind period



All patients who were randomised and received at least one dose of study treatment (Full Analysis Set). Summary statistics are based on non-missing observed data.
Abbreviations: SE: standard error; VMS: vasomotor symptoms.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs.[83, 84]

Table 29: Percent change from baseline in mean frequency of moderate to severe VMS per 24 hours at selected visits (Full Analysis Set); 12-week double-blind period

Analysis Visit	Statistic	SKYLIGHT 1		SKYLIGHT 2		Pooled studies (SKYLIGHT 1 AND SKYLIGHT 2)	
		Placebo (n=175)	Fezolinetant 45 mg (n=174)	Placebo (n=167)	Fezolinetant 45 mg (n=167)	Placebo (n=342)	Fezolinetant 45 mg (n=341)
Baseline	n	175	174	167	167	342	341
	Mean (SD)	10.51 (3.79)	10.44 (3.92)	11.59 (5.02)	11.79 (8.26)	11.04 (4.46)	11.10 (6.45)
Week 4	n	166	164	151	155	317	319
	Mean (SD)	7.25 (4.29)	5.20 (4.48)	8.08 (6.50)	5.67 (7.29)	7.64 (5.46)	5.43 (6.00)
	Percentage change from baseline[a]						
	n	166	164	151	155	317	319
	Mean (SD)	−30.46 (35.30)	−50.63 (35.44)	−33.60 (34.14)	−55.16 (36.25)	−31.96 (34.73)	−52.84 (35.85)
	LS mean (SE)	−30.59 (2.67)	−51.65 (2.69)	−34.72 (2.78)	−56.37 (2.77)	−32.59 (1.93)	−53.90 (1.94)
	95% CI (two-sided)	−35.83, −25.35	−56.94, −46.35	−40.18, −29.26	−61.82, −50.92	−36.37, −28.80	−57.70, −50.10
	Difference in LS Means: fezolinetant versus placebo[b]						
	LS mean (SE)		−21.05 (3.79)		−21.65 (3.92)		−21.31 (2.73)
	95% CI (two-sided)		−28.50, −13.61		−29.36, −13.94		−26.67, −15.95
	P value (two-sided unadjusted)[c]		<0.001		<0.001		<0.001
Week 12	n	139	146	140	145	279	291
	Mean (SD)	6.85 (4.66)	4.06 (3.85)	6.73 (7.58)	4.49 (5.39)	6.79 (6.28)	4.27 (4.68)
	Percentage change from baseline[a]						
	n	139	146	140	145	279	291
	Mean (SD)	−34.97 (39.65)	−61.35 (32.71)	−45.35 (39.79)	−64.27 (34.92)	−40.18 (39.99)	−62.80 (33.80)
	LS mean (SE)	−37.06 (2.89)	−61.24 (2.86)	−46.91 (2.87)	−65.85 (2.85)	−41.93 (2.04)	−63.41 (2.02)
	95% CI (two-sided)	−42.73, −31.38	−66.85, −55.62	−52.56, −41.26	−71.45, −60.25	−45.93, −37.93	−67.37, −59.44

	Difference in LS Means: fezolinetant versus placebo[b]						
	LS mean (SE)		−24.18 (4.06)		−18.94 (4.05)		−21.48 (2.87)
	95% CI (two-sided)		−32.16, −16.20		−26.89, −10.98		−27.11, −15.85
	P value (two-sided unadjusted)[c]		<0.001		<0.001		<0.001

All patients who were randomised and received at least one dose of study treatment (Full Analysis Set). The LS means, SE, CI, and p values come from a MMRM analysis of covariance model with percent change from baseline as the dependent variable and treatment group, week and smoking status (current versus former/never) as factors, with baseline measurement and baseline weight as covariates, as well as an interaction of treatment by week and an interaction of baseline measurement by week. [a]A negative percent change indicated a reduction/improvement from baseline (that is, a favourable outcome). [b]Differences were calculated by subtracting the LS mean of placebo group from the LS mean of fezolinetant group. [c]P value is for comparison of fezolinetant with placebo from the above described MMRM model.
Abbreviations: CI: confidence interval; LS: least squares; MMRM: mixed model for repeated measures; NA: not applicable.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs and ISE.[83-85]

B.2.6.2.3.3 Frequency and severity of moderate to severe VMS, change from baseline to Week 24

Patients treated with fezolinetant 45 mg for the entire study had greater reductions in the frequency and severity of moderate to severe VMS compared with placebo during the 12-week double-blind period. This improvement in VMS frequency was sustained through Week 24 (Table 30 and Table 31); the timepoint used for all efficacy analyses in DAYLIGHT. In patients who received placebo for the 12-week double-blind period, there was an improvement in VMS frequency from baseline to Week 24 of fezolinetant exposure (study Week 36) once these patients were re-randomised from placebo to fezolinetant 45 mg for the active treatment extension period. This improvement included the placebo effect during the 12-week double-blind period and the effect from the re-randomisation from placebo to fezolinetant 45 mg. Overall, this improvement in VMS frequency was sustained through Week 24 of fezolinetant exposure, with no evidence of reduced effect size suggestive of tachyphylaxis.

Table 30: Change from baseline to Week 24 (24 weeks of fezolinetant exposure) in mean frequency of moderate to severe VMS per 24 hours (Full Analysis Set, fezolinetant); 52-week period

Analysis Visit	Statistic	SKYLIGHT 1		SKYLIGHT 2	
		Placebo/ Fezolinetant 45 mg (n=76)	Fezolinetant 45 mg (n=174)	Placebo/ Fezolinetant 45 mg (n=75)	Fezolinetant 45 mg (n=167)
Baseline[a]	n	76	174	75	167
	Mean (SD)	10.57 (4.12)	10.44 (3.92)	10.91 (2.95)	11.79 (8.26)
24 weeks of fezolinetant exposure[b]	n	62	138	60	134
	Mean (SD)	3.68 (4.48)	3.32 (3.70)	3.89 (6.07)	3.12 (4.50)
	Change from Baseline[c]				
	n	62	138	60	134
	Mean (SD)	−7.32 (4.53)	−7.32 (4.58)	−7.08 (5.40)	−7.96 (4.53)

All patients who were randomised and received at least one dose of fezolinetant during either the 12-week double-blind or the active treatment extension periods (Full Analysis Set, Fezolinetant). Results in this table include only data from 24 weeks of fezolinetant exposure. For the placebo/fezolinetant groups, data from the 12-week placebo-controlled period (placebo data) were not included. These patients were re-randomised in a 1:1 ratio to 30 mg or 45 mg of fezolinetant after 12 weeks of placebo and only data from the active treatment extension period (while they were receiving fezolinetant) were analysed for these patients. For the fezolinetant groups, data from the 12-week period were included as these patients were on fezolinetant for the entire duration of the study. [a]Baseline for all treatment groups is the value taken at the start of the study, on or prior to first dose of placebo or fezolinetant. [b]Week 24 from study start for patients in the fezolinetant 45 mg group; Week 36 from study start for patients in the placebo/fezolinetant group. [c]A negative change indicated a reduction/improvement from baseline (that is, a favourable outcome).
Abbreviations: SD: standard deviation; VMS: vasomotor symptoms.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs.[83, 84]

Table 31: Change from baseline to Week 24 (24 weeks of fezolinetant exposure) in mean severity of moderate to severe VMS per 24 hours (Full Analysis Set, fezolinetant); 52-week period

Analysis Visit	Statistic	SKYLIGHT 1		SKYLIGHT 2	
		Placebo/ Fezolinetant 45 mg (n=76)	Fezolinetant 45 mg (n=174)	Placebo/ Fezolinetant 45 mg (n=75)	Fezolinetant 45 mg (n=167)
Baseline[a]	n	76	174	75	167
	Mean (SD)	2.42 (0.36)	2.40 (0.35)	2.44 (0.31)	2.41 (0.34)
24 weeks of fezolinetant exposure[b]	n	62	138	60	134
	Mean (SD)	1.65 (0.84)	1.64 (0.84)	1.49 (0.86)	1.52 (0.87)
	Change from baseline[c]				
	n	62	138	60	134
	Mean (SD)	−0.76 (0.89)	−0.77 (0.90)	−0.95 (0.88)	−0.90 (0.80)

Baseline included moderate to severe incidences. Post-baseline includes mild, moderate and severe incidences. All patients who were randomised and received at least one dose of fezolinetant during either the 12-week double-blind or the active treatment extension periods (Full Analysis Set, Fezolinetant). Results in this table include only data from 24 weeks of fezolinetant exposure. For the placebo/fezolinetant groups, data from the 12-week placebo-controlled period (placebo data) were not included. These patients were re-randomised in a 1:1 ratio to 30 mg or 45 mg of fezolinetant after 12 weeks of placebo and only data from the active treatment extension period (while they were receiving fezolinetant) were analysed for these patients. For the fezolinetant groups, data from the 12-week period were included as these patients were on fezolinetant for the entire duration of the study. [a]Baseline for all treatment groups is the value taken at the start of the study, on or prior to first dose of placebo or fezolinetant. [b]Week 24 from study start for patients in the fezolinetant 45 mg group; Week 36 from study start for patients in the placebo/fezolinetant group. [c]A negative change indicated a reduction/improvement from baseline (that is, a favourable outcome).
Abbreviations: SD: standard deviation; VMS: vasomotor symptoms.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs.[83, 84]

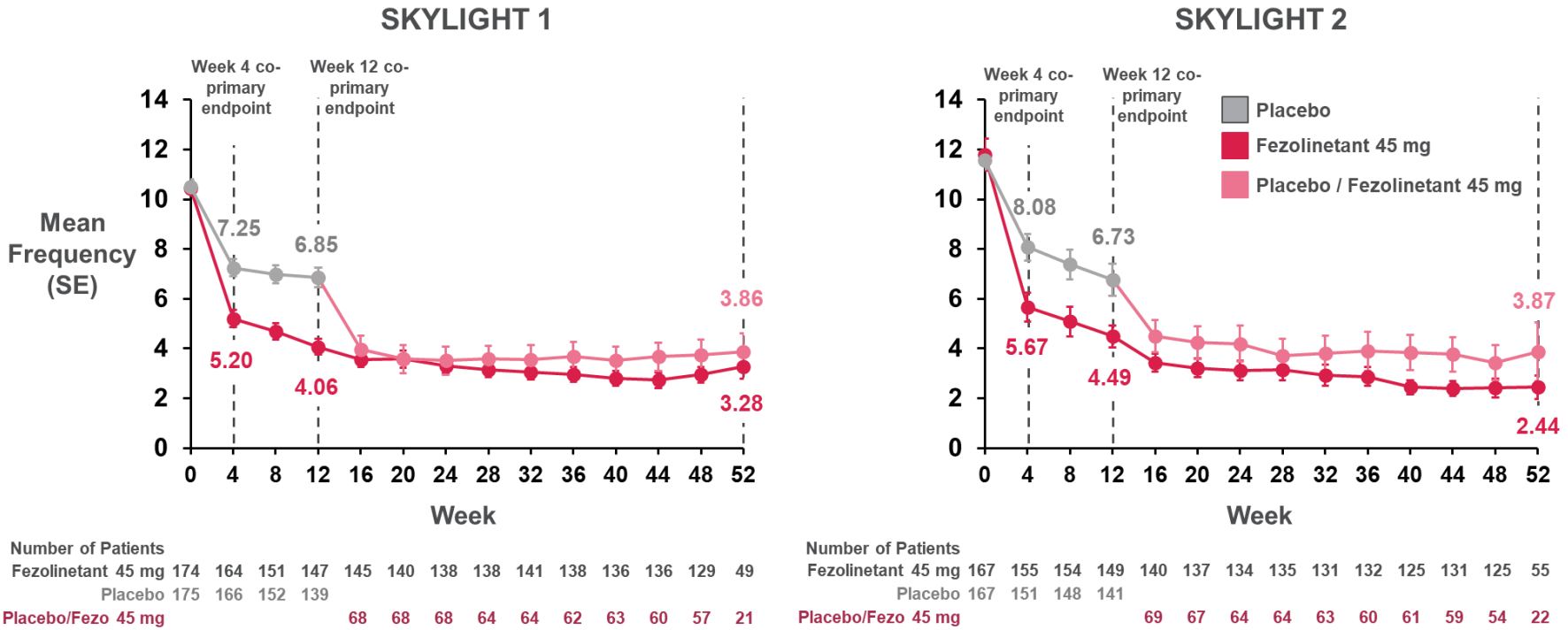
B.2.6.2.4 Exploratory endpoints

B.2.6.2.4.1 Frequency and severity of moderate to severe VMS, change from baseline to each visit in the active treatment extension period and the follow-up visit

Long-term efficacy and persistence of effect of fezolinetant 30 mg and 45 mg were assessed in the active treatment extension period of the pivotal phase 3 studies.

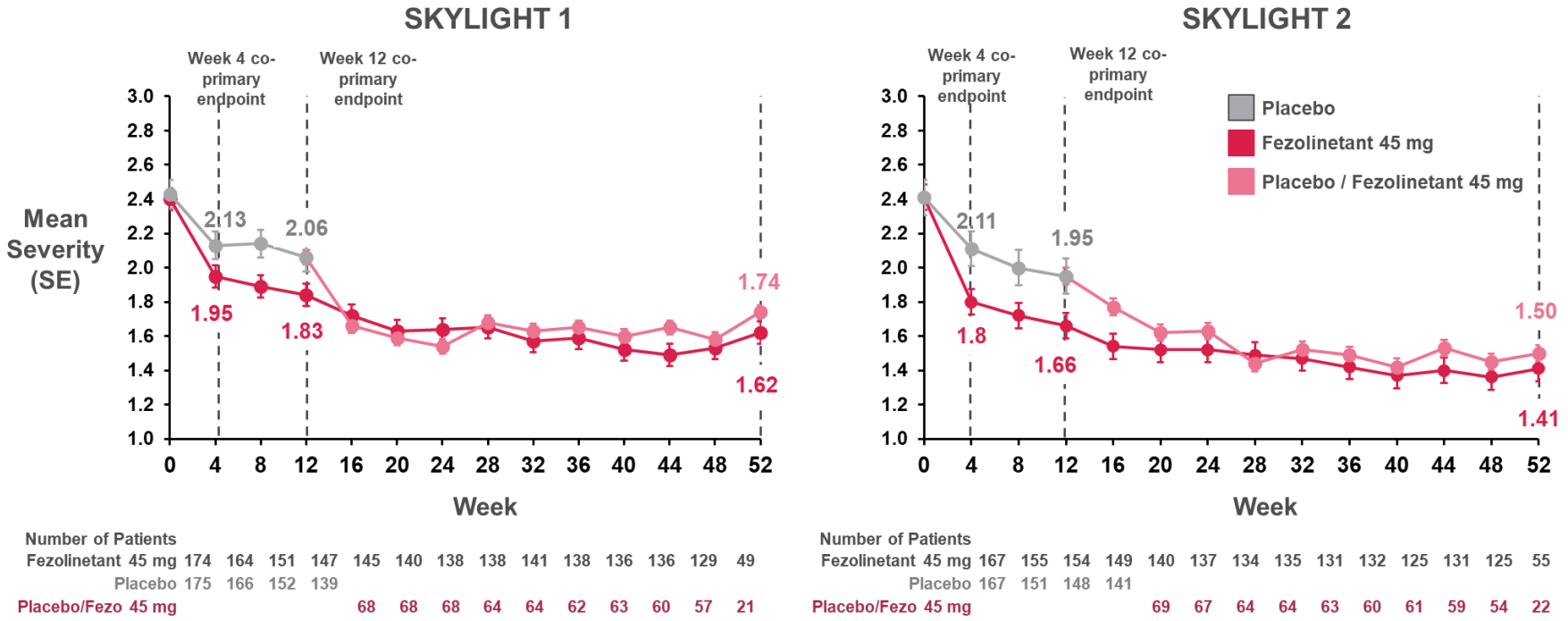
Results up to Week 52 of SKYLIGHT 1 and SKYLIGHT 2 demonstrate that the statistically significant reduction in the LS mean daily frequency and severity of moderate to severe VMS achieved during the 12-week placebo-controlled period was maintained throughout the 52-week study, with no evidence of reduced effect size over time suggestive of tachyphylaxis (Figure 15 and Figure 16). Patients on placebo when re-randomised to active fezolinetant demonstrated additional and sustained benefit from fezolinetant treatment on VMS frequency and severity from baseline throughout the 52-week study. Patients who continued fezolinetant 45 mg since the start of the trial (Week 0) also demonstrated a reduction in VMS frequency and severity over time up to 52 weeks. There was also no evidence of reduced effect size over time suggestive of tachyphylaxis through the study.

Figure 15: Mean frequency of moderate to severe VMS per 24 hours by visit (Full Analysis Set, Fezolinetant); 52-week Period



All patients who were randomised and received at least one dose of fezolinetant during either the 12-week double-blind or the active treatment extension periods (Full Analysis Set, Fezolinetant). Baseline for all treatment groups was the value taken at the start of the study, on or prior to first dose of placebo or fezolinetant. At Week 12, patients in the placebo group were re-randomised in a 1:1 ratio to 30 mg or 45 mg of fezolinetant for the active treatment extension period. For the fezolinetant groups, patients were on fezolinetant for the entire duration of the study. Weeks shown indicate weeks relative to initial randomisation.
Abbreviations: SE: standard error; VMS: vasomotor symptoms.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs.[83, 84]

Figure 16: Mean severity of moderate to severe VMS per 24 hours by visit (Full Analysis Set, Fezolinetant); 52-week Period



All patients who were randomised and received at least one dose of fezolinetant during either the 12-week double-blind or the active treatment extension periods (Full Analysis Set, Fezolinetant). Baseline for all treatment groups was the value taken at the start of the study, on or prior to first dose of placebo or fezolinetant. At Week 12, patients in the placebo group were re-randomised in a 1:1 ratio to 30 mg or 45 mg of fezolinetant for the active treatment extension period. For the fezolinetant groups, patients were on fezolinetant for the entire duration of the study. Weeks shown indicate weeks relative to initial randomisation.
Abbreviations: SE: standard error.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs.[83, 84]

B.2.6.2.4.2 Change from baseline to Week 12 in PROMIS SRI SF 8a total score

In SKYLIGHT 1 and SKYLIGHT 2, there was a numerical reduction (improvement) from baseline in PROMIS SRI SF 8a relative to placebo at Week 4 and Week 12 in patients treated with fezolinetant 45 mg (Table 32).

Table 32: Change from baseline to Week 12 in patient-reported sleep related impairment by the PROMIS SRI SF 8b total – Final Analysis Set

Analysis Visit	Statistic	SKYLIGHT 1		SKYLIGHT 2	
		Placebo (n=175)	Fezolinetant 45 mg (n=174)	Placebo (n=167)	Fezolinetant 45 mg (n=167)
Baseline	n	175	174	166	167
	Mean (SD)	20.8 (7.3)	22.0 (7.3)	22.4 (7.5)	21.5 (7.2)
Week 4	n	160	159	151	158
	Mean (SD)	18.5 (7.3)	16.8 (6.5)	18.9 (6.9)	17.3 (6.3)
	Change from baseline[a]				
	n	160	159	151	158
	Mean (SD)	−2.1 (6.6)	−4.9 (7.3)	−3.2 (7.1)	−4.2 (7.6)
	LS mean change from baseline[a]				
	LS mean (SE)	−2.2 (0.5)	−4.5 (0.5)	−3.1 (0.5)	−4.5 (0.5)
	95% CI (two−sided)	−3.1, −1.3	−5.4, −3.6	−4.1, −2.1	−5.4, −3.5
	Difference in LS Means: fezolinetant versus placebo[b]				
	LS mean (SE)		−2.3 (0.6)		−1.3 (0.7)
	95% CI (two−sided)		−3.5, −1.0		−2.7, 0.0)
	P value (two−sided unadjusted)[c]		<0.001		0.054
Week 12	n	148	156	144	145
	Mean (SD)	18.0 (7.8)	16.9 (6.9)	18.5 (6.7)	17.1 (5.9)
	Change from baseline[a]				
	n	148	156	143	145
	Mean (SD)	−2.6 (7.7)	−4.7 (6.9)	−3.8 (7.6)	−4.2 (7.1)
	LS mean change from baseline[a]				
	LS mean (SE)	−2.5 (0.5)	−4.2 (0.5)	−3.7 (0.5)	−4.7 (0.5)
	95% CI (two−sided)	−3.4, −1.5	−5.2, −3.2	−4.7, −2.7	−5.7, −3.7
	Difference in LS Means: fezolinetant versus placebo[b]				
	LS mean (SE)		−1.8 (0.7)		−1.0 (0.7)
	95% CI (two−sided)		−3.1, −0.4		−2.4, 0.4

	P value (two−sided unadjusted)[c]		0.013		0.164

All patients who were randomised and received at least one dose of study treatment (Full Analysis Set). The LS means, SE, CI and P values come from a MMRM analysis of covariance model with change from baseline as the dependent variable and treatment group, week and smoking status (current versus former/never) as factors, with baseline measurement and baseline weight as covariates, as well as an interaction of treatment by week and an interaction of baseline measurement by week. [a]A negative change indicated a reduction/improvement from baseline (that is, a favourable outcome). [b]Differences were calculated by subtracting the LS mean of placebo group from the LS mean of fezolinetant group. [c]P value is for comparison of fezolinetant with placebo from the above described MMRM model.
Abbreviations: CI: confidence interval; LS: least squares; MMRM: mixed model for repeated measures; NA: not applicable; PROMIS SRI SF 8a: Patient−Reported Outcome Measurement Information System Sleep Related Impairment – Short Form 8a.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs.[83, 84]

B.2.6.2.4.3 Change from baseline to Week 12 in MENQOL (vasomotor domain)

In patients treated with fezolinetant 45 mg, there was a reduction (improvement) from baseline in MENQOL vasomotor domain score relative to placebo at Week 4 of the SKYLIGHT 1 and SKYLIGHT 2 trials (Table 33), both p<0.001 (unadjusted). The greatest reductions were in the MENQOL vasomotor domain total score: least squares mean reduction over placebo (SE) for fezolinetant 45 mg at Week 4 was −1.01 (0.15; 95% CI −1.30 to −0.73) and at Week 12 was −0.86 (0.16; 95% CI −1.17 to −0.56).[91] The treatment effect on the MENQOL vasomotor domain score in the fezolinetant 45 mg groups was sustained throughout the active treatment extension period. Patients who received placebo for the 12−week double−blind period demonstrated an improvement from baseline in MENQOL vasomotor score, once these patients were re−randomised to fezolinetant 45 mg for the active treatment extension period.

Table 33: Change from baseline in MENQOL vasomotor domain scores (Full Analysis Set); 12−Week double blind period

Analysis Visit	Statistic	SKYLIGHT 1		SKYLIGHT 2	
		Placebo (n=175)	Fezolinetant 45 mg (n=174)	Placebo (n=167)	Fezolinetant 45 mg (n=167)
Baseline	n	175	174	165	167
	Mean (SD)	6.46 (1.50)	6.35 (1.71)	6.61 (1.38)	6.60 (1.40)
Week 4	n	160	158	151	158
	Mean (SD)	5.42 (1.86)	4.36 (1.80)	5.35 (1.90)	4.30 (2.04)
	Change from baseline[a]				
	n	160	158	150	158
	Mean (SD)	−1.07 (1.83)	−2.00 (2.12)	−1.26 (1.87)	−2.30 (2.08)
	LS mean change from baseline[a]				
	LS mean (SE)	−1.03 (0.14)	−2.03 (0.14)	−1.28 (0.15)	−2.33 (0.15)
	95% CI (two−sided)	−1.31, −0.76	−2.30, −1.75	−1.58, −0.98	−2.63, −2.04
	Difference in LS Means: fezolinetant versus placebo[b]				

	LS mean (SE)		−1.00 (0.20)		−1.05 (0.21)
	95% CI (two−sided)		(−1.38, −0.61)		−1.48, −0.93
	P value (two−sided unadjusted)[c]		<0.001		<0.001
Week 12	n	148	156	144	145
	Mean (SD)	5.08 (2.13)	4.05 (1.98)	4.76 (2.02)	3.96 (1.96)
	Change from baseline[a]				
	n	148	156	142	145
	Mean (SD)	−1.36 (1.91)	−2.30 (2.19)	−1.86 (2.05)	−2.61 (2.00)
	LS mean change from baseline[a]				
	LS mean (SE)	−1.31 (0.16)	−2.31 (0.15)	−1.89 (0.16)	−2.66 (0.16)
	95% CI (two−sided)	−1.62, −1.01	−2.61, −2.01	−2.21, −1.58	−2.96, −2.35
	Difference in LS Means: fezolinetant versus placebo[b]				
	LS mean (SE)		−0.99 (0.22)		−0.76 (0.22)
	95% CI (two−sided)		−1.42, −0.57		−1.20, −0.32
	P value (two−sided unadjusted)[c]		<0.001		<0.001

The LS Means, standard errors (SE), confidence intervals (CI), and p values come from an MMRM analysis of covariance model with change from baseline as the dependent variable and treatment group, week and smoking status (current versus former/never) as factors, with baseline measurement and baseline weight as covariates, as well as an interaction of treatment by week and an interaction of baseline measurement by week. [a]A negative change indicates a reduction/improvement from baseline (that is, a favourable outcome). [b]Differences are calculated by subtracting the LS mean of placebo group from the LS mean of fezolinetant group. [c]P value was for comparison of fezolinetant with placebo from the above described MMRM model.
Abbreviations: CI: confidence interval; LS: least squares; MENQOL: menopause−specific quality−of−life; MMRM: mixed model for repeated measures.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs. [83, 84]

B.2.6.2.4.4 Change from baseline to Week 12 in EQ-5D-5L VAS

In patients treated with fezolinetant 45 mg, there was an increase from baseline in EQ-5D-5L VAS score relative to placebo at Week 4 of the SKYLIGHT 1 and SKYLIGHT 2 trials (Table 34), unadjusted p=0.025 and p=0.015, respectively. The treatment effect on the EQ-5D-5L VAS score in the fezolinetant 45 mg groups was generally sustained throughout the active treatment extension period. Patients who received placebo for the 12-week double-blind period had an increase from baseline in EQ-5D-5L VAS score once these patients were re-randomised from placebo to fezolinetant 45 mg for the active treatment extension period (see SKYLIGHT 1 and SKYLIGHT 2 clinical study reports for further details).

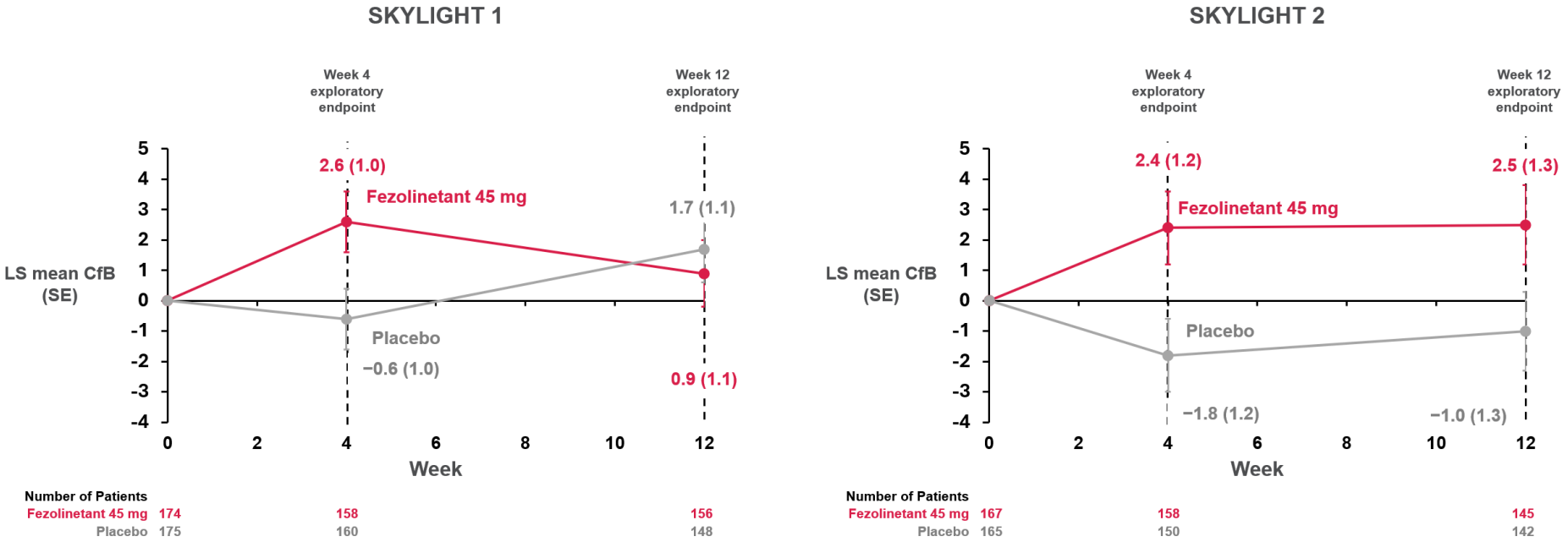
Table 34: Change from baseline in EQ-5D-5L VAS Score (Full Analysis Set); 12-Week double blind period

Analysis Visit	Statistic	SKYLIGHT 1		SKYLIGHT 2	
		Placebo (n=175)	Fezolinetant 45 mg	Placebo (n=167)	Fezolinetant 45 mg

			(n=174)		(n=167)
Baseline	n	175	174	165	167
	Mean (SD)	80.5 (14.4)	79.9 (15.3)	76.4 (19.2)	78.0 (15.9)
Week 4	n	160	158	151	158
	Mean (SD)	80.3 (17.1)	83.0 (15.9)	75.0 (19.6)	79.9 (15.0)
	Change from baseline[a]				
	n	160	158	150	158
	Mean (SD)	−0.5 (14.1)	2.8 (14.8)	−1.7 (19.9)	1.9 (14.8)
	LS mean change from baseline[a]				
	LS mean (SE)	−0.6 (1.0)	2.6 (1.0)	−1.8 (1.2)	2.4 (1.2)
	95% CI (two-sided)	−2.6, 1.4	0.6, 4.6	−4.2, 0.6	0.1, 4.8
	Difference in LS Means: fezolinetant versus placebo[b]				
	LS mean (SE)		3.2 (1.4)		4.2 (1.7)
	95% CI (two-sided)		0.4, 6.0		0.8, 7.7
	P value (two-sided unadjusted)[c]		0.025		0.015
Week 12	n	148	156	144	145
	Mean (SD)	82.9 (15.4)	81.3 (17.3)	75.2 (19.6)	80.1 (15.7)
	Change from baseline[a]				
	n	148	156	142	145
	Mean (SD)	1.7 (13.8)	1.1 (18.4)	−0.3 (22.7)	2.2 (16.0)
	LS mean change from baseline[a]				
	LS mean (SE)	1.7 (1.1)	0.9 (1.1)	−1.0 (1.3)	2.5 (1.3)
	95% CI (two-sided)	−0.5, 3.8	−1.3, 3.0	−3.5, 1.6	−0.1, 5.0
	Difference in LS Means: fezolinetant versus placebo[b]				
	LS mean (SE)		−0.8 (1.5)		3.4 (1.8)
	95% CI (two-sided)		−3.8, 2.2		−0.2, 7.0
	P value (two-sided unadjusted)[c]		0.598		0.061

The LS Means, standard errors (SE), confidence intervals (CI), and p values come from an MMRM analysis of covariance model with change from baseline as the dependent variable and treatment group, week and smoking status (current versus former/never) as factors, with baseline measurement and baseline weight as covariates, as well as an interaction of treatment by week and an interaction of baseline measurement by week. [a]A positive change indicates an improvement from baseline (that is, a favourable outcome). [b]Differences are calculated by subtracting the LS mean of placebo group from the LS mean of fezolinetant group. [c]P value was for comparison of fezolinetant with placebo from the above described MMRM model.
Abbreviations: CI: confidence interval; EQ-5D-5L: EuroQoL five dimensions five level version; LS: least squares; MMRM: mixed model for repeated measures; NA: not applicable; VAS: visual analogue scale.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs.[83, 84]

Figure 17: Change from baseline to Week 4 and Week 12 in EQ-5D-5L VAS Score – Full Analysis Set



Abbreviations: CfB: change from baseline; EQ-5D-5L: EuroQoL five dimensions five level version; SE: standard error; VAS: visual analogue scale.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSR.[83, 84]

B.2.6.2.4.5 Change from baseline to Week 12 in WPAI-VMS

In patients treated with fezolinetant 45 mg, there was an improvement from baseline in WPAI-VMS domain scores relative to placebo at Week 4 and Week 12, including in the overall work productivity loss (Table 35) and activity impairment (Table 36) domains. The effect of fezolinetant treatment on both domains was sustained throughout the active treatment extension period (see SKYLIGHT 1 and SKYLIGHT 2 clinical study reports for further details).

Table 35: Change from baseline in WPAI-VMS domain (overall work productivity loss) scores (Full Analysis Set); 12-Week double blind period

Analysis Visit	Statistic	SKYLIGHT 1		SKYLIGHT 2	
		Placebo (n=175)	Fezolinetant 45 mg (n=174)	Placebo (n=167)	Fezolinetant 45 mg (n=167)
Baseline	n	95	97	111	107
	Mean (SD)	49.18 (26.18)	43.78 (29.80)	49.05 (26.85)	46.47 (28.33)
Week 4	n	82	94	95	91
	Mean (SD)	32.34 (25.65)	20.12 (23.05)	34.62 (29.00)	22.00 (23.81)
	Change from baseline[a]				
	n	71	79	86	83
	Mean (SD)	−14.65 (27.27)	−21.08 (31.10)	−14.71 (35.69)	−25.06 (31.64)
	LS mean change from baseline[a]				
	LS mean (SE)	−10.65 (2.82)	−21.34 (2.66)	−16.22 (2.77)	−26.21 (2.82)
	95% CI (two-sided)	−16.21, −5.10	−26.57, −16.10	−21.68, −10.76	−31.76, −20.66
	Difference in LS Means: fezolinetant versus placebo[b]				
	LS mean (SE)		−10.68 (3.88)		−9.99 (3.95)
	95% CI (two-sided)		−18.32, −3.04		−17.77, −2.21
	P value (two-sided unadjusted)[c]		0.006		0.012
Week 12	n	71	82	72	84
	Mean (SD)	33.33 (27.21)	16.52 (21.77)	31.49 (27.41)	19.06 (24.83)
	Change from baseline[a]				
	n	62	71	66	72
	Mean (SD)	−13.92 (29.73)	−26.11 (28.55)	−14.92 (32.24)	−27.88 (33.46)
	LS mean change from baseline[a]				
	LS mean (SE)	−9.27 (2.89)	−27.34 (2.70)	−15.90 (3.01)	−29.82 (2.89)
	95% CI (two-sided)	−14.98, −3.56	−32.68, −22.01	−21.83, −9.98	−35.52, −24.13
	Difference in LS Means: fezolinetant versus placebo[b]				
	LS mean (SE)		−18.08 (3.97)		−13.92 (4.16)

	95% CI (two-sided)		−25.90, −10.26		−22.13, −5.71
	P value (two-sided unadjusted)[c]		<0.001		<0.001

The LS Means, standard errors (SE), confidence intervals (CI), and p values come from an MMRM analysis of covariance model with change from baseline as the dependent variable and treatment group, week and smoking status (current versus former/never) as factors, with baseline measurement and baseline weight as covariates, as well as an interaction of treatment by week and an interaction of baseline measurement by week. [a]A positive change indicates an improvement from baseline (that is, a favourable outcome). [b]Differences are calculated by subtracting the LS mean of placebo group from the LS mean of fezolinetant group. [c]P value was for comparison of fezolinetant with placebo from the above described MMRM model.
Abbreviations: CI: confidence interval; LS: least squares; MMRM: mixed model for repeated measures; NA: not applicable; VMS: vasomotor symptoms; WPAI-VMS: Work Productivity and Activity Impairment questionnaire specific to VMS.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs.[83, 84]

Table 36: Change from baseline in WPAI-VMS domain (activity impairment) scores (Full Analysis Set); 12-Week double blind period

Analysis Visit	Statistic	SKYLIGHT 1		SKYLIGHT 2	
		Placebo (n=175)	Fezolinetant 45 mg (n=174)	Placebo (n=167)	Fezolinetant 45 mg (n=167)
Baseline	n	175	174	165	167
	Mean (SD)	47.94 (27.88)	43.79 (29.73)	50.79 (25.99)	49.10 (28.11)
Week 4	n	160	158	150	158
	Mean (SD)	33.94 (28.06)	22.72 (23.26)	33.67 (26.56)	27.22 (26.27)
	Change from baseline[a]				
	n	160	158	149	158
	Mean (SD)	−13.75 (29.60)	−20.51 (28.88)	−16.04 (30.62)	−21.84 (28.10)
	LS mean change from baseline[a]				
	LS mean (SE)	−11.47 (1.78)	−20.99 (1.78)	−16.24 (2.00)	−22.75 (1.95)
	95% CI (two-sided)	−14.96, −7.98	−24.48, −17.50	−20.17, −12.31	−26.58, −18.92
	Difference in LS Means: fezolinetant versus placebo[b]				
	LS mean (SE)		−9.52 (2.52)		−6.51 (2.79)
	95% CI (two-sided)		(−14.46, −4.57)		−11.99, −1.03
	P value (two-sided unadjusted)[c]		<0.001		0.020
Week 12	n	148	156	144	145
	Mean (SD)	29.66 (28.51)	20.32 (23.78)	31.67 (26.50)	23.10 (25.40)
	Change from baseline[a]				
	n	148	156	142	145
	Mean (SD)	−17.23 (30.70)	−23.21 (30.94)	−19.72 (31.93)	−26.07 (26.47)
	LS mean change from baseline[a]				

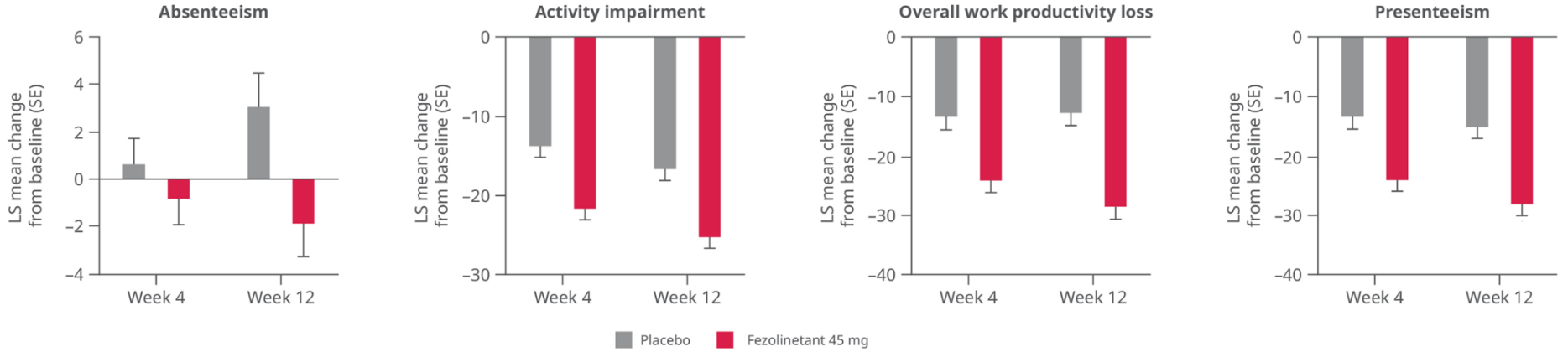
	LS mean (SE)	−14.80 (1.95)	−23.76 (1.90)	−18.78 (2.03)	−27.01 (2.00)
	95% CI (two-sided)	−18.63, −10.97	−27.51, −20.02	−22.76, −14.80	−30.95, −23.07
	Difference in LS Means: fezolinetant versus placebo[b]				
	LS mean (SE)		−8.97 (2.73)		−8.23 (2.85)
	95% CI (two-sided)		−14.32, −3.61		−13.83, −2.63
	P value (two-sided unadjusted)[c]		0.001		0.004

The LS Means, standard errors (SE), confidence intervals (CI), and p values come from an MMRM analysis of covariance model with change from baseline as the dependent variable and treatment group, week and smoking status (current versus former/never) as factors, with baseline measurement and baseline weight as covariates, as well as an interaction of treatment by week and an interaction of baseline measurement by week. [a]A positive change indicates an improvement from baseline (that is, a favourable outcome). [b]Differences are calculated by subtracting the LS mean of placebo group from the LS mean of fezolinetant group. [c]P value was for comparison of fezolinetant with placebo from the above described MMRM model.

Abbreviations: CI: confidence interval; LS: least squares; MMRM: mixed model for repeated measures; NA: not applicable; VMS: vasomotor symptoms; WPAI-VMS: Work Productivity and Activity Impairment questionnaire specific to VMS

Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs.[83, 84]

Figure 18: Change from baseline to Week 4 and Week 12 in WPAI-VMS domain scores – Final Analysis Set



Abbreviations: WPAI-VMS: Work Productivity and Activity Impairment questionnaire specific to VMS.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs.[83, 84]

B.2.7 Subgroup analysis

B.2.7.1 DAYLIGHT

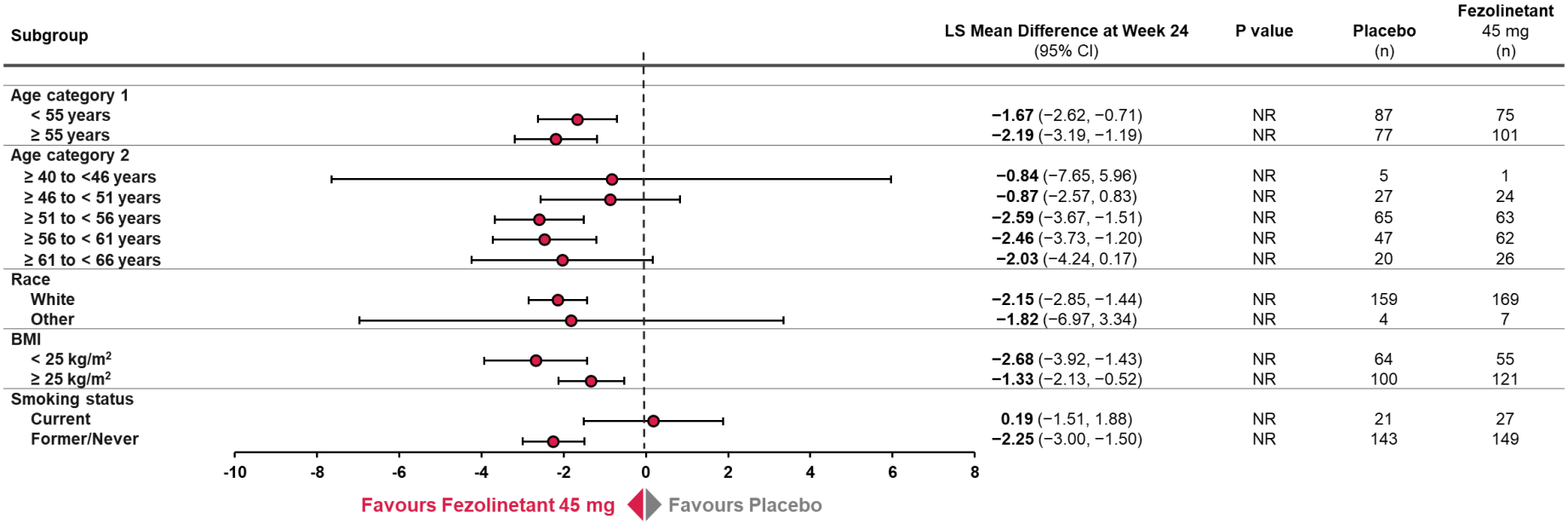
B.2.7.1.1 Consistency of treatment effect on the co-primary endpoints

Results from pre-specified subgroup analyses of the primary efficacy endpoint at Week 24 in DAYLIGHT are presented in Figure 19.

DAYLIGHT was not designed to detect treatment differences in the subgroups; therefore, the results should be interpreted with caution. Moreover, interpretation of the results is also limited due to the small number of patients in some subgroups.

Improvement in mean frequency of moderate to severe VMS was broadly consistent across clinically relevant and pre-specified subgroups in the fezolinetant 45 mg group. As compared with the full trial population (see section B.2.6.1.1), subgroup results were broadly aligned. The only exception to the general observation that fezolinetant was associated with a greater improvement in mean frequency of moderate to severe VMS than placebo was the subgroup analysis of current smokers.

Figure 19: Forest plot of subgroup analysis of difference in mean change in frequency of moderate to severe VMS from baseline to Week 24 compared with placebo (Full Analysis Set); DAYLIGHT



A negative change indicated a reduction/improvement from baseline (that is, a favourable outcome). Categories with fewer than 15 patients are not presented.
Abbreviations: BMI: body mass index; CI: confidence interval; VMS: vasomotor symptoms.
Source: Astellas (Data on File). DAYLIGHT CSR.[82]

B.2.7.1.2 Subgroup analyses excluding the HRT-averse population

As discussed in Section B.1.1, the decision problem addressed in this submission excludes the population of patients who are deemed to be suitable for HRT, but made an informed decision not to take it following risk-benefit discussions of HRT. This is in line with clinical expert opinion (discussed in Section B.1.1 and B.1.3.3) who agreed that fezolinetant should not be positioned as a treatment for patients who are HRT-averse, as it is expected that HRT remains the relevant comparator for this patient population.

However, the design of the DAYLIGHT trial permitted the inclusion of patients who were deemed to be HRT-averse, which includes 168 patients, or 37% of all patients randomised to treatment in DAYLIGHT. As the efficacy results presented in Section B.2.6.1 are inclusive of the HRT-averse subgroup, it is assumed that excluding this large cohort of patients from the efficacy analyses would introduce considerable uncertainty to trial results as the trial was not powered to explore differences within HRT-unsuitable subgroups, and the subsequent economic analysis.

However, to assess the impact of the HRT-averse population on the results from the DAYLIGHT trial, subgroup analyses were conducted which excluded HRT-averse patients. The baseline characteristics and clinical efficacy outcomes for the subgroup analysis excluding HRT-averse patients are presented in Appendix E.

B.2.7.2 SKYLIGHT 1 and SKYLIGHT 2

B.2.7.2.1 Consistency of treatment effect on the co-primary endpoints

The consistency of treatment effect on the co-primary endpoints was evaluated in the pooled Week 12 analysis of SKYLIGHT 1 and SKYLIGHT 2 as supportive efficacy data. The list of prespecified and post-hoc subgroup analyses are presented in Table 37 below. Pre-specified and post-hoc subgroup analyses demonstrated that the intrinsic and extrinsic factors evaluated had no impact on the efficacy of fezolinetant on the co-primary endpoints of VMS frequency and severity at Week 12.

Table 37: Subgroup analyses performed in the pooled analysis of SKYLIGHT 1 and SKYLIGHT 2

Subgroup analysis	Factor(s)
Prespecified subgroup analyses	
Intrinsic factors	• Age category 1 (<55 years vs ≥55 years) • Race: category 1 (White vs Non-White), category 2 (Black vs Non-Black) and category 3 (Asian vs Non-Asian) • Ethnicity: Hispanic or Latino vs Not Hispanic or Latino • BMI: ≥18.5 kg/m^2 to <25.0 kg/m^2 vs ≥25.0 kg/m^2 to <30.0 kg/m^2 vs ≥30.0 kg/m^2
Extrinsic factors	• Smoking status: current vs former/never • Region: Europe vs North America
Additional post-hoc subgroup analyses	
Intrinsic factors	• Age category 2 (<50 years vs 50 to 60 years vs >60 years) • Time since amenorrhea: <2 years vs 2 to 10 years vs >10 years • VMS duration: <1 year vs 1 to 10 years vs >10 years • VMS frequency at baseline: ≥median vs <median (median = 9.78)

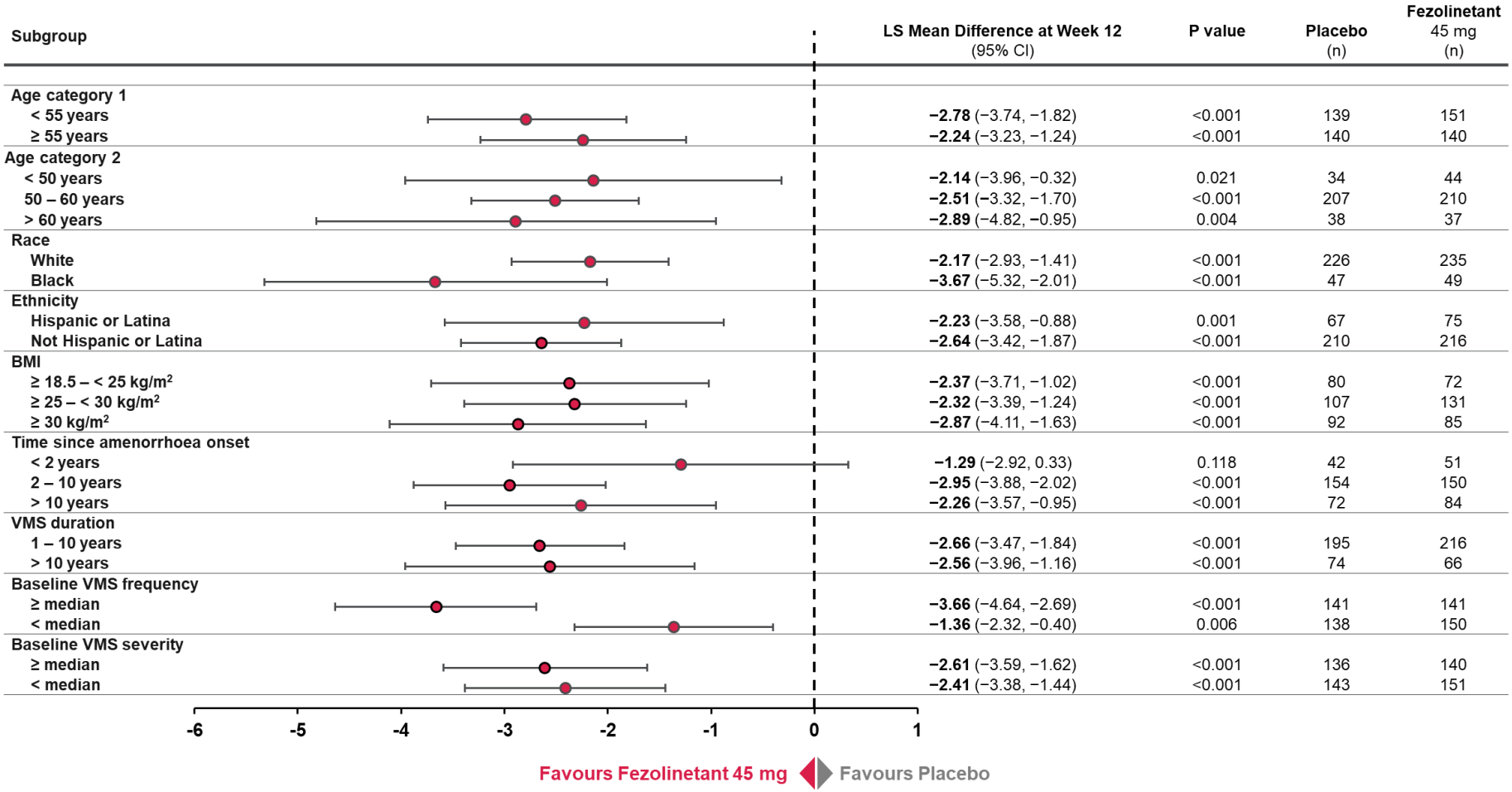
	• VMS severity at baseline: ≥median vs <median (median = 2.36)
Extrinsic factors	• Sleep disturbance (PROMIS SD SF 8b): <1st quartile (= 23) vs 1st quartile (= 23) to <median (= 27) vs median (= 27) to <3rd quartile (= 32) vs ≥3rd quartile • Medical history of hypertension: yes vs no • Oophorectomy history: yes vs no • Hysterectomy history: yes vs no • Hysterectomy/oophorectomy history: hysterectomy without oophorectomy vs hysterectomy with oophorectomy • Previous HRT treatment: yes vs no • Reason for stopping HRT: lack of improvement in symptoms vs other • Concurrent use of SSRIs: yes vs no • Alcohol history: current vs former/never

Abbreviations: BMI: body mass index; HRT: hormone replacement therapy; PROMIS SD SF 8b: Patient-reported Outcomes Measurement Information System Sleep Disturbance – Short Form 8b; SSRI: selective serotonin receptor inhibitors; VMS: vasomotor symptoms.

For VMS frequency, the subgroup analyses showed generally similar results across the intrinsic (Figure 20) and extrinsic factors (Figure 21). The magnitude of the treatment effect was similar for all the categories analysed. The treatment by subgroup interaction for each intrinsic and each extrinsic factor was tested at an alpha level of 0.1. In general, the interaction between treatment by subgroups was not significant. The wide confidence intervals were due to the small patient numbers in each subgroup.

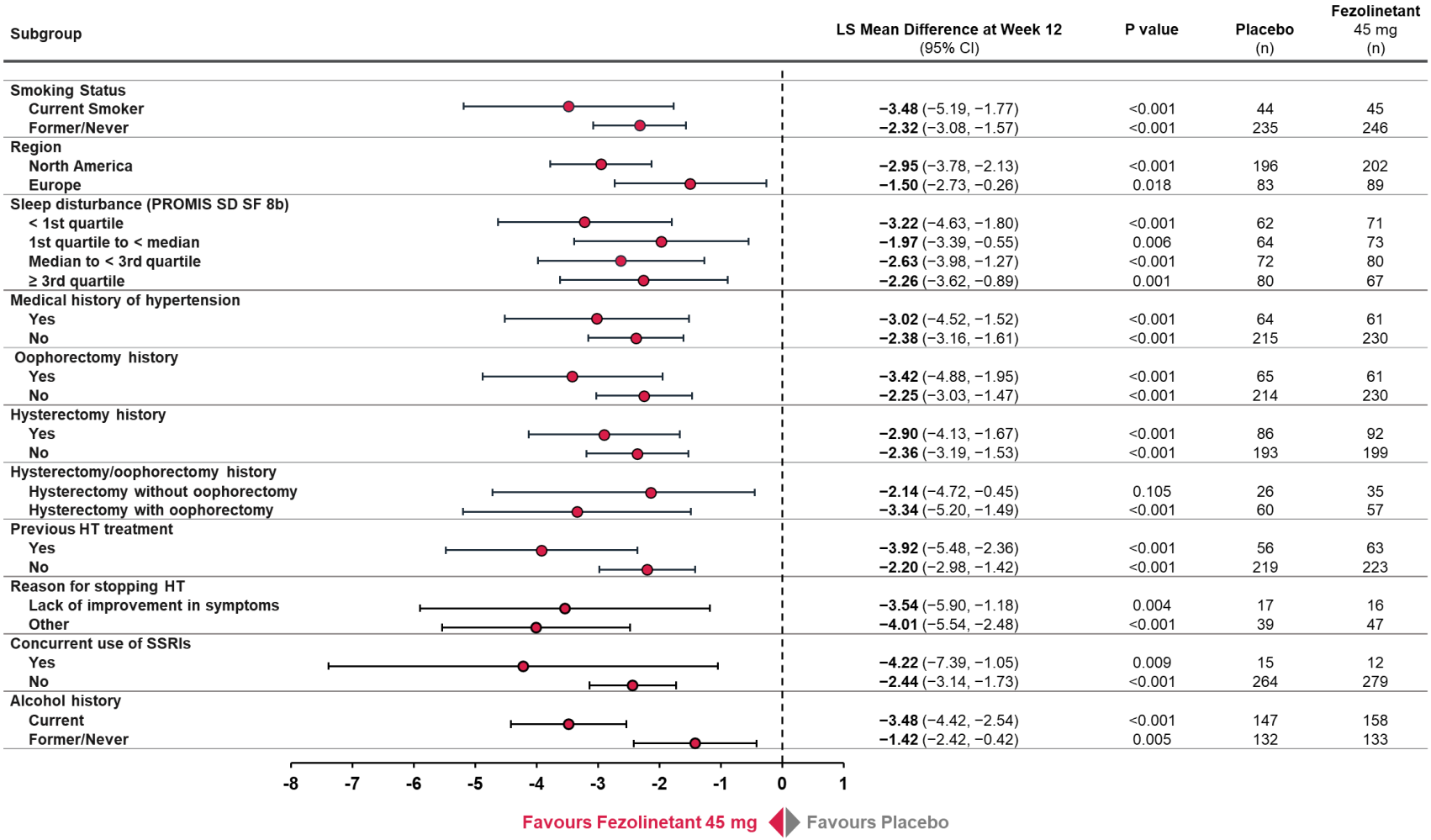
Similarly, there was no clinically meaningful impact of intrinsic (Figure 22) and extrinsic factors (Figure 23) on the efficacy of fezolinetant as measured by VMS severity at Week 12. The magnitude of the treatment effect was generally similar for the categories analysed except for the subgroup categories of age (<50 years and >60 years), ethnicity (Hispanic or Latino), time since amenorrhoea onset (>0 years) and VMS duration (>10 years); however, the 95% CIs included 0 and the number of patients was small in some of these categories. Furthermore, the point estimates still favoured fezolinetant versus placebo.

Figure 20: Intrinsic factors: Forest plot of subgroup analysis of difference in mean change in frequency of moderate to severe VMS from baseline to Week 12 compared with placebo (Full Analysis Set); pooled SKYLIGHT 1 and SKYLIGHT 2



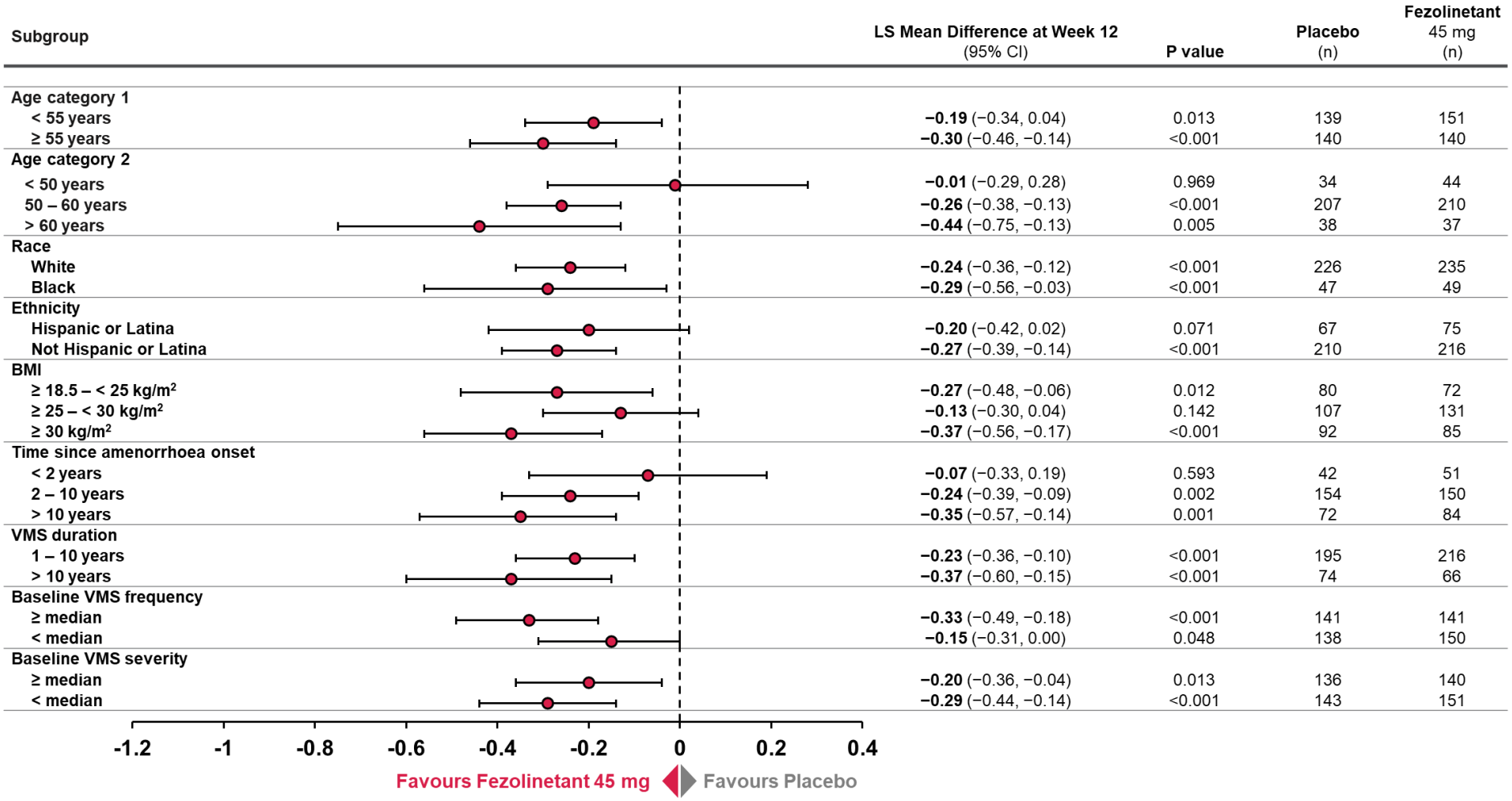
All patients who were randomised and received at least one dose of study drug (Full Analysis Set). A negative change indicated a reduction/improvement from baseline (that is, a favourable outcome). Categories with fewer than 15 patients are not presented.
Abbreviations: BMI: body mass index; CI: confidence interval; VMS: vasomotor symptoms.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs and ISE.[83-85]

Figure 21: Extrinsic factors: Forest plot of subgroup analysis of difference in mean change in frequency of moderate to severe VMS from baseline to Week 12 compared with placebo (Full Analysis Set); pooled SKYLIGHT 1 and SKYLIGHT 2



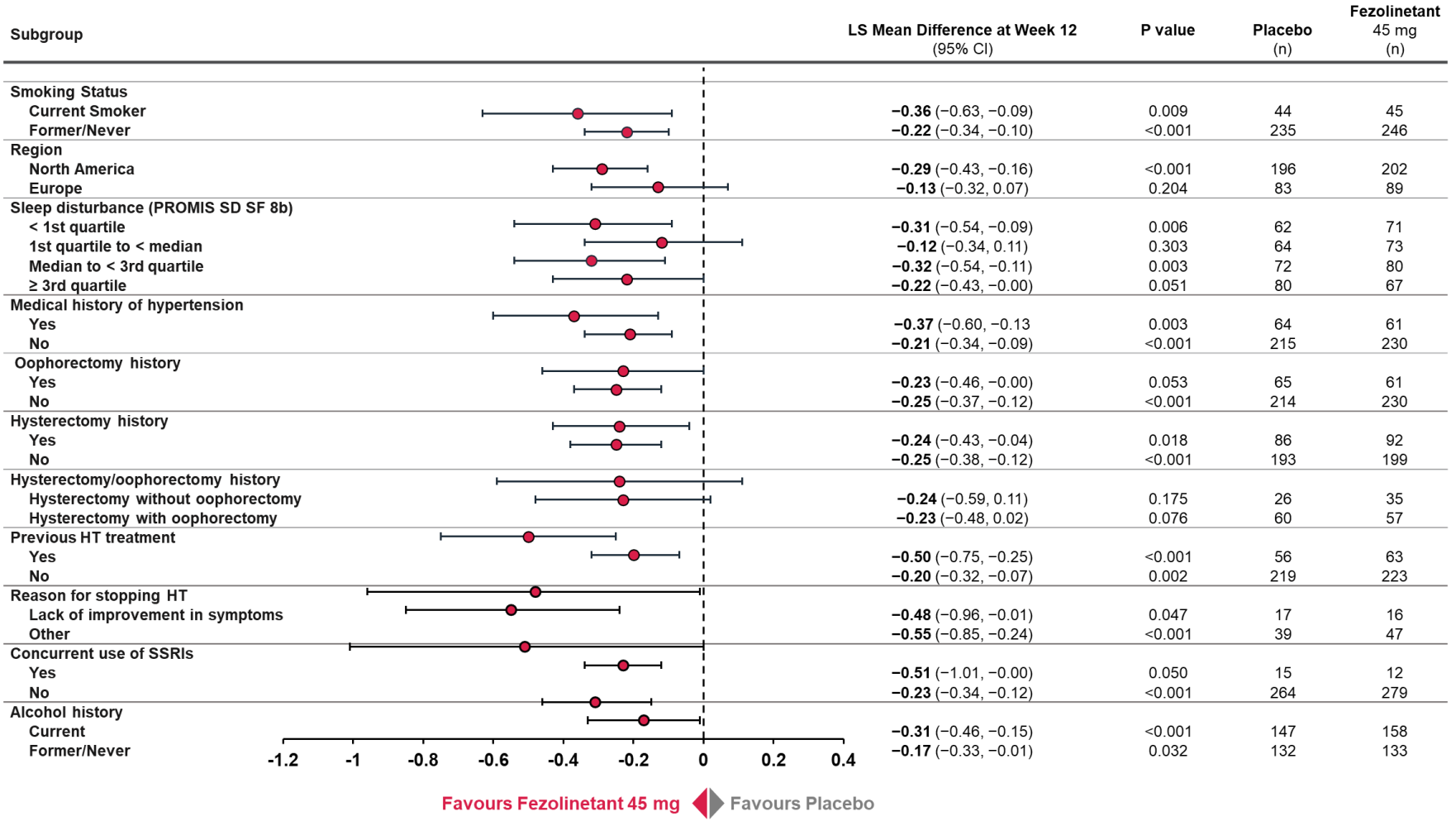
All patients who were randomised and received at least one dose of study drug (Full Analysis Set). A negative change indicated a reduction/improvement from baseline (that is, a favourable outcome). Categories with fewer than 15 patients are not presented.
Abbreviations: HRT: hormone replacement therapy; SSRI: selective serotonin reuptake inhibitor.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs and ISE.[83-85]

Figure 22: Intrinsic factors: Forest plot of subgroup analysis of difference in mean change in severity of moderate to severe VMS from baseline to Week 12 compared with placebo (Full Analysis Set); pooled SKYLIGHT 1 and SKYLIGHT 2



All patients who were randomised and received at least one dose of study drug (Full Analysis Set). A negative change indicated a reduction/improvement from baseline (that is, a favourable outcome). Categories with fewer than 15 patients are not presented.
Abbreviations: BMI: body mass index; CI: confidence interval; LS: least squares; VMS: vasomotor symptoms.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs and ISE.[83-85]

Figure 23: Extrinsic factors: Forest plot of subgroup analysis of difference in mean change in severity of moderate to severe VMS from baseline to Week 12 compared with placebo (Full Analysis Set); pooled SKYLIGHT 1 and SKYLIGHT 2



All patients who were randomised and received at least one dose of study drug (Full Analysis Set). A negative change indicated a reduction/improvement from baseline (that is, a favourable outcome). Categories with fewer than 15 patients are not presented.
Abbreviations: HRT: hormone replacement therapy; SSRI: selective serotonin reuptake inhibitor.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs and ISE.[83-85]

B.2.7.2.2 Subgroup analyses in the pooled HRT-unsuitable subpopulation

For the economic analysis, the DAYLIGHT trial provides direct clinical evidence in the target population for this submission, as women enrolled in the trial had moderate to severe VMS and were deemed HRT-unsuitable. Given that DAYLIGHT provides trial data for up to 24 weeks, it was necessary to source model inputs from the pooled SKYLIGHT 1 and 2 trials with a longer follow-up of 52 weeks. As such, to be consistent with the target population of this submission and to align more closely with the population in the DAYLIGHT trial, pooled SKYLIGHT 1 and 2 trial data were sourced from the subpopulation of women who were deemed HRT-unsuitable, using the Week 24 and Week 52 analyses to inform the economic modelling. This approach was considered appropriate as most women (n=875, 85.6%) were deemed unsuitable for HRT in the pooled trials, providing reassurance that the sample size of the efficacy data is still suitably large to inform the model base case.

The methodology for defining HRT-unsuitability in the SKYLIGHT trials has previously been discussed in Section B.2.3.2.2, and the baseline characteristics of patients in the HRT-unsuitable subgroup of the pooled analysis of SKYLIGHT 1 and SKYLIGHT 2 are presented in Appendix E.

Key efficacy data for the HRT-unsuitable subpopulation from the pooled analysis of SKYLIGHT 1 and SKYLIGHT 2 are presented in Table 38 for Weeks 4 and 12, and Table 39 for Weeks 24 and 52. Additional HRT-unsuitable subgroup data are presented in Appendix E.

Frequency of moderate to severe VMS, change from baseline to Weeks 4 and 12

Table 38: Analysis of change in mean frequency of moderate to severe VMS from baseline to Week 4 and Week 12 in the subgroup of patients unsuitable for HRT (FAS)

	Fezolinetant 45 mg (N=341)			
	N	n	LS mean difference vs placebo (95% CI) †	2-sided un-adjusted p value ‡
Week 4				
HRT-unsuitable	[redacted]	[redacted]	[redacted]	[redacted]
Week 12				
HRT-unsuitable	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CI: confidence interval; FAS: final analysis set; HRT: hormone replacement therapy; LS: least square; MMRM: mixed model repeated measurements.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 Pooled Analysis.[92]

Frequency of moderate to severe VMS, change from baseline to Weeks 24 and 52 *(used in model base-case)*

Data on the change from baseline to Week 24 and Week 52, in the frequency of moderate to severe VMS, were used to generate the clinical effectiveness inputs in the base-case economic analysis discussed in Section B.3. The trial period in the pooled analysis of SKYLIGHT 1 and SKYLIGHT 2, between Week 24 and Week 52, was not placebo controlled, and therefore comparative efficacy results between fezolinetant and placebo are not available for this period. However, in the absence of longer term data of relevance from DAYLIGHT, Week 24 and Week 52 data from the HRT-unsuitable subgroup of the pooled analysis of SKYLIGHT 1 and SKYLIGHT 2 were considered a suitable alternative to generate longer term efficacy inputs in the model (see further details in Section B.3.3.2.1).

The Week 24 and Week 52 results observed in the pooled SKYLIGHT 1 and 2 HRT-unsuitable subpopulation were consistent with those of the overall pooled SKYLIGHT 1 and 2 population. Specifically, the direction of the treatment effect and point estimate for the subpopulation aligned with the corresponding values in the overall population, demonstrating similar efficacy across both populations.

Table 39: Analysis of change in mean frequency of moderate to severe VMS from baseline to Week 24 and Week 52 in the subgroup of patients unsuitable for HRT (FAS)

Subgroup	Fezolinetant 45 mg (N=341)	
	n	Mean (SD)
Week 24		
HRT-unsuitable	[redacted]	[redacted]
Week 52		
HRT-unsuitable	[redacted]	[redacted]

Abbreviations: FAS: final analysis set; HRT: hormone replacement therapy, VMS: vasomotor symptoms.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 Pooled Analysis Table 9.3.2.3.1.[93]

B.2.7.2.3 Subgroup analyses excluding the HRT-averse population

For consistency with the DAYLIGHT data used in the base case economic analysis, the efficacy results for the HRT-unsuitable subgroup analysis of the pooled SKYLIGHT 1 and SKYLIGHT 2 trials presented in Section B.2.7.2.2 also included patients who were classed as HRT-averse. Similar subgroup analyses to those presented in Section B.2.7.1.2 excluding the HRT-averse population were also conducted for the pooled population of SKYLIGHT 1 and SKYLIGHT 2 to inform the HRT-averse-excluded scenario in the economic analyses.

The baseline characteristics and clinical efficacy outcomes for this subgroup analysis excluding HRT-averse patients are presented in Appendix E.

B.2.8 Meta-analysis

In the base-case economic analysis, clinical inputs were primarily sourced from the DAYLIGHT trial, making a meta-analysis unnecessary. Additional clinical data were sourced from the HRT-unsuitable subpopulation of the pooled SKYLIGHT 1 and SKYLIGHT 2 trials. As SKYLIGHT 1 and SKYLIGHT 2 were identically designed (discussed in B.2.2 and Section B.2.7.1.2) naïve pooling of data from both trials was deemed appropriate and used in the model base case.

Further pooled analyses using SKYLIGHT 1, SKYLIGHT 2 and DAYLIGHT were also conducted to assess the impact of using pooled trial data (i.e., using a larger sample size) on the results in the economic analysis. Given the similarities in the trial design – particularly in terms of population, intervention, comparator and outcomes – naïve pooling of data from DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 was considered appropriate. Baseline characteristics and data from the analyses are presented in Appendix E.3.

The results from these analyses informed model scenario analyses and are discussed in Section B.3.11.3.

While a formal meta-analysis of SKYLIGHT 1 and SKYLIGHT 2 could be conducted for this appraisal, it would not provide additional advantages beyond the naïve pooling already performed as both trials involved the same population, interventions, comparators, outcomes, and time points.

It should be noted that the model submitted with this appraisal utilises individual patient-level data from all three trials and does not include functionality to incorporate alternative comparative efficacy estimates such as odds ratios from meta-analyses.

B.2.9 Indirect and mixed treatment comparisons

As discussed in Section B.1.3.3, fezolinetant is positioned for use in people with moderate to severe VMS for whom HRT is deemed unsuitable for medical reasons. UK clinical experts have confirmed that no active treatment is the sole relevant comparator to fezolinetant in this setting. Therefore, it was not necessary to conduct an indirect treatment comparison, as clinical evidence on the efficacy of no active treatment was sourced from the placebo arm of the DAYLIGHT trial.

B.2.10 Adverse reactions

Summary of safety evidence for fezolinetant

- In the DAYLIGHT trial, patients receiving fezolinetant 45mg showed comparable frequencies of TEAEs comparable to those receiving placebo. Both groups had a low incidence of serious TEAEs, with only one serious TEAE attributed to the study drug in the fezolinetant 45 mg group. No deaths occurred in the study. Additionally, no significant differences were observed in the occurrence of specific TEAEs of special interest, including liver test elevations, uterine bleeding, endometrial hyperplasia, cancer or disordered proliferative endometrium, and thrombocytopenia. Notably, the incidence of liver enzyme elevations was low across both groups (ALT or AST levels >3 x ULN), and no patients had bilirubin levels >2 ULN. Furthermore, there were no instances of Hy's law observed in the trial.
- In both SKYLIGHT 1 and SKYLIGHT 2 trials, during the 12-week double-blind period, the incidence of TEAEs for patients on fezolinetant 45 mg was comparable to the placebo group. No deaths were reported, and serious TEAEs were infrequent and comparable between the fezolinetant 45 mg and placebo groups. None of the serious TEAEs were deemed related to the study treatment by the investigators. TEAEs leading to withdrawal of treatment were low and similar between the treatment groups in both trials. No notable differences were observed in the frequency of any TEAEs of special interest across the treatment groups in both trials. This included factors such as liver enzyme elevations, uterine bleeding, endometrial hyperplasia, cancer or disordered proliferative endometrium and thrombocytopenia. Furthermore, the incidence of liver enzyme elevations (ALT or AST levels >3 x ULN) was low, with these events generally being asymptomatic, transient, and resolving either during treatment or after treatment discontinuation. None of the patients in either trial exhibited bilirubin levels >2 ULN, and there were no cases of Hy's law observed.
- Overall, the safety profile of fezolinetant 45 mg demonstrates a low incidence of adverse events, with no concerning patterns or severe adverse effects detected.

Trial results relating to AEs in this submission are presented for the SAF populations in the DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 trials. The key safety results from these SAF populations are presented below.

Additional safety data were also available from the SKYLIGHT 4 trial, further details of which are presented in Appendix F. However, as a summary, after 52 weeks of treatment, treatment emergent adverse events (TEAEs) were comparable between all treatment groups in SKYLIGHT 4 and only a small proportion of these TEAEs were considered to be serious (3.8% and 2.3% of

patients in the fezolinetant 45 mg and placebo groups, respectively). Likewise, the proportion of patients who withdrew from the study due to a TEAE was similar across the treatment groups and no statistically significant differences in incidence of treatment-emergent adverse events of special interest (AESIs) were observed between treatment groups.

Full details of safety data for DAYLIGHT, SKYLIGHT 1 and 2 and SKYLIGHT 4 are available in the respective CSRs.

B.2.10.1 DAYLIGHT

B.2.10.1.1 Exposure of study treatment

A summary of exposure for the 24-week double-blind period of DAYLIGHT is presented in Table 40.

In the SAF population, the median duration of exposure to study treatment (placebo or fezolinetant 45 mg) was 168 days in each treatment group during this period. Most patients (fezolinetant 45 mg: 68.1%; placebo: 65.5%) had cumulative study treatment exposure of ≥168 days.

Table 40: Study treatment exposure – 24-week double blind period (Safety Analysis Set)

Category	Category/ Statistic	Placebo (n=226)	Fezolinetant 45 mg (n=226)
Duration (days)[a]	n	226	226
	Mean (SD)	144.3 (51.8)	157.4 (34.8)
	Median	168.0	168.0
	Min–Max	5–191	6–199
Cumulative duration category, n (%)[a]	≥1	226 (100.0)	226 (100.0)
	≥7	225 (99.6)	225 (99.6)
	≥14	220 (97.3)	224 (99.1)
	≥21	212 (93.8)	223 (98.7)
	≥28	211 (93.4)	223 (98.7)
	≥42	206 (91.2)	220 (97.3)
	≥56	196 (86.7)	218 (96.5)
	≥84	189 (83.6)	209 (92.5)
	≥112	181 (80.1)	201 (88.9)
	≥140	179 (79.2)	197 (87.2)
	≥168	148 (65.5)	154 (68.1)

[a]Duration was defined as (last dose date) − (the first dose date) + 1.
Abbreviations: Max: maximum; Min: minimum; SD: standard deviation.
Source: Astellas (Data on File). DAYLIGHT CSR.[82]

B.2.10.1.2 Overview of TEAEs and death

As shown in Table 41, the frequencies of TEAEs in the patients treated with fezolinetant 45mg in DAYLIGHT, were broadly similar to those who received placebo. The proportion of patients receiving placebo who experienced one or more TEAE was 61.1%, as compared with 65.0% of those receiving fezolinetant 45 mg. Only a small minority of TEAEs were considered serious in

nature in both treatment arms, with patients experiencing marginally higher serious TEAEs in the fezolinetant arm (3.5% versus 4.4%, respectively). Only one serious TEAE was attributed to the study drug in the DAYLIGHT trial.

Additionally, there was a greater, albeit marginal, proportion of patients who withdrew from treatment due to a TEAE in the placebo arm compared to the fezolinetant arm (6.2% versus 4.9%). The proportion of patients where treatment withdrawal was attributed to a study drug was equal for both arms of the trial (7.1% each). No deaths occurred across both arms of the DAYLIGHT trial.

Table 41: Overview of TEAEs and Death in the DAYLIGHT trial – Safety Analysis Set

TEAE, n (%)	Placebo (N=226)	Fezolinetant 45mg (N=226)
Patients with ≥1 TEAE	138 (61.1)	147 (65.0)
Drug-related TEAE[a]	25 (11.1)	39 (17.3)
Serious TEAE[b]	8 (3.5)	10 (4.4)
Drug-related Serious TEAE	0	1 (0.4)
TEAE leading to withdrawal of treatment	14 (6.2)	11 (4.9)
Drug-related TEAE leading to withdrawal of treatment	7 (3.1)	7 (3.1)
Death[c]	0	0

TEAEs were defined as an AE observed after starting administration of the study drug and 21 days after the last dose of study drug. [a]Any TEAE with a causal relationship assessed as "yes" by the investigator. If relationship was missing, then the TEAE was also considered to be drug-related. [b]An AE designated as serious by either the investigator or determined by Important Medical Event process was treated as an SAE. [c]All reported deaths after the first study drug administration.
Abbreviations: AE: adverse event; SAE: serious adverse event; TEAE: treatment emergent adverse event.
Source: Astellas (Data on File). DAYLIGHT CSR.[82]

B.2.10.1.3 Common TEAEs

TEAEs occurring in ≥2% of patients who received fezolinetant 45 mg in the SAF population of the DAYLIGHT trial are presented in Table 42. The most common TEAEs reported by patients who received fezolinetant 45 mg included COVID-19 (13.3%) and headache (8.8%). Likewise, the most commonly reported TEAEs in the placebo arm were also COVID-19 (12.8%) and headache (9.3%). While the observed safety profile of fezolinetant in the DAYLIGHT trial was broadly in line with that observed for placebo, a higher proportion of patients experienced fatigue in the fezolinetant arm versus the placebo arm (5.8% versus 0.4%, respectively). Fatigue is a documented potential side effect for fezolinetant, and was noted as such in the informed consent forms shared with patients during screening. However, most of the events of fatigue experienced by patients who received fezolinetant in DAYLIGHT were considered mild in severity; only one (0.4%) event was considered to be moderate in severity and less than half were causally attributed to the study treatment.

Table 42: Most common TEAEs, occurring in ≥2% of patients in the fezolinetant group (MedDRA V25.0) – Safety Analysis Set

Preferred Term, n (%)	Placebo (N=226)	Fezolinetant 45mg (N=226)
Patients with ≥1 TEAE	**138 (61.1)**	**147 (65.0)**

COVID-19	29 (12.8)	30 (13.3)
Headache	21 (9.3)	20 (8.8)
Fatigue	1 (0.4)	13 (5.8)
Nasopharyngitis	11 (4.9)	9 (4.0)
Nausea	4 (1.8)	6 (2.7)
Diarrhoea	3 (1.3)	6 (2.7)
Upper respiratory tract infection	4 (1.8)	6 (2.7)
Influenza	6 (2.7)	6 (2.7)
Dizziness	5 (2.2)	6 (2.7)
Insomnia	1 (0.4)	6 (2.7)
Anxiety	2 (0.9)	6 (2.7)
Urinary tract infection	5 (2.2)	5 (2.2)
Bronchitis	3 (1.3)	5 (2.2)
Weight increased	1 (0.4)	5 (2.2)
Oropharyngeal pain	2 (0.9)	5 (2.2)

AEs were coded using MedDRA.
Abbreviations: AE: adverse event; TEAE: treatment emergent adverse event.
Source: Astellas (Data on File). DAYLIGHT CSR.[82]

B.2.10.1.4 Serious TEAEs

Serious TEAEs occurring in patients in the SAF population of the DAYLIGHT trial are presented in Table 43 below. Across both treatment arms, only a small minority of patients were reported as experiencing a serious TEAE, however this was marginally higher in the fezolinetant-treated group than the placebo-treated group (4.4% versus 3.5%, respectively). No single serious TEAE was reported in >1 patient.

Table 43: Serious TEAEs (MedDRA V25.0) – Safety Analysis Set

Preferred Term, n (%)	Placebo (N=226)	Fezolinetant 45mg (N=226)
Any serious TEAE	**8 (3.5)**	**10 (4.4)**
Pericardial effusion	0	1 (0.4)
Enteritis	0	1 (0.4)
Gastritis	0	1 (0.4)
Pyelocystitis	0	1 (0.4)
Contusion	0	1 (0.4)
Spinal column injury	0	1 (0.4)
Alanine aminotransferase increased	0	1 (0.4)
Hepatic enzyme increased	0	1 (0.4)
Intervertebral disc protrusion	0	1 (0.4)
Intracranial aneurysm	0	1 (0.4)
Sciatica	0	1 (0.4)
Nephrolithiasis	0	1 (0.4)
Acute kidney injury	0	1 (0.4)
Hernia hiatus repair	0	1 (0.4)

Myocardial infarction	1 (0.4)	0
Coronary artery dissection	1 (0.4)	0
General physical health deterioration	1 (0.4)	0
Cellulitis	1 (0.4)	0
Pneumonia	1 (0.4)	0
Pyelonephritis	1 (0.4)	0
Mental impairment	1 (0.4)	0
Psychotic disorder	1 (0.4)	0
Ureterolithiasis	1 (0.4)	0
Respiratory distress	1 (0.4)	0
Haematoma	1 (0.4)	0

AEs were coded using MedDRA.
Abbreviations: AE: adverse event; TEAE: treatment emergent adverse event.
Source: Astellas (Data on File). DAYLIGHT CSR.[82]

B.2.10.1.5 TEAEs of special interest

The TEAEs listed in Table 44 were predetermined as TEAEs of special interest in the DAYLIGHT trial for the SAF population. Overall, uterine bleeding, endometrial hyperplasia, cancer or disordered proliferative endometrium and thrombocytopenia occurred more frequently in patients receiving placebo in DAYLIGHT, as compared with those treated with fezolinetant 45 mg. Patients who received fezolinetant demonstrated elevated liver test levels with greater frequency than those patients who received placebo.

Table 44: TEAEs of special interest (MedDRA V25.0) – Safety Analysis Set

Preferred Term, n (%)	Placebo (N=226)	Fezolinetant 45mg (N=226)
Liver test elevations	6 (2.7)	10 (4.4)
Uterine bleeding	10 (4.4)	6 (2.7)
Endometrial hyperplasia, cancer or disordered proliferative endometrium	2 (0.9)	1 (0.4)
Thrombocytopenia	1 (0.4)	0

AEs were coded using MedDRA.
Abbreviations: AE: adverse event; TEAE: treatment emergent adverse event.
Source: Astellas (Data on File). DAYLIGHT CSR.[82]

B.2.10.1.6 TEAEs leading to withdrawal of treatment

TEAEs which led to treatment withdrawal in the SAF population of the DAYLIGHT trial are presented in Table 45. Overall, a higher proportion of treatment withdrawals due to TEAEs occurred in patients receiving placebo as compared with patients receiving fezolinetant 45 mg treatment (6.2% versus 4.9%, respectively). Only diarrhoea and increased hepatic enzymes (in the fezolinetant 45 mg treatment arm) and headache (in the placebo arm) resulted in >1 patient withdrawing from the study treatment, all n=2 (0.9%).

Table 45: TEAEs (MedDRA V25.0) leading to withdrawal of treatment – Safety Analysis Set

Preferred Term, n (%)	Placebo (N=226)	Fezolinetant 45mg (N=226)

TEAE leading to withdrawal of treatment	14 (6.2)	11 (4.9)
Diarrhoea	0	2 (0.9)
Hepatic enzyme increased	0	2 (0.9)
Abdominal distension	0	1 (0.4)
Constipation	1 (0.4)	1 (0.4)
Alanine aminotransferase increased	0	1 (0.4)
Arthralgia	0	1 (0.4)
Pain in extremity	0	1 (0.4)
Headache	2 (0.9)	1 (0.4)
Paraesthesia	0	1 (0.4)
Vaginal haemorrhage	0	1 (0.4)
Hirsutism	0	1 (0.4)
Varicose vein	0	1 (0.4)
Insomnia	0	1 (0.4)
Myocardial infarction	1 (0.4)	0
Coronary artery dissection	1 (0.4)	0
Abdominal pain	1 (0.4)	0
Swelling face	1 (0.4)	0
Helicobacter infection	1 (0.4)	0
COVID-19	1 (0.4)	0
Fibroadenoma of breast	1 (0.4)	0
Dizziness	1 (0.4)	0
Postmenopausal haemorrhage	1 (0.4)	0
Respiratory distress	1 (0.4)	0
Alopecia	1 (0.4)	0
Depressed mood	1 (0.4)	0
Panic attack	1 (0.4)	0
Hot flush	1 (0.4)	0

AEs were coded using MedDRA.
Abbreviations: AE: adverse event; TEAE: treatment emergent adverse event.
Source: Astellas (Data on File). DAYLIGHT CSR.[82]

B.2.10.1.7 Other safety evaluations

Endometrial safety

In DAYLIGHT, transvaginal ultrasound (TVU) was required for all patients to assess endometrial thickness at screening and changes from baseline in endometrial thickness are described in Table 46. TVU data were generally unremarkable and there were no clinically relevant differences in the change from baseline to Week 24 in endometrial thickness across both treatment groups.

Table 46: Endometrial thickness (mm): change from baseline to Week 24 (Safety Analysis Set, Fezolinetant)

Statistic	Placebo (n=226)	Fezolinetant 45 mg (n=226)
Baseline		
n	204	185
Mean (SD)	2.75 (1.49)	2.70 (1.39)
Median	2.30	2.50
Min–Max	0.0–7.9	0.0–8.0
Change from baseline		
n	129	129
Mean (SD)	−0.16 (1.29)	0.01 (−1.94)
Median	0.00	0.00
Min–Max	−6.5–3.2	−3.4–17.3

Transvaginal ultrasound is not required for patients who have had a partial (supra-cervical) or full hysterectomy.
Abbreviations: Max: maximum; Min: minimum; SD: standard deviation.
Source: Astellas (Data on File). DAYLIGHT CSR.[82]

Liver tests

During the 24-week double-blind period of DAYLIGHT, there was 3 (1.3%) out of 224 patients with ALT or AST levels >3 x ULN in the fezolinetant 45 mg group. None of the patients had bilirubin levels >2 x ULN. Alkaline phosphatase levels were >1.5 x ULN in 1 (0.5%) out of 220 patients in the placebo group and 3 (1.3%) out of 224 patients in the fezolinetant 45 mg group. There were also no cases of Hy's law (definition provided in 'Outcome definitions' Section B.2.3.3).

B.2.10.2 SKYLIGHT 1 and SKYLIGHT 2

B.2.10.2.1 Exposure to study treatment

A summary of exposure for the 12-week double-blind period of SKYLIGHT 1 and SKYLIGHT 2 is presented in Table 47.

In the pooled 12-Week population, the median duration of exposure to study treatment (placebo or fezolinetant) was 84 days in each treatment group during this period. Most patients (fezolinetant 45 mg: 94.4%; placebo: 92.1%) had cumulative study treatment exposure of ≥56 days. Of which, 231 and 246 patients in the fezolinetant 45 mg and placebo groups had a cumulative study treatment exposure of ≥84 days.

Median duration of fezolinetant exposure throughout the 52-week period for patients who continued on fezolinetant during the active extension period was 364 days in both trials (Table 48). Of the patients re-randomised from placebo to fezolinetant 45 mg, median duration of fezolinetant exposure was 280 days on both trials. Most patients in the active treatment extension period had cumulative fezolinetant exposure of ≥280 days. In the fezolinetant groups, >80% of the patients had cumulative exposure of ≥336 days.

Table 47: Study treatment exposure – 12-week double blind period (Safety Analysis Set)

Category	Category/ Statistic	SKYLIGHT 1		SKYLIGHT 2		Pooled studies (SKYLIGHT 1 AND SKYLIGHT 2)	
		Placebo (n=175)	Fezolinetant 45 mg (n=173)	Placebo (n=167)	Fezolinetant 45 mg (n=167)	Placebo (n=342)	Fezolinetant 45 mg (n=341)
12-week Double-blind Period							
Duration (days)[a]	n	175	173	167	167	342	341
	Mean (SD)	78.7 (16.8)	81.4 (14.8)	79.2 (17.5)	80.7 (14.7)	79.0 (17.1)	81.1 (14.7)
	Median	84.0	84.0	84.0	84.0	84.0	84.0
	Min–Max	8–91[b]	1–112[b]	1–96[c]	2–94	1–96	1–112
Cumulative duration category, n(%)[a]	≥ 1 day	175 (100.0)	173 (100.0)	167 (100.0)	167 (100.0)	342 (100.0)	341 (100.0)
	≥ 7 days	175 (100.0)	170 (98.3)	165 (98.8)	165 (98.8)	340 (99.4)	336 (98.5)
	≥ 14 days	174 (99.4)	170 (98.3)	164 (98.2)	165 (98.8)	338 (98.8)	336 (98.5)
	≥ 21 days	169 (96.6)	168 (97.1)	161 (96.4)	164 (98.2)	330 (96.5)	333 (97.7)
	≥ 28 days	167 (95.4)	168 (97.1)	160 (95.8)	163 (97.6)	327 (95.6)	332 (97.4)
	≥ 42 days	163 (93.1)	165 (95.4)	156 (93.4)	159 (95.2)	319 (93.3)	325 (95.3)
	≥ 56 days	161 (92.0)	164 (94.8)	154 (92.2)	157 (94.0)	315 (92.1)	322 (94.4)
	≥ 84 days	115 (65.7)	128 (74.0)	116 (69.5)	117 (70.1)	231 (67.5)	246 (72.1)

All randomised patients who took at least 1 dose of study treatment (Safety Analysis Set). All patients who were randomised and received at least one dose of fezolinetant during either the 12-week double-blind or the active treatment extension periods (Safety Analysis Set, Fezolinetant). Patients in the placebo/fezolinetant groups were patients who were on placebo for 12 weeks and were then re-randomised in a 1:1 ratio to 30 mg or 45 mg of fezolinetant for the active treatment extension period. Results for these groups in this table include only data from fezolinetant exposure. Data from the 12-week placebo-controlled period (placebo data) are not included in the placebo/fezolinetant columns but are shown in the placebo column in the 12-week double-blind period rows. [a]Duration was calculated as (the date of last dosing during the double-blind treatment period) − (the date of first dosing) + 1. [b]Patients took the last dose for the 12-week double-blind period after Day 85. [c]One patient returned to the clinic and the last dose for the 12-week double-blind period was on Day 96.
Abbreviations: Max: maximum; Min: minimum; SD: standard deviation.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs and ISE.[83-85]

Figure 24: Study treatment exposure – 52-week period (Safety Analysis Set)

Category	Category/ Statistic	SKYLIGHT 1		SKYLIGHT 2	
		Placebo/ Fezolinetant 45 mg (n=76)	Fezolinetant 45 mg (n=173)	Placebo/ Fezolinetant 45 mg (n=75)	Fezolinetant 45 mg (n=167)
52-week Period					
Duration (days)[a]	n	76	173	75	167
	Mean (SD)	263.7 (53.4)	326.2 (98.6)	254.1 (67.4)	319.9 (102.1)
	Median	280.0	364.0	280.0	364.0
	Min–Max	62–294	1–402	23–289	2–386
Cumulative duration category, n(%)[a]	≥ 7 days	76 (100.0)	170 (98.3)	75 (100.0)	165 (98.8)
	≥ 14 days	76 (100.0)	170 (98.3)	75 (100.0)	165 (98.8)
	≥ 21 days	76 (100.0)	168 (97.1)	75 (100.0)	164 (98.2)
	≥ 28 days	76 (100.0)	168 (97.1)	73 (97.3)	163 (97.6)
	≥ 42 days	76 (100.0)	165 (95.4)	72 (96.0)	160 (95.8)
	≥ 56 days	76 (100.0)	164 (94.8)	71 (94.7)	158 (94.6)
	≥ 84 days	73 (96.1)	162 (93.6)	70 (93.3)	155 (92.8)
	≥ 112 days	71 (93.4)	156 (90.2)	69 (92.0)	151 (90.4)
	≥ 140 days	70 (92.1)	155 (89.6)	68 (90.7)	149 (89.2)
	≥ 168 days	69 (90.8)	154 (89.0)	66 (88.0)	144 (86.2)
	≥ 196 days	69 (90.8)	153 (88.4)	66 (88.0)	142 (85.0)
	≥ 224 days	69 (90.8)	150 (86.7)	64 (85.3)	141 (84.4)
	≥ 252 days	68 (89.5)	149 (86.1)	63 (84.0)	140 (83.8)
	≥ 280 days	53 (69.7)	147 (85.0)	42 (56.0)	138 (82.6)
	≥ 308 days	0	147 (85.0)	0	137 (82.0)
	≥ 336 days	0	145 (83.8)	0	136 (81.4)
	≥ 364 days	0	120 (69.4)	0	109 (65.3)

[a]Duration was defined as (the date of last dosing of fezolinetant) − (the date of first dosing of fezolinetant) + 1. Duration was only calculated for patients with a complete date of first and last dose of fezolinetant.
Abbreviations: Max: maximum; Min: minimum.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs.[83, 84]

B.2.10.2.2 Overview of TEAEs and death

The safety of fezolinetant 45 mg versus placebo was evaluated in the 12-week double-blind period and are presented below.

A summary of the TEAE in SKYLIGHT 1 and SKYLIGHT 2 is presented in Table 48. Further safety data from SKYLIGHT 4 are presented in Appendix F.

In SKYLIGHT 1 and SKYLIGHT 2, the proportion of patients that experienced at least one TEAE was comparable between the fezolinetant 45 mg group and the placebo group. There were no reported deaths in either trial. The frequency of serious TEAEs was low and the same across the fezolinetant 45 mg arm of both trials and comparable to the placebo arm. No serious TEAEs experienced by patients in the fezolinetant 45 mg group were considered by the investigator to

be related to study treatment (referred to as drug-related in the table below). The frequency of TEAE leading to withdrawal was low and similar across treatment arms and between trials.

Table 48: Overview of treatment-emergent adverse events (Safety Analysis Set); 12-week double-blind period

Category	SKYLIGHT 1		SKYLIGHT 2	
	Placebo (n=175)	Fezolinetant 45 mg (n=173)	Placebo (n=167)	Fezolinetant 45 mg (n=167)
Patients with ≥1 TEAE	78 (44.6%)	75 (43.4%)	54 (32.3%)	60 (35.9%)
Drug-related[a] TEAE	22 (12.6%)	13 (7.5%)	11 (6.6%)	25 (15.0%)
Serious TEAE[b]	1 (0.6%)	2 (1.2%)	0	2 (1.2%)
Drug-related[a] serious TEAE[b]	0	0	0	0
TEAE leading to death	0	0	0	0
Drug-related[a] TEAE leading to death	0	0	0	0
TEAE leading to withdrawal of treatment	9 (5.1%)	4 (2.3%)	1 (0.6%)	5 (3.0%)
Drug-related[a] TEAE leading to withdrawal of treatment[c]	7 (4.0%)	3 (1.7%)	0	5 (3.0%)
Death[d]	0	0	0	0

All randomised patients who took at least one dose of study treatment (Safety Analysis Set). A TEAE was defined as an adverse event observed after starting administration of study treatment and either up to 21 days after the last dose of study treatment, or first dose of study treatment during the active treatment extension period, whichever occurred first. [a]A reasonable possibility that the event may have been caused by the study treatment as assessed by the investigator. If relationship was missing, then it was considered as drug-related. [b]Includes serious adverse events upgraded by the sponsor based on review of the sponsor's list of Always Serious terms, if any upgrade was done. [c]Outcome for the TEAE of alexithymia was updated from "drug withdrawal" to "dose not changed" in the 52-week database since treatment discontinuation occurred later than TEAE end date. This patient was removed from the 52-week data listing on TEAEs leading to withdrawal of treatment. [d]All reported deaths after the first study treatment administration.
Abbreviations: TEAE: treatment-emergent adverse event.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs.[83, 84]

B.2.10.2.3 Common TEAEs

Overall, in SKYLIGHT 1, the three most frequently reported TEAEs by preferred term were headache, blood glucose increased and upper abdominal pain. In SKYLIGHT 2, headache, upper respiratory tract infection, dry mouth and nausea were the three most frequently reported TEAEs by preferred term.

Table 49: Most frequently reported (fezolinetant total) treatment-emergent adverse events in SKYLIGHT 1 by preferred term (Safety Analysis Set); 12-week double-blind period

MedDRA (v23.0) Preferred Term	Placebo (n=175)	Fezolinetant 45 mg (n=173)
Headache	13 (7.4%)	11 (6.4%)
Blood glucose increased	0	6 (3.5%)
Abdominal pain upper	2 (1.1%)	4 (2.3%)
Arthralgia	1 (0.6%)	2 (1.2%)
Blood creatine phosphokinase increased	0	4 (2.3%)

Dizziness	2 (1.1%)	3 (1.7%)
Upper respiratory tract infection	3 (1.7%)	2 (1.2%)
Nasopharyngitis	2 (1.1%)	1 (0.6%)
Alanine aminotransferase increased	4 (2.3%)	3 (1.7%)
Fatigue	3 (1.7%)	3 (1.7%)
Sinusitis	2 (1.1%)	2 (1.2%)
Dry mouth	1 (0.6%)	2 (1.2%)

AEs were coded using MedDRA. All randomised patients who took at least one dose of study treatment (Safety Analysis Set). Treatment-emergent adverse events are reported at the discretion of the investigator. Sorting order: descending by the number of patients of total fezolinetant group by preferred term. In case of ties, ascending order by preferred term code was applied.
Source: Astellas (Data on File). SKYLIGHT 1 CSR.[83]

Table 50: Most frequently reported (fezolinetant total) treatment-emergent adverse events in SKYLIGHT 2 by preferred term (Safety Analysis Set); 12-week double-blind period

MedDRA (v23.0) Preferred Term	Placebo (n=167)	Fezolinetant 45 mg (n=167)
Headache	4 (2.4%)	6 (3.6%)
Upper respiratory tract infection	7 (4.2%)	5 (3.0%)
Dry mouth	0	4 (2.4%)
Nausea	0	4 (2.4%)
Arthralgia	1 (0.6%)	1 (0.6%)
Weight increased	1 (0.6%)	1 (0.6%)
Abdominal distension	1 (0.6%)	2 (1.2%)
Constipation	0	1 (0.6%)
Dry skin	2 (1.2%)	2 (1.2%)
Hypertension	3 (1.8%)	1 (0.6%)
Fatigue	2 (1.2%)	1 (0.6%)
Somnolence	1 (0.6%)	1 (0.6%)

AEs were coded using MedDRA. All randomised patients who took at least one dose of study treatment (Safety Analysis Set). Treatment-emergent adverse events are reported at the discretion of the investigator. Sorting order: descending by the number of patients of total fezolinetant group by preferred term. In case of ties, ascending order by preferred term code was applied.
Source: Astellas (Data on File). SKYLIGHT 2 CSR.[84]

B.2.10.2.4 Study drug-related TEAEs

No drug-related TEAEs were reported for ≥2% of patients in any treatment group during the 12-week double-blind period of SKYLIGHT 1; the individual study drug-related TEAEs were similar between treatment groups.

In SKYLIGHT 2, study drug-related TEAEs occurring in ≥2% of patients in any treatment group were headache (fezolinetant 45 mg: 1 [0.6%]; placebo: 2 [1.2%]) and weight increased (fezolinetant 45 mg: 1 [0.6%]; placebo: 0). Individual study drug-related TEAEs were similar between treatment groups.

Full list of study drug-related TEAEs are provided in the individual CSRs for SKYLIGHT 1 and SKYLIGHT 2. Additional details for drug-related TEAEs from SKYLIGHT 4 are presented in Appendix F and in the SKYLIGHT 4 CSR.

B.2.10.2.5 TEAEs by severity

During the 12-week double-blind period of SKYLIGHT 1 and SKYLIGHT 2, most TEAEs were of mild or moderate severity per severity grading scale as described in the protocol. The following TEAEs were classified as severe for SKYLIGHT 1: abdominal pain upper, nausea, non-cardiac chest pain, cholelithiasis for 4 patients (2.3%) in the placebo group; and pain, headache, dizziness, sciatica and breast pain for 4 patients (2.3%) in the fezolinetant 45 mg group. For SKYLIGHT 2, severe TEAEs were vomiting, colitis and abdominal pain for 1 patient (0.6%) in the fezolinetant 45 mg group. None of the TEAEs observed in the placebo group of SKYLIGHT 2 were classed as severe.

B.2.10.2.6 Serious TEAEs

A summary of serious TEAEs during the 12-week double-blind period of SKYLIGHT 1 and SKYLIGHT 2 is presented in Table 51 and Table 52, respectively. Individual serious TEAEs were reported for 2 patients in the fezolinetant 45 mg group in each trial. None of the serious TEAEs reported were considered related to study treatment by the investigator in either trial.

Table 51: Serious treatment-emergent adverse events (Safety Analysis Set): 12-week double-blind period of SKYLIGHT 1

MedDRA (v23.0) System Organ Class Preferred Term	Placebo (n=175)	Fezolinetant 45 mg (n=173)
Overall serious TEAEs	1 (0.6%)	2 (1.2%)
Hepatobiliary disorders	1 (0.6%)	0
Cholelithiasis	1 (0.6%)	0
Nervous system disorders	0	1 (0.6%)
Paraesthesia	0	1 (0.6%)
Vascular disorders	0	1 (0.6%)
Varicose vein	0	1 (0.6%)

AEs were coded using MedDRA. All randomised patients who took at least one dose of study treatment (Safety Analysis Set).
Abbreviations: TEAE: treatment-emergent adverse event.
Source: Astellas (Data on File). SKYLIGHT 1 CSR.[83]

Table 52: Serious treatment-emergent adverse events (Safety Analysis Set): 12-week double-blind period of SKYLIGHT 2

MedDRA (v23.0) System Organ Class Preferred Term	Placebo (n=167)	Fezolinetant 45 mg (n=167)
Overall serious TEAEs	0	2 (1.2%)
Hepatobiliary disorders	0	1 (0.6%)
Biliary dyskinesia	0	1 (0.6%)
Injury, poisoning and procedural complications	0	1 (0.6%)

Posterior tibial nerve injury	0	1 (0.6%)

AEs were coded using MedDRA. All randomised patients who took at least one dose of study treatment (Safety Analysis Set).
Abbreviations: COVID-19: coronavirus disease 2019; TEAE: treatment-emergent adverse event.
Source: Astellas (Data on File). SKYLIGHT 2 CSR.[84]

B.2.10.2.7 TEAEs of special interest

During the 12-week double-blind period, the frequency of any of the treatment-emergent AESIs were similar between treatment groups (Table 53).

Table 53: Treatment-emergent adverse events of special interest (Safety Analysis Set); 12-week double-blind period

MedDRA (v23.0) Treatment-emergent AESI	SKYLIGHT 1		SKYLIGHT 2	
	Placebo (n=175)	Fezolinetant 45 mg (n=173)	Placebo (n=167)	Fezolinetant 45 mg (n=167)
Liver test elevations[a]	5 (2.9%)	7 (4.0%)	0	3 (1.8%)
Uterine bleeding	2 (1.1%)	2 (1.2%)	1 (0.6%)	1 (0.6%)
Endometrial hyperplasia, cancer or disordered proliferative endometrium	0	0	0	0
Thrombocytopenia	1 (0.6%)	1 (0.6%)	0	0
Bone fractures	0	1 (0.6%)	1 (0.6%)	0
Potential abuse liability	0	0	1 (0.6%)	0
Depression	2 (1.1%)	3 (1.7%)	4 (2.4%)	1 (0.6%)
Wakefulness	1 (0.6%)	1 (0.6%)	1 (0.6%)	1 (0.6%)
Effect on memory	0	0	0	0

AEs were coded using MedDRA. All randomised patients who took at least one dose of study treatment (Safety Analysis Set). [a]Liver tests included ALT, AST, ALP and TBL.
Abbreviations: AESI: adverse event of special interest; ALP: alkaline phosphatase; ALT: alanine aminotransferase; AST: aspartate aminotransferase; TBL: total bilirubin.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs.[83, 84]

B.2.10.2.8 TEAEs leading to withdrawal of treatment

TEAEs leading to withdrawal of treatment during the 12-week double-blind period of SKYLIGHT 1 and SKYLIGHT 2 were reported in 9 and 1 patient in the placebo group and 4 and 5 patients in the fezolinetant 45 mg group, respectively, as described in Table 48.

No specific TEAE leading to withdrawal of treatment (by preferred term) was reported for >2 patients in the fezolinetant 45 mg group; nausea and headache leading to withdrawal of treatment were reported for 3 patients (1.7%, each) in the placebo group. In SKYLIGHT 2, no specific TEAE leading to withdrawal of treatment was reported for >2 patients in any treatment group.

Drug-related TEAEs leading to withdrawal of treatment were reported for 7 patients and 0 patients in the placebo group in SKYLIGHT 1 and SKYLIGHT 2, respectively, 3 and 5 patients in the fezolinetant 45 mg group of SKYLIGHT 1 and SKYLIGHT 2, respectively.

B.2.10.2.9 Other safety evaluations

CNS safety

No suicidal ideation or behaviour or self-injurious behaviour with suicidal intent as measured by the Columbia Suicide Severity Rating Scale were reported during the 12-week double-blind period of SKYLIGHT 1, except for 1 patient in the fezolinetant 45 mg group (suicidal ideation at Week 12). In SKYLIGHT 2, 1 patient in the placebo group (suicidal ideation at baseline and Week 12) and 1 patient in the fezolinetant 45 mg group (suicidal ideation at Week 12). In both trials, none of the patients engaged in non-suicidal self-injurious behaviour.

Endometrial safety

In SKYLIGHT 1 and SKYLIGHT 2, TVU was required for all patients to assess endometrial thickness at screening and changes from baseline in endometrial thickness are described in Table 54. TVU data were generally unremarkable and there were no clinically relevant differences in the change from baseline to Week 52 in endometrial thickness across all treatment groups. In either trial, half or more than half of patients in any treatment group with evaluable or available data were assessed for endometrial thickness. The most common reason for missing TVU assessments was a history of hysterectomy since TVU was not required for patients who had a partial (supra-cervical) or full hysterectomy Week 52 or ED visit for patients who withdrew from the study. The rest of the patients did not have a history of hysterectomy, but they were missing a TVU assessment at baseline or postbaseline or both.

Table 54: Endometrial thickness: change from baseline to Week 52 (Safety Analysis Set, Fezolinetant); 52-week period

Statistic	SKYLIGHT 1		SKYLIGHT 2	
	Placebo/ Fezolinetant 45 mg (n=76)	Fezolinetant 45 mg (n=173)	Placebo/ Fezolinetant 45 mg (n=75)	Fezolinetant 45 mg (n=167)
History of hysterectomy[a]	24 (31.6%)	57 (32.9%)	23 (30.7%)	56 (33.5%)
Evaluable patients/ available data at baseline	43	106	49	103
Evaluable patients/ available data at Week 52	39	77	43	80
Endometrial thickness (mm)				
Mean (SD)	0.21 (2.19)	−0.42 (1.84)	−0.15 (2.62)	−0.46 (1.80)
Median	−0.03	−0.14	0.25	−0.20
Min–Max	−7.8–7.5	−6.4–4.9	−12.5, 5.1	−6.5, 3.1

All randomised patients who took at least one dose of fezolinetant during either the 12-week double-blind or active treatment extension periods (Safety Analysis Set, Fezolinetant). Results in this table include only data from fezolinetant exposure. For the placebo/fezolinetant group, data from the 12-week placebo-controlled period (placebo data) were not included. These patients were re-randomised in a 1:1 ratio to 30 mg or 45 mg of fezolinetant after 12 weeks of placebo and only data from the active treatment extension period are analysed for these patients (up to 40 weeks of fezolinetant exposure). For the fezolinetant groups, data from the 12-week period were included as these patients were on fezolinetant for the entire duration of the study. [a]In SKYLIGHT 1, 6 patients from the placebo group who had a hysterectomy are not included in this table since they did not receive a dose of study treatment during the active treatment extension period and 1 patient from the fezolinetant 45 mg group who had a hysterectomy was included in this table during the 52-week period. In SKYLIGHT 2, 3 patients from the placebo group who had a hysterectomy are not included in this table since they did not receive a dose of study treatment during the active treatment extension period. Therefore, the number of patients in SKYLIGHT 1 and 2 in this table differs from those in Table 11.
Abbreviations: Max: maximum; Min: minimum; SD: standard deviation.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs.[83, 84]

In SKYLIGHT 1 and SKYLIGHT 2 trials, 2 and 1 patient in the fezolinetant 45 mg group (up to 52 weeks of fezolinetant exposure) were classified with a final diagnosis of disorder proliferative pattern, respectively. In each trial, 1 patient in the placebo/fezolinetant 45 mg group (up to 40 weeks of fezolinetant exposure) was classified with a final diagnosis of disorder proliferative pattern. None of the patients had a final diagnosis of malignancy during the 52-week period.

Liver tests

During the 12-week double-blind period of SKYLIGHT 1, there was 1 patient with ALT or AST levels >3 x ULN in the placebo group and 0 patients in the fezolinetant 45 mg group. In SKYLIGHT 2, 1 patient in the placebo group and 3 in the fezolinetant 45 mg group had ALT or AST levels >3 x ULN.

None of the patients had bilirubin levels >2 x ULN in either trial. There were also no cases of Hy's law (definition provided in ‘Outcome definitions’ Section B.2.3.3).

B.2.11 Ongoing studies

The fezolinetant clinical development programme comprises 19 completed studies (11 phase 1 studies, 4 phase 2 studies and 4 phase 3 studies) to date. In addition to these, a randomised, placebo-controlled, double-blind, phase 3 clinical study (NCT06440967) has been initiated to investigate the efficacy and safety of fezolinetant for treatment of moderate to severe VMS in women with Stage 0 to 3 hormone receptor-positive breast cancer who are receiving adjuvant endocrine therapy.[94]

B.2.12 Interpretation of clinical effectiveness and safety evidence

B.2.12.1 Principal findings from the clinical evidence base

The efficacy and safety of fezolinetant in treating moderate to severe VMS in postmenopausal women have been assessed through a comprehensive phase 3 clinical development programme, encompassing four confirmatory RCTs. The phase 3b trial, DAYLIGHT, along with the pivotal phase 3 trials, SKYLIGHT 1 and SKYLIGHT 2, have been summarised below, given their direct relevance to the decision problem. Although not of direct relevance to the submission, confirmatory results from SKYLIGHT 4 on the long-term safety of fezolinetant have been summarised in the appendices for completeness.

Baseline demographics of patients in the SKYLIGHT trials were broadly generalisable to those observed in UK clinical practice. For DAYLIGHT, clinical expert opinion sought by Astellas noted that the patients in the trial were older than patients seen in UK practice. Given that the target HRT-unsuitable for medical reasons population in this submission includes patients who had stopped previous HRT (HRT-stoppers), and that a considerable proportion of women across the Phase 3 trial programme had prior HRT use, Astellas believe that it is likely that the populations in these studies are indicative of the type of women that will initiate fezolinetant i.e., not just a homogeneous recently diagnosed population.

The DAYLIGHT trial demonstrated the efficacy and safety of fezolinetant in addressing moderate to severe VMS in postmenopausal women, for those deemed unsuitable for HRT. Administered at a dosage of 45 mg once daily, fezolinetant met its primary endpoint at Week 24, demonstrating statistically significant reductions in both the daily frequency and severity of moderate to severe VMS when compared to the placebo. A response was observed as early as Week 1 following treatment initiation, with continued improvements through Week 4 and a sustained benefit throughout the 24-week double-blind, placebo-controlled period. This reduction in daily VMS frequency and severity translated to substantial improvements in sleep, quality of life, daily functioning and psychological outcomes, as measured by PROMIS SD SF 8b, EQ-5D-5L VAS, WPAI-VMS and PHQ-4. Of note, patients on fezolinetant reported reduced sleep disturbance, improved health state and daily functioning and a reduced likelihood of suffering from anxiety and depression, demonstrating the broader benefits of fezolinetant 45 mg. Furthermore, the safety data were consistent with the established safety profile of fezolinetant. Through Week 24, there were no discernible safety concerns associated with the 45 mg fezolinetant dose.

The results of DAYLIGHT were consistent with the findings from the pivotal SKYLIGHT 1 and SKYLIGHT 2 trials, in which fezolinetant 45 mg met the four co-primary endpoints by demonstrating a statistically significant reduction in both daily frequency and severity of VMS at Weeks 4 and 12 compared with placebo. Similarly, a response was observed within a week since treatment initiation, sustained through to Week 52. This led to meaningful improvements in both quality of life and daily functioning, as measured by EQ-5D-5L VAS and WPAI-VMS. The key secondary endpoint was also consistent with the co-primary efficacy results. Of note, fezolinetant 45 mg demonstrated improvements in sleep disturbance compared with placebo at Week 12, with benefits sustained through Week 52. Subgroup analyses revealed consistent efficacy across various parameters, including age, BMI, time to VMS onset, and severity of VMS at onset and prior HRT history. Across both studies, treatment with fezolinetant 45 mg was well-tolerated with an unremarkable safety profile. Additionally, in the fezolinetant 45 mg group, the incidence of TEAEs was low and comparable to placebo.

Finally, the results from the SKYLIGHT 4 trial were supportive of conclusions of comparable long-term safety between fezolinetant and placebo.

B.2.12.2 Strengths and limitations of the clinical evidence base

As discussed in Section B.2.5, the DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 trials were deemed to be methodologically robust and well reported and the results of which were therefore considered to be at a low risk of bias.

All three trials were sufficiently powered to detect a true treatment effect, owing to the large sample size. The trials collectively provide data on fezolinetant efficacy and safety in more than

1,000 postmenopausal women with moderate to severe VMS. The enrolled population of DAYLIGHT was aligned with the population addressed in the decision problem. In SKYLIGHT 1 and SKYLIGHT 2, although specific restrictions on HRT suitability were not imposed, most women (n=875, 85.6%) were deemed HRT-unsuitable. Pre-specified subgroup analyses were conducted for these patients who were deemed HRT-unsuitable, in a pooled analysis of SKYLIGHT 1 and SKYLIGHT 2, to further support any conclusions reached following analyses of the data reported in DAYLIGHT.

Overall, baseline characteristics were broadly similar across treatment groups in each trial, and, at the time of the primary analysis, treatment discontinuation was low and broadly similar across treatment groups, which limited any imbalances that would have otherwise been caused by this. Furthermore, appropriate estimand strategies were used to handle post-randomisation events, limiting further biases. While a high placebo effect was observed in the reduction of moderate to severe VMS frequency and severity in the placebo group, significant differences favouring fezolinetant 45 mg versus placebo were still evident at the primary endpoint analysis of all three trials, underscoring the robustness of its efficacy in addressing moderate to severe VMS.

Additionally, the long-term data from the SKYLIGHT trials, spanning a follow-up period from Week 13 to Week 52, demonstrate that the therapeutic effect of fezolinetant 45 mg was consistently maintained throughout this period. This sustained efficacy supports the expectation that the real-world effectiveness of fezolinetant will likely mirror the outcomes observed during this phase of the trials, particularly since participants were aware they were receiving the active treatment.

A conceptual framework was developed through an international steering group to define the key treatment priorities of menopausal women with VMS. [95] In line with this Core Outcome Set (COS) for VMS, the trials considered a diverse range of clinically relevant endpoints, including frequency and severity of moderate to severe VMS, sleep disturbance, quality of life, daily functioning and psychological outcomes.[95] These outcomes hold paramount importance due to their high prevalence among people who are seeking treatment for menopausal symptoms and due to their importance on continuity of healthcare and further impact on long-term outcomes. As noted in Section B.1.3.2, VMS is the most common and bothersome symptom of menopause, which negatively impacts other aspects of health, including sleep, mental health and daily functioning. As such, the benefits demonstrated in these trials will therefore translate to meaningful improvements for patients in clinical practice.

While the trials discussed in support of fezolinetant (DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2) predominantly enrolled a white population, the absence of treatment-modifying effects based on race implies that the effect of fezolinetant 45 mg is anticipated to be generalisable to other ethnic groups. This includes people of African and Caribbean descent, who tend have a higher prevalence, longer duration and greater severity of moderate to severe VMS.[49]

As discussed in Section B.1.3.3, no active treatment represents the relevant comparator to fezolinetant in people with moderate to severe VMS for whom HRT is deemed unsuitable for medical reasons (HRT-contraindicated, HRT-caution and HRT-stoppers). As such, clinical evidence informing the no active treatment arm of the model base case included efficacy data from the placebo arm of the DAYLIGHT trial. However, there are limitations to modelling no active treatment using placebo trial data. As noted in Section B.2.6.1.1, patients in the placebo arm experienced improvements in VMS frequency from baseline through Week 24 in DAYLIGHT despite receiving pharmacologically inert tablets. While high placebo effects are common in

clinical trials in moderate to severe VMS, there are issues regarding the generalisability of these results to no active treatment, given that the target population are unlikely to be receiving placebo in UK clinical practice for ethical reasons and additionally, they are also unlikely to be receiving any additional follow up or clinician interactions associated with partaking in a clinical trial. Consequently, the use of placebo efficacy data to inform the no active treatment arm of the model is, at best, conservative (Section B.3.3.2.2).

B.2.12.3 Conclusion

Fezolinetant is a first-in-class nonhormonal therapy that is proven to significantly reduce both the frequency and severity of moderate to severe VMS, which translate to improved outcomes in terms of sleep, quality of life, daily functioning and mental health. Results from DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 demonstrate the persistence of response to fezolinetant 45 mg from baseline up to Week 52. The mechanism of fezolinetant is targeted, acting specifically on the thermoregulatory pathways that underpin menopause-related VMS. Therefore, the use of fezolinetant in people with moderate to severe VMS for whom HRT is deemed unsuitable for medical reasons would represent a step change in treatment care pathway, given the current unmet need for non-hormonal alternatives to HRT and the current lack of robust, licensed and evidenced-based therapies for treating the moderate to severe VMS. The reimbursement of fezolinetant would also address the urgent concerns of patients with moderate to severe VMS who do not have access to HRT or other non-hormonal therapies, either due to unsuitability to HRT or barriers to access, respectively.

B.3 Cost effectiveness

Summary of the cost-effectiveness analysis

De novo cost-effectiveness model

- A *de* novo economic model was developed to assess the cost effectiveness of fezolinetant for treatment of moderate to severe VMS in menopausal people who are deemed unsuitable for HRT for medical reasons.
- The model used a Markov cohort structure to best capture the key features of moderate to severe VMS and to reflect the clinical pathway of care for the target population of relevance to the decision problem.
- The model structure comprised six health states. Four health states were defined based on average daily moderate to severe VMS frequency and two health states captured natural VMS cessation and death (absorbing state), respectively. The use of a granular, VMS frequency health state model was validated by three UK clinical experts who considered the modelling approach appropriate, and further validated by four patient experts.[7, 12]
- Upon entering the model, the modelled cohort either received fezolinetant or no active treatment and were distributed across the on-treatment VMS frequency health states based on the observed baseline distribution in DAYLIGHT.
- On-treatment VMS frequency health state transition probabilities were sourced primarily for DAYLIGHT, with supporting efficacy inputs from pooled SKYLIGHT 1 and SKYLIGHT 2 trials. In the on-treatment VMS frequency health states, the model cohort either remain in the same health state or transition to other on treatment health states until treatment discontinuation, VMS cessation, death, or the end of the model time horizon.
- All-cause discontinuation rates were used to model transitions to the off-treatment VMS frequency health states. In the off-treatment VMS frequency health states, the modelled cohort followed the natural history of VMS until VMS cessation, death, or the end of the model time horizon.
- The natural history distribution captured the variation of VMS for the modelled cohort in the off-treatment VMS frequency health states, representing the 'true' no active treatment effect for menopausal people with moderate to severe VMS not receiving any active treatments. These distributions were based on inputs derived from structured expert elicitation (SEE) estimates, aligned with UK clinical expert and patient expert opinion.
- The VMS cessation health state ensured that the model captured the natural cessation of VMS in menopausal women. People in any VMS frequency health states (whether on treatment or off treatment) can experience cessation of VMS, after which VMS cannot return (and treatment is discontinued if people were previously receiving treatment). These people remain in the VMS cessation health state until death or until the end of the model time horizon, whichever comes first.
- Death was represented by an absorbing health state, accumulating patient flows from all other health states, and patients were at risk of death at any time during the simulation, as modelled by UK general population mortality.
- In line with the NICE reference case, the cost perspective adopted was of the NHS and personal social services (PSS), and the outcomes perspective was all relevant health effects over a 10-year time horizon. Costs and outcomes were discounted at 3.5% per annum.
- Utility values for moderate to severe VMS frequency health states were derived from ED-5D-5L data collected directly from patients in the DAYLIGHT trial and mapped onto EQ-5D-3L. However, EQ-5D has been observed to be insensitive to treatments for VMS. As such, utility values were adjusted in line with the views of three UK clinical experts to best reflect values that would be expected in UK clinical practice.[7] These adjusted values were validated by four patient experts.[12]
- Health state unit costs and resource use were sourced from NHS reference costs (2021/22) and Personal Social Services Research Unit (PSSRU 2023),[96] with inputs from UK expert elicitation and validation by patient experts where necessary.[12, 91]

Base case cost-effectiveness results

- The base case probabilistic ICER for fezolinetant 45 mg versus no active treatment was £10,355 per QALY gained, which is below the £30,000 threshold considered cost effective.
- Overall, the results demonstrate that fezolinetant is a cost-effective option for the treatment of moderate to severe VMS versus no active treatment, the comparator relevant to UK clinical practice.

Sensitivity and scenario analysis

- The base case probabilistic ICER did not differ meaningfully from the deterministic ICER (£10,364 per QALY gained). The probabilistic sensitivity analysis (PSA) results indicated that the base-case results showed little variation with regard to incremental costs in each iteration, as most iterations generally fell between £1,000 and £1,500. On the effectiveness axis, all iterations fell within the northeast quadrant and below the £30,000/QALY cost-effectiveness threshold, despite the modest variation in incremental QALYs gained.
- The deterministic sensitivity analyses (DSA) results showed that the incremental net health benefit (INHB) was most sensitive to health state utility values for VMS frequency health states and, to a lesser extent, health care resource use. None of the ten variables crossed the point of indifference (i.e., when INHB is zero) for either their upper bound or lower bound values.
- The probabilistic scenario analyses results were all below the cost-effectiveness threshold used by NICE £30,000/QALY, providing reassurance that the results of the base-case analysis were robust to uncertainties in both model inputs and assumptions.

Conclusion

- The cost-effectiveness analysis demonstrates fezolinetant to be a cost-effective use of NHS resources versus no active treatment. Overall, fezolinetant represents a step change in the treatment care pathway for moderate to severe VMS, given the current unmet need for non-hormonal alternatives and the lack of robust, licensed, and evidenced-based therapies.

B.3.1 Published cost-effectiveness studies

A *de novo* SLR was conducted to identify economic evaluations for the treatment of VMS associated with menopause. The original SLR searches were performed in June 2023 and updated in April 2024.

In the original SLR, a total of 2,865 records were retrieved from the electronic databases, of which 864 were duplicates, resulting in 2,001 novel records that were screened at the title/abstract review stage. Supplementary searches (HTA body websites, economic websites and conference searches) retrieved no additional articles relevant for inclusion. It is worth noting that although the NICE NG23 clinical guideline was identified, it was excluded because the publication type of the supplementary searches was restricted to HTA submissions. The bibliographies of six relevant SLRs/(N)MAs or HTAs were searched, which identified an additional 26 articles, but these were all deemed ineligible for inclusion in the review. No articles reporting economic evidence were included at the abstract review stage, hence no records were excluded at full text review.

In the SLR update, a total of 3,085 records were retrieved from the electronic database, of which 2,879 were duplicates (including duplicates from the original search), resulting in 206 novel

records being screened at the title/abstract review stage, with one suitable* article identified. Supplementary searches (HTA body websites, economic websites and conference searches) retrieved no additional articles relevant for inclusion. The bibliographies of one relevant HTA was searched, which identified six additional articles, but these were all deemed ineligible for inclusion in the review. Details of the economic SLR can be found in Appendix G.

In summary, one article reporting economic evaluation evidence was included in the SLR update, summarised in Table 55.

Table 55: Summary list of published cost-effectiveness studies

Study	Summary of model	Patient population (average age in years)	QALYs (intervention, comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
Beaudoin (2023)[97]	• Cost-effectiveness • *De novo* decision analytical model using an intention-to-treat analysis informed by clinical trials and previous relevant economic models • Model health states: on treatment: responding or not responding (those that discontinued due to the intervention not improving symptoms); off treatment: (discontinued due to symptom resolution); all-cause death • Intervention: Fezolinetant • Comparators: No pharmacologic treatment (prescription or non-prescription)/MHT • Time horizon: Lifetime horizon • Cycle length: Model cycle length varied in the first year (e.g., 3 months to correspond with response rates from trials) but then converged on an annual model cycle length there after • Source of inputs: Model inputs were estimated from the clinical review, published literature and information from stakeholders. • Clinical: Duration of VMS and treatment sourced from Avis *et al.* 2015[98]; on treatment health	Women with VMS associated with menopause being treated with fezolinetant and comparators. Women had an approximate age of 54 years with a wide range of months since onset of VMS. The population focus included patients comparable to those in the fezolinetant clinical studies.	**Fezolinetant**: QALY: 16.43 Life years: 19.88 Equal value life years: 16.43 **No pharmacologic treatment** QALY: 16.33 Life years: 19.88 Equal value life years: 16.33 **Incremental (Fezolinetant versus No pharmacological treatment)** QALY: 0.10 Life years: 0.00 Equal value life years: 0.10	**Fezolinetant**: Intervention cost: $45,000[a] Total cost: $198,000 **No pharmacologic treatment** Intervention cost: $0 Total cost: $157,000 **Incremental (Fezolinetant versus No pharmacological treatment)** Intervention cost: $45,000[a] Total cost: $41,000 **Other Non-intervention costs** Fezolinetant: $153,000 No pharmacologic treatment: $157,000 Incremental (Fezolinetant versus No pharmacological treatment): -$4,000	Treatment: Fezolinetant Comparator: No pharmacologic treatment Cost per QALY Gained: $390,000 Cost per equal value life years: $390,000 Cost per symptom-free day: $500

	state utility sourced from Coon *et al.* 2018[99]; discontinuation in first year of model; key treatment effects from fezolinetant clinical studies (including reduction in VMS frequency and severity) • Utility: Evidence from a mapping instrument between the MENQoL and EQ-5D was used to derive utility scores and differences across treatment and no treatment of VMS. • Costs/resource use: Focus was on direct medical care costs and included drug costs (Fezolinetant: $6,000.00; MHT: $123.45) sourced from the placeholder price and IBM Micromedex, respectively; mean VMS-related direct costs for treated and untreated patients per year ($1,731) sourced from Sarrel *et al.* 2015.[100] Costs for managing and treating future unrelated complications were also included • Epidemiology: NR • Treatment sequencing: NR • Discount rate: 3% per year • Cost year: NR • One way sensitivity analysis and probabilistic sensitivity				

[a]Based on the annual placeholder price of $6,000.00. Interpret cost with caution.
Abbreviations: DSA, deterministic sensitivity analysis; ICER, incremental cost-effectiveness ratio; NR, not reported; NHS, national health service; PSA, probabilistic sensitivity analysis; QALY, quality adjusted life year; VMS: vasomotor symptoms .

B.3.2 Economic analysis

As noted in Section B.3.1, one study was identified in the economic SLR that assessed the cost-effectiveness of fezolinetant.[101] However, the model structure (three mutually exclusive and exhaustive health states: on treatment, off treatment and all-cause death) reported in the base case analysis was deemed overly simplistic and failed to reflect the value that fezolinetant is expected to offer to menopausal people with moderate to severe vasomotor-predominant symptoms who are deemed unsuitable for HRT for medical reasons. The model assumes that people who respond to therapy remain on treatment until symptoms resolve, without accounting for intermittent symptom recurrence or varying responses to treatment over time. Additionally, the model does not incorporate dynamic transitions between different health states based on changing conditions and treatment responses, which can lead to oversimplified and potentially less accurate outcomes.

As such, a *de novo* cost-utility analysis, in line with the NICE reference case, using a Markov cohort model structure was considered to best capture the key features of moderate to severe vasomotor-predominant menopausal symptoms and the current clinical pathway of care for the target population addressed in the decision problem. This model type can incorporate probabilities of transitioning between health states, thus offering a more detailed and realistic projection of long-term costs and health outcomes. By capturing these dynamic changes, a Markov cohort model enhances the robustness and reliability of the cost-effectiveness analysis for fezolinetant.

B.3.2.1 Patient population

In line with the target population of relevance to the decision problem discussed in Section B.1.1, the economic analysis evaluates the cost-effectiveness of fezolinetant in a cohort of menopausal people with moderate to severe VMS for whom HRT is deemed unsuitable for medical reasons.

The DAYLIGHT trial provides direct clinical evidence in the target population, as women enrolled in the trial had moderate to severe VMS and were deemed HRT-unsuitable. Given that DAYLIGHT provides trial data for up to 24 weeks, it was necessary to source model inputs from the pooled SKYLIGHT 1 and 2 trial with a longer follow-up of 52 weeks. As such, to be consistent with the target population of this submission, pooled SKYLIGHT 1 and 2 trial data was sourced from the subpopulation of women who were deemed HRT-unsuitable. This approach was deemed appropriate as most women (n=875, 85.6%) were deemed unsuitable for HRT in the pooled trials, providing reassurance that the sample size of the efficacy data is still suitably large to inform the model base case.

Overall, both DAYLIGHT and pooled SKYLIGHT 1 and 2 trials are considered generalisable to the target population of relevance to the decision problem and thus suitable to inform the modelled cohort. Further information on the HRT-unsuitable subpopulation from SKYLIGHT 1 and SKYLIGHT 2 can be found in Section B.2.7.1.2.

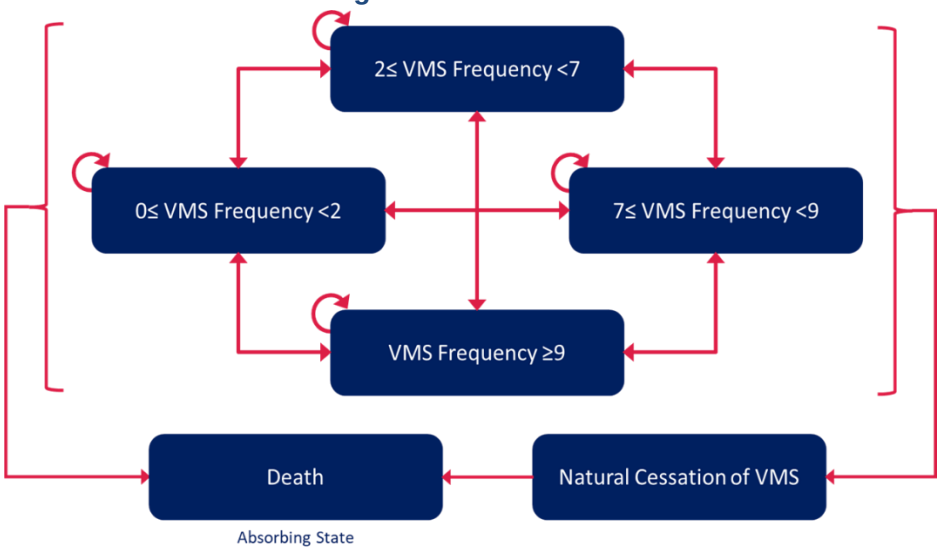
B.3.2.2 Model structure

The model comprised six health states. Four health states were defined based on average daily moderate to severe VMS frequency, while the remaining two health states captured natural VMS cessation and death (absorbing state).

The use of a granular, VMS frequency health state model was chosen to best capture the health effects and costs of fezolinetant versus no active treatment in VMS. The model structure and the use of VMS frequency to define health states was validated by three UK clinical experts who considered both methods appropriate for evaluating the cost-effectiveness of fezolinetant in this indication.[7] UK clinical experts emphasised that VMS frequency was the most objective measure of the health effects of treatment and confirmed that the method used to define VMS frequency health states was appropriate.[7] Four patient experts were consulted in August 2024 and also agreed that the model structure based on VMS frequency health states was appropriate.[12] The published literature and clinical guidelines, such as NICE NG23, support this decision by recognising VMS frequency as a proxy for VMS severity.[15, 102, 103] For instance, NICE NG23 states that "*three flushes per day would be considered as a mild/moderate level of flushing.*"[15] Further information on the health states used in the model is detailed in Section B.3.2.2.1.

The health state transition diagram for people receiving fezolinetant and no active treatment in the base case model is presented in Figure 25.

Figure 25: Health state transition diagram



Abbreviations: VMS: vasomotor symptoms

Upon entering the model, the modelled cohort received either fezolinetant or no active treatment and were distributed across the VMS frequency health states as observed in DAYLIGHT. A scenario analysis was performed that applies the baseline distribution from pooled DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 trials (Section B.3.11.3).

At each four-week cycle, menopausal people may remain in the same health state or transition to any other health state. Menopausal people in any of the VMS frequency health states (whether on treatment or off treatment) can experience cessation of VMS, after which VMS cannot return (and treatment is discontinued if people were previously receiving treatment). These menopausal people remain in the VMS cessation health state until death or until the end of the model time horizon, whichever comes first. The VMS cessation health state ensured that the model captured the natural cessation of VMS in menopausal women. The per cycle probability of VMS cessation was derived using the median duration of VMS reported by Avis *et al.*, (2015).[49] Cessation of VMS inputs are summarised and justified in B.3.3.2.5.

For menopausal people receiving fezolinetant, VMS frequency transition probabilities were derived from Week 0–24 DAYLIGHT efficacy data and Week 24–52 pooled SKYLIGHT 1 and SKYLIGHT 2 efficacy data (from the HRT-unsuitable subpopulation), detailed in Section B.3.3.2.1. For those receiving no active treatment, VMS frequency transition probabilities were derived from Week 0–12 DAYLIGHT placebo arm data, detailed in Section B.3.3.2.2.

The natural history distribution captured the variation of VMS for menopausal people in the off-treatment VMS frequency health states. These distributions were based on structured expert elicitation (SEE) estimates, representing the 'true' no active treatment effect for women with VMS not receiving any active treatments. Natural history distribution is further detailed in Section B.3.3.2.3.

For menopausal people in the fezolinetant arm, all-cause discontinuation rates were used to model transitions to the off-treatment VMS frequency health states. In the off-treatment VMS frequency health states, menopausal people followed the natural history of VMS until VMS cessation, death, or the end of the model time horizon. Discontinuation rates up to Week 24 were derived from Week 0–24 DAYLIGHT data, and rates from Week 24 onwards were based on pooled Week 24–52 SKYLIGHT 1 and 2 data (HRT-unsuitable subpopulation). In the no active treatment arm, discontinuation rates from Week 0–24 DAYLIGHT data were applied up to Week 12, after which all remaining menopausal people transitioned to the off-treatment VMS frequency health states, in which they revert to the natural history distribution and then follow natural history as described above for fezolinetant. Treatment discontinuation is detailed in Section B.3.3.2.4.

No adverse events were considered in the model, further information on safety is included in Section B.3.3.2.6.

Death was represented by an absorbing health state, accumulating patient flows from all other health states, and patients were at risk of death at any time during the simulation, as modelled by UK general population mortality.[104] Menopausal people may transition to the death state from any of the five health states and remain there for the remainder of model horizon. The approach to modelling mortality is detailed in Section B.3.3.2.7.

The model also considered resource use by VMS frequency health states (including VMS cessation), as detailed in Section B.3.5.

Features of the economic analysis

The key features of the economic analysis and their justifications are presented in Table 56.

Effectiveness measures include life years (LYs) and quality-adjusted life years (QALYs). The incremental cost-effectiveness ratio (ICER) of fezolinetant versus no active treatment was evaluated in terms of the incremental cost per QALY gained. An annual discount of 3.5% was applied for both costs and QALYs. The cost perspective was of the NHS and personal social services (PSS), and the outcomes perspective was all relevant health effects over a 10-year time horizon. A time horizon of 10 years was considered appropriate, given that there is no differential mortality effect between technologies and the differences in costs and clinical outcomes relate to a relatively short period, in line with the NICE reference case. The model considers one scenario analysis of a shorter 3-year time horizon to assess the impact of assumptions around the time horizon on the model results (see Section B.3.11.3).

Health state utility values were derived based on EQ-5D-5L data sourced directly from DAYLIGHT and mapped onto EQ-5D-3L, in line with the NICE reference case. These DAYLIGHT utility values were adjusted based on feedback from UK clinical experts, which were validated by patient experts who deemed the estimates conservative,[12] details of which can be found in Section B.3.4.5.

Table 56: Features of the economic analysis

	Current evaluation	
Factor	**Chosen values**	**Justification**
Time horizon	10 years	There is no differential mortality effect between technologies and the differences in costs and clinical outcomes relate to a relatively short period. The median treatment duration for fezolinetant is modelled to be approximately [redacted] months ([redacted] years). In line with the NICE reference case, a time horizon shorter than a lifetime was justified. This approach was further validated by three UK clinical experts who indicated that a 10-year time horizon was appropriate, given that they expect the vast majority of people to have discontinued treatment by the 10-year timepoint.[7] UK clinical experts further confirmed that the modelled treatment duration was in line with the expected treatment duration for fezolinetant in UK clinical practice.[7] A scenario is provided that applies a 3-year time horizon, in order to remove the uncertainty associated with the natural history of VMS applied in the post-discontinuation health states, which is detailed in Section B.3.11.3.
Cycle length	4 weeks	A four-week cycle length was used in the model in line with the clinical assessment timepoints used in DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2.
Treatment waning effect?	Fezolinetant: treatment waning effect is captured in the treatment discontinuation rate No active treatment waning of placebo effect up to Week 12 is captured by applying a treatment discontinuation rate in each model cycle. At the end of Week 12, the placebo effect is lost, and all menopausal	It is expected that loss of treatment response in menopausal people with VMS is clinically measurable and will therefore lead to treatment discontinuation. Discontinuation was therefore considered a suitable proxy for treatment waning. The decision not to model treatment waning for fezolinetant was validated by UK clinical experts as part of NICE's early engagement process through the OMA, who indicated that it was reasonable to assume no treatment waning while on fezolinetant due to its mechanism of action; 52 weeks of fezolinetant efficacy data from pooled SKYLIGHT 1 and SKYLIGHT 2 trials showed that fezolinetant is effective at both eliciting symptom relief from day one of treatment and maintaining that response up to Week 52, with no indication of treatment waning (see B.3.3.2.4 for further details).

	people follow natural history	For the no active treatment arm, all-cause discontinuation rates were also applied to menopausal people in the on-treatment VMS frequency health states to reflect the proportion of people returning to their natural history distribution in each model cycle up to Week 12 (Section B.3.3.2.4). All patients remaining in the on-treatment VMS frequency health states at the end of Week 12 immediately discontinue treatment and revert to natural history, then follow their natural history trajectory until VMS cessation, death, or the end of the model time horizon. This modelling approach was in line with the views of UK clinical experts who unanimously confirmed that they did not expect long-term placebo effect to persist in clinical practice, which was corroborated by patient experts who also agreed that they do not expect the placebo effect to persist (Section B.3.3.2.2).[7, 12]
Source of utilities	EQ-5D-5L data were collected directly from patients enrolled in DAYLIGHT	In line with the NICE reference case, EQ-5D data informed the base case economic analysis. However, health state utility values derived from DAYLIGHT EQ-5D data were adjusted in line with the views of UK clinical experts; this adjustment was required as a result of the insensitivity of the EQ-5D measure to the improvements brought by pharmacological treatments in menopausal women, as noted in the published literature.[105, 106] These clinician-adjusted values were further validated by four patient experts, who felt the estimates to be conservative.[12]This is detailed in Section B.3.4.5. To explore the impact of unadjusted EQ-5D utilities on the economic analysis, scenario analyses using utilities from DAYLIGHT and the DAYLIGHT and SKYLIGHT pooled analysis were also conducted. These are detailed in Section B.3.11.3.
Source of costs	Healthcare resource unit costs were sourced from PSSRU and NHS costs	In line with the NICE reference case. This is detailed in Section B.3.5.
Health effects measure	QALYs	In line with the NICE reference case.
Half-cycle correction	Yes	A half-cycle correction was applied to the calculation of LYs, QALYs and costs to account for events that occur part way through a cycle.

Abbreviations: EQ-5D: EuroQoL-5 Dimension; HRQoL: health-related quality of life; LYs: life years; NHS: National Health Service; NICE: National Institute for Clinical Excellence; OMA: Office for Market Access; PSSRU: Personal Social Services Research Unit; QALYs; quality-adjusted life-years; SAE: serious adverse event; UK: United Kingdom; VMS: vasomotor symptoms.

B.3.2.2.1 Health states

Justification for defining VMS health states based on frequency rather than severity

During the early model conceptualisation, feedback from expert interviews conducted in 2021 and 2022 indicated that the use of VMS frequency to define health states in the model was

appropriate.[7] This approach was supported by the fact that VMS frequency is a commonly reported outcome in the published literature and is considered a more objective measure than VMS severity. Experts also highlighted the heterogeneity in definitions of VMS severity in the literature, which restricts data combination and comparison of results across trials.

Given that some experts suggested capturing VMS severity alongside VMS frequency in the model, Astellas assessed the feasibility of directly capturing both outcomes in the model. In earlier versions of the model, health states were defined using the hot flash score (HFS) outcome, which includes both VMS frequency and severity, based on a phase 2a proof of concept study (ESN364_HF_204).[107] However, regulatory guidance from the US Food and Drug Administration (FDA) and European Medicines Agency (EMA) emphasise the assessment of VMS frequency as the primary endpoint for efficacy trials.[108, 109] Consequently, the phase 3 confirmatory trials for fezolinetant did not use HFS as the primary or key secondary outcome, but measured VMS frequency and severity separately in DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2. This separation of outcomes limits the construction of a singular endpoint that captures both outcomes. It would in theory be possible to combine both VMS frequency and severity into a single outcome; however, it would not be possible to then generate statistical measures of dispersion around the mean, such as SE and SD, because data would only be available for SE and SD separately for both outcomes. These measures of dispersion are critical because they inform the probability sensitivity analysis (PSA). Additionally, given that VMS frequency is more commonly reported than HFS in the published literature, variation in outcomes measures in clinical trials of interventions for VMS restricts the ability to combine data from different trials and to compare results across trials.[110]

Considering the expert opinion and limitations with *post hoc* construction of a composite endpoint based on phase 3 confirmatory trial data, Astellas deemed it appropriate to model VMS health states based on VMS frequency. The published literature and clinical guidelines, such as NICE NG23, support this decision by recognising VMS frequency as a proxy for VMS severity.[15, 102, 103] For instance, NICE NG23 states that "*three flushes per day would be considered as a mild/moderate level of flushing*".[3, 15] Recent interviews with UK clinical experts in April 2024 and patient experts in August 2024 further confirmed that using VMS frequency to define health states is appropriate.[7, 12] The patient experts unanimously agreed that frequency of VMS matters most to patients, reporting that it is more debilitating for patients to experience multiple VMS events than a single severe VMS event, and that patients find it much easier to quantify the frequency of their VMS over severity.[12]

Justification for VMS frequency health state cut-offs

In the base case, VMS frequency health states were defined as follows:

- 0 ≤ VMS Frequency < 2
- 2 ≤ VMS Frequency < 7
- 7 ≤ VMS Frequency < 9
- VMS Frequency ≥ 9

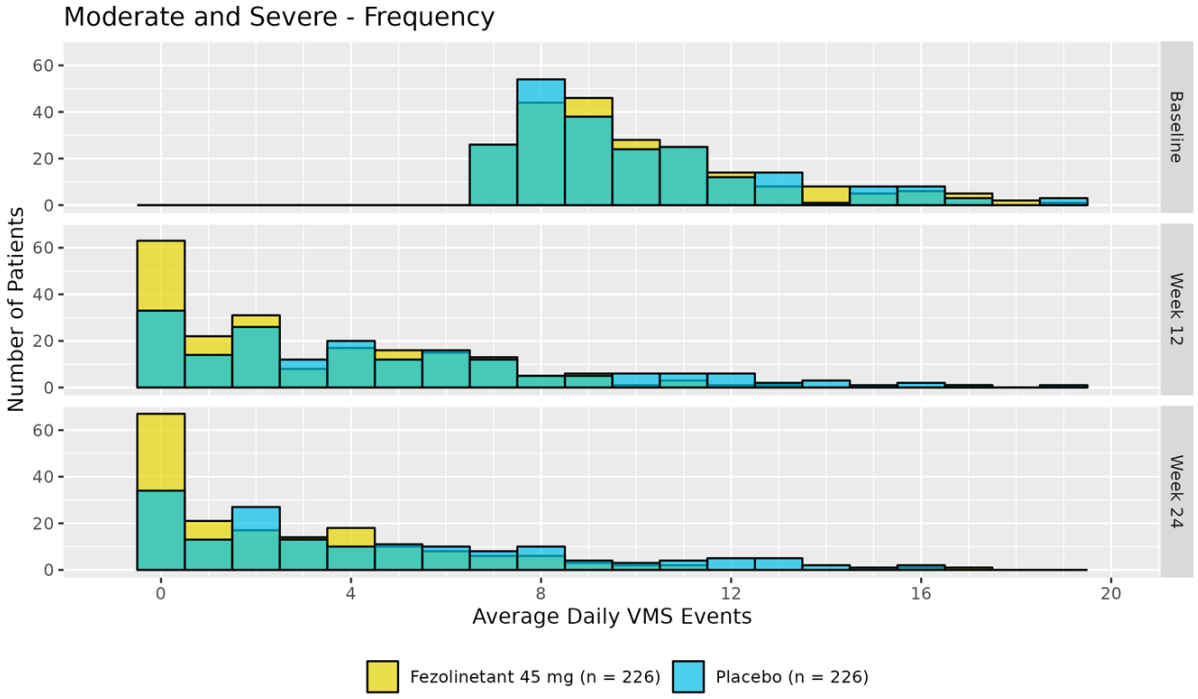
The cut-off thresholds were defined based on statistical analysis of the distribution of average daily moderate to severe VMS frequency in DAYLIGHT (Figure 26) and utility values associated with each average daily moderate to severe VMS frequency health state (Table 57), and validated by UK clinical experts for clinical plausibility.

Utility values based on UK tariffs were derived using a generalised estimating equation (GEE) model with an exchangeable correlation structure to account for the correlation within patients' repeated assessments, with VMS frequency category as the independent variable and using pooled data from all included people in the fezolinetant 45 mg and placebo arms of DAYLIGHT.

The rationale for each cut-off is provided below based on both statistical and clinical perspectives:[111]

- Lower threshold of 2:
    - Our analysis revealed no statistical difference (i.e., p>0.05) between the frequency categories of 0 and >0 to <1, supporting their combination into a single category, 0 to <1. The frequency category 1 to <2 was relatively similar to 0 to <1, justifying their combination in a broader category, 0 to <2. Additionally, the four categories (3 to <4, 4 to <5, 5 to <6, and 6 to <7) showed no significant differences, allowing their combination into a single category, 3 to <7. Moreover, given that 2 to <3 was significantly different from 0 to <2 but not from 3 to <7; 2 to <3 was combined with 3 to <7 into a single category, 2 to <7.
    - Analysis of EQ-5D utility values for the 0 to <2 group showed a statistical difference in utility (p<0.05) compared with the 2 to <7 group. As such, a VMS frequency of 2 was chosen as a cut-off. This was further validated by UK clinical expert opinion which indicated that a 75% reduction from baseline in moderate to severe VMS frequency is a clinically meaningful improvement.
- Threshold of 7:
    - Given that all people enter DAYLIGHT with a minimum average of 7 moderate to severe VMS per day, it was necessary to use this as a threshold for moderate to severe VMS frequency health state definitions to ensure that most people in the trial who improve from baseline, are modelled to experience a change of health state in the model
    - Moreover, the threshold of 7 reflects the inclusion criteria of the DAYLIGHT, which was mandated by guidance from FDA on clinical trials for VMS[109]
- Upper threshold of 9:
    - Our analysis revealed no statistical difference between the frequency categories of 7 to <8 and 8 to <9, supporting their combination into a single category, 7 to <9. There was a considerable decrease in utility values starting from the 9 to <10 category. Additionally, the categories 9 to <10, 10 to <11, 11 to <12, and 12 to <13 were statistically similar, with no significant difference when using 11, 12, or 13 as thresholds. However, using 14+ as a threshold was problematic due to a small sample size. Therefore, the 9+ categories were combined into a single category.
    - UK clinical experts confirmed that a threshold at 9 was clinically plausible.

Figure 26: VMS frequency distributions at baseline, Week 12, and Week 24, fezolinetant 45 mg and placebo in DAYLIGHT



Abbreviations: VMS: vasomotor symptoms.
Source: Astellas (Data on File). DAYLIGHT CSR.[82]

Table 57: Mean ED-5D utility values by VMS frequency based on UK tariffs and estimates using GEE models accounting for repeated measures

VMS frequency	EQ-5D utility mean	EQ-5D utility 95% CI	Number of EQ-5D observations	EQ-5D utility mean value of VMS health states (95% CI)
0	0.835	(0.808, 0.863)	244	0.833 (0.814,0.852)
>0 to <1	0.848	(0.817, 0.878)	146	
1 to <2	0.824	(0.803, 0.846)	173	
2 to <3	0.806	(0.781, 0.830)	145	0.793 (0.774, 0.812)[a]
3 to <4	0.793	(0.765, 0.821)	115	
4 to <5	0.795	(0.770, 0.821)	129	
5 to <6	0.787	(0.758, 0.816)	113	
6 to <7	0.775	(0.749, 0.801)	97	
7 to <8	0.784	(0.759, 0.809)	171	0.785 (0.763, 0.806)
8 to <9	0.785	(0.759, 0.810)	153	
9 to <10	0.747	(0.712, 0.781)	104	0.747 (0.725, 0.770)[a]
10 to <11	0.746	(0.710, 0.782)	69	
11 to <12	0.755	(0.722, 0.788)	72	
12 to <13	0.742	(0.703, 0.782)	47	
13 to <14	0.786	(0.733, 0.839)	32	
≥14	0.724	(0.682, 0.767)	95	

Footnotes: [a] Indicates a significant difference in EQ-5D in the VMS frequency category compared with the previous category (p<0.05).

Abbreviations: CI: confidence interval; GEE: generalised estimating equation; UK: United Kingdom; VMS: vasomotor symptoms.

B.3.2.3 Intervention technology and comparators

Intervention

The intervention assessed in the base case model was fezolinetant 45 mg once daily. This was in line with the decision problem and the licensed dose for fezolinetant in moderate to severe VMS associated with menopause. The dosing schedule for fezolinetant used in the model is also consistent with the fezolinetant 45 mg arm of DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2.

Comparators

As noted in Section B.3.1, the target population of relevance to the decision problem is menopausal people with moderate to severe vasomotor-predominant menopausal symptoms for whom HRT is deemed unsuitable for medical reasons. For this population, UK clinical experts confirmed that no active treatment represents the sole relevant comparator to fezolinetant, given the lack of effective and safe treatment options available in UK clinical practice. Therefore, in keeping with UK clinical expert opinion, Astellas considers no active treatment to be the main comparator of relevance to the decision problem.

B.3.3 Clinical parameters and variables

The model used clinical data primarily sourced from DAYLIGHT, with supporting evidence from sources, including:

- Pooled SKYLIGHT 1 and SKYLIGHT 2
- Outputs from SEE of natural history of untreated VMS associated with menopause
- Pooled DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2

B.3.3.1 Baseline characteristics

B.3.3.1.1 Age

In the base case, the mean starting age of the modelled cohort was 51 years based on the average age of menopause onset in the UK.[112]

It was deemed appropriate given that UK clinical experts consulted by Astellas indicated that the mean age of women enrolled in DAYLIGHT (54.5 years) was higher than the mean age of people seen in UK clinical practice (51.0 years). This discrepancy arose because DAYLIGHT included not only women experiencing the onset of VMS at enrolment but also those who had been experiencing moderate to severe VMS for a few years before enrolment. Two scenarios were explored using baseline age from DAYLIGHT trial data and pooled SKYLIGHT and DAYLIGHT trial data (Section B.3.11.3).

Table 58: Baseline characteristics of modelled cohort

Parameter	Value	Source
Mean age, years	51.0	Clinical input; NHS inform: Menopause[112]

B.3.3.1.2 Baseline moderate to severe VMS frequency distribution

Upon entering the model, menopausal people are distributed into each moderate to severe VMS frequency 'on-treatment' health state based on the baseline distribution observed in DAYLIGHT (Table 59). A scenario analysis was provided that applies the baseline distribution from pooled DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 trials (Section B.3.11.3).

To have been eligible for inclusion in DAYLIGHT, women needed to have reported a minimum average of 7 moderate to severe VMS per day within 10 days prior to randomisation. However, one person in the study was reported as having a VMS frequency between 2 and 7, due to this time lag between inclusion criteria and baseline measurement, which occurred at Day 1 (i.e., randomisation).

Table 59: Baseline distribution of people used in the model

Source		0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
DAYLIGHT	Number of people	0	1	189	262
	% of people	0.00%	0.22%	41.81%	57.96%

Abbreviations: VMS: vasomotor symptoms.

B.3.3.2 Efficacy

Because DAYLIGHT was strictly conducted in a population deemed unsuitable for HRT, efficacy inputs while 'on treatment' were primarily sourced from this trial. When data were unavailable due to the short trial follow-up period, longer-term efficacy inputs from the HRT-unsuitable subgroup of the pooled SKYLIGHT 1 and 2 were used. A scenario analysis was also conducted which excluded patients who were classified as HRT-averse from the DAYLIGHT trial and the pooled population of patients in SKYLIGHT 1 and SKYLIGHT 2. This analysis was conducted to assess the impact of excluding this large cohort of patients on the clinical efficacy data from the respective trials, and subsequently the impact on the results from this economic analysis. This scenario is explored further in Section B.3.11.3.

B.3.3.2.1 Fezolinetant

In the base case, whilst 'on treatment', the model applies four-weekly (per cycle) transition probability matrices at each cycle up until Week 24, derived using Week 0–24 DAYLIGHT efficacy data for fezolinetant. For Weeks 24–52, four-weekly (per cycle) transition probability matrices are applied to each cycle, derived using pooled Week 24–52 SKYLIGHT 1 and SKYLIGHT 2 efficacy data (HRT-unsuitable subpopulation). From Week 52 onwards, the model applies the average of four-weekly transition probability matrices to each cycle, derived using pooled SKYLIGHT 1 and SKYLIGHT 2 Week 24–52 efficacy data (HRT-unsuitable

subpopulation). A summary of the transition probabilities used in the base case are presented in Appendix N.

For menopausal people who discontinue treatment with fezolinetant at any timepoint in the model (see Section B.3.3.2.4), they were modelled to transition to the off-treatment VMS frequency health states, in which they revert to the natural history distribution of VMS and follow natural history transitions (see Section B.3.3.2.3).

It was deemed appropriate to inform efficacy inputs beyond Week 24 using pooled SKYLIGHT 1 and SKYLIGHT 2 efficacy data (HRT-unsuitable subpopulation), given that the trial period for DAYLIGHT lasted up to Week 24, with no further follow-up efficacy data beyond this timepoint. This approach made use of the best available data over a full year to inform the efficacy inputs for fezolinetant in the model. Treatment waning was not directly considered in the model for fezolinetant, as it is expected that loss of treatment response in menopausal people with moderate to severe VMS is clinically measurable and will therefore lead to treatment discontinuation. As such, discontinuation was considered a suitable proxy for treatment waning. The decision not to model treatment waning was validated by UK clinical experts as part of NICE's early engagement process through the OMA, who indicated that it was reasonable to assume no treatment waning while on fezolinetant due to its mechanism of action. Moreover, 52 weeks of fezolinetant efficacy data from pooled SKYLIGHT 1 and SKYLIGHT 2 trials showed that fezolinetant is effective at both eliciting a response at Week 1 and maintaining that response up to Week 52, with no indication of treatment waning.

In the model, per cycle transition probabilities were estimated using the distribution of menopausal people across the four VMS frequency health states observed in DAYLIGHT or pooled SKYLIGHT 1 and SKYLIGHT 2 trials at four-week intervals. A multinomial model was fitted to the transition counts observed in each four-week cycle of pooled SKYLIGHT 1 and 2 to generate average four-weekly transition probabilities. Given that transition probabilities for Week 24–52 of pooled SKYLIGHT 1 and 2 were applied for the remainder of the model time horizon, this approach provided a more robust estimation of health state transitions than using per cycle transition probabilities by considering multiple cycles collectively. This smoothing effect reduces the impact of random fluctuations within specific cycles and provides a more reliable estimate of average transition probabilities over a longer period, which is particularly useful for application over the model time horizon.

A scenario analysis was also considered focusing on data from the pooled DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 trials, including a combination of baseline patient distribution and age, efficacy, discontinuation rates and utility values (Section B.3.11.3).

B.3.3.2.2 No active treatment

In the base case, for menopausal people on treatment in the no active treatment arm, transitions between VMS frequency health states were modelled based on four-weekly per cycle transition probabilities derived from Week 0–12 DAYLIGHT placebo efficacy data, up until Week 12. At the end of Week 12, the model assumes that all remaining menopausal people in the on-treatment VMS frequency health states transition to the off-treatment VMS frequency health states, in which they revert to their natural history distribution of VMS and follow natural history transitions (Section B.3.3.2.3). A summary of the transition probabilities used in the base case are presented in Appendix N.

The approach of using placebo data up to Week 12 for modelling moderate to severe VMS frequency health state transitions in people receiving no active treatment was conservative for several reasons. Despite receiving pharmacologically inert tablets, women in the placebo arm experienced notable improvements in moderate to severe VMS frequency from baseline through Week 12 in DAYLIGHT, exceeding expectations from sampling variation. While high placebo effects are common in clinical trials in moderate to severe VMS,[113, 114] there are concerns about generalising these results to no active treatment in real-world UK clinical practice, given that the target population are unlikely to be receiving a placebo in clinical practice for ethical reasons, and are also unlikely to be receiving any additional follow-up or GP interactions associated with partaking in a clinical trial. As such, it is likely that the benefits of no active treatment are overestimated in the model. A scenario analysis was considered that explores the modelling of no transitions in the no active treatment arm from Week 0–24 using the baseline distribution, and a scenario analysis was explored where discontinuation was applied at Week 24 as an automatic reversion to 'post discontinuation' efficacy (Section B.3.11.3).

Additionally, the loss of the placebo effect at the end of Week 12 was validated with UK clinical experts, who acknowledged that although a long-term placebo effect has been shown in menopause clinical trials, likely in part due to resource use and healthcare provider interaction as part of the trial, and a placebo effect may be observed in the short-term by menopausal people taking non-pharmacological supplements, the effect is not expected to persist in clinical practice and menopausal people receiving no active treatment would be expected to follow natural history more closely (Section B.3.3.2.3).[7] This opinion was also shared by patient experts, who did not expect the placebo effect to persist for a long time beyond the treatment period of a clinical trial.[12]

B.3.3.2.3 Natural history

As noted in the sections above, in the base case, once menopausal people in the on-treatment health states transition to the off-treatment health states, they are modelled to follow natural history.

Natural history data were not available from the phase 3 clinical trials for fezolinetant, given that neither DAYLIGHT, SKYLIGHT 1 nor SKYLIGHT 2 were designed to directly capture natural history data. Additionally, no relevant studies on the natural history of VMS for the target population of interest to the decision problem were identified in the clinical SLR conducted in line with the NICE reference case (discussed in Section B.2.1 and in Appendix D).

This evidence gap was therefore addressed by a two-stage process:

1. First, six clinical experts from England and Wales (a consultant gynaecologist, a primary care GP, a consultant gynaecology nurse, two pharmacists and a consultant obstetrician/gynaecologist) provided their judgements on the natural history of postmenopausal women who were not currently receiving any treatment for VMS and were deemed HRT-unsuitable.[115] This was conducted via SEE using Structured Expert Elicitation Resources (STEER) guidance (based on the York reference protocol).[116]
2. Following this, the outputs of the SEE study were validated by three UK clinical experts in order to generate natural history inputs for the model base case.[7] While the SEE process is methodologically robust and designed to rigorously gather expert input using structured protocols, further validation was deemed necessary to ensure that the findings were not

overly influenced by the views of a single, relatively small group of experts. The second phase of validation provided an opportunity to confirm that the initial outputs reflected broader clinical consensus and were not limited by the subjective experiences of the original SEE participants.

The resulting SEE natural history estimates were then further validated by a group of four patient experts in order to ensure that the estimates were appropriate from a patient perspective.[12]

SEE estimates

The SEE project recruited six UK clinical experts via pre-specified recruitment criteria, and included general practitioners, gynaecologists, nurses, and pharmacists with a strong interest in women's health who have provided care and advice for women who are going through menopause, or menopause specialists conferred by the British Menopause Society, as well as experience in treatment modalities for VMS. The outputs of the SEE inform a number of the assumptions in the model base case (as reported below) and further details on the methodology and key outputs from the SEE exercises are presented in a report.[115]

During the SEE exercise, UK clinical experts were asked to consider a hypothetical cohort of 1,000 postmenopausal women in the UK who were not currently receiving any treatment for moderate to severe VMS and were not suitable for HRT. They were then asked to estimate the proportion (i.e., the number of women out of 1,000) who experienced different daily VMS frequencies across three timepoints (Year 1, Year 3 and Year 6) so that an overall trajectory may be mapped out. These time points were selected based on the observed median VMS durations reported by Politi *et al.,* (2008)[117] (4 years) and Avis *et al.,* (2015)[49] (7.4 years). This timeframe was considered sufficiently long to capture plausible changes in VMS frequency alongside the natural history of menopause.

The median number of women elicited in each VMS frequency health state at the aforementioned timepoints are presented in Table 60, and the corresponding proportion of women in each state is presented in Table 61 and Figure 27, with the changes in the frequency of moderate-to-severe VMS over time presented in Figure 28.

Table 60: SEE natural history health state distribution estimates[a]

Time	0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9	Total
Year 1	135.67 (SE 114.57)	378.24 (SE 172.35)	137.73 (SE 178.73)	150.87 (SE 160.06)	802.51
Year 3	350.15 (SE 183.04)	384.12 (SE 160.75)	127.22 (SE 153.92)	90.74 (SE 90.29)	952.23
Year 6	399.77 (SE 159.49)	276.90 (SE 90.90)	102.91 (SE 159.82)	51.12 (SE 87.39)	830.70

[a] Values are reported as the median value of clinician estimates, based on a hypothetical cohort of 1,000 women. The SE is calculated as the standard deviation divided by the square root of the sample size.
Abbreviations: SE: standard error; SEE: structured expert elicitation; VMS: vasomotor symptoms.

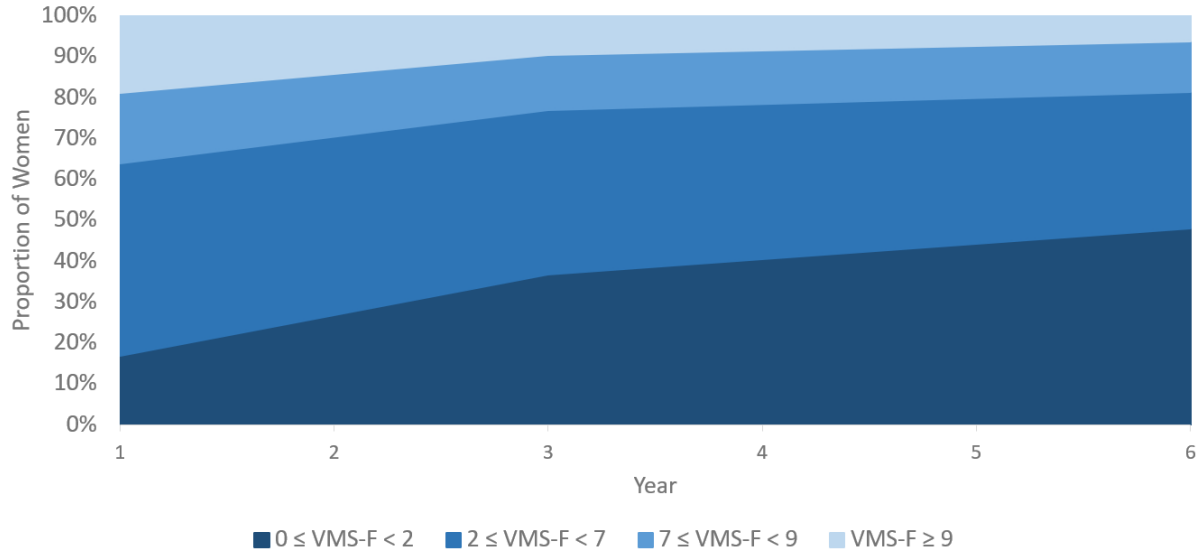
Table 61: SEE natural history health state distribution proportions[a]

Time	0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
Year 1	16.91%	47.13%	17.16%	18.80%

Year 3	36.77%	40.34%	13.36%	9.53%
Year 6	48.12%	33.33%	12.39%	6.15%

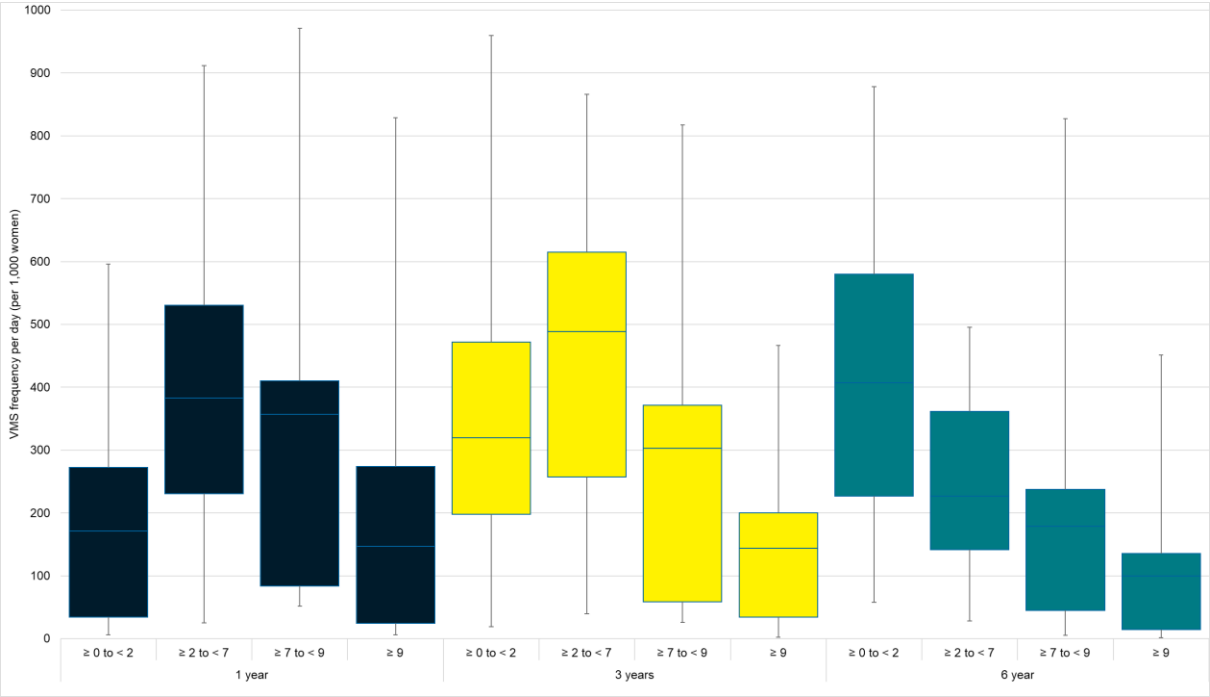
[a] In order to calculate proportions in each health state and at each timepoint, the total in Table 60 was used as the denominator.
Abbreviations: SEE: structured expert elicitation; VMS: vasomotor symptoms.

Figure 27: SEE natural history estimates



Abbreviations: SEE: structured expert elicitation; VMS: vasomotor symptoms.

Figure 28: Pooled proportion of menopausal people experiencing different frequencies of VMS over time and the associated subjective uncertainty



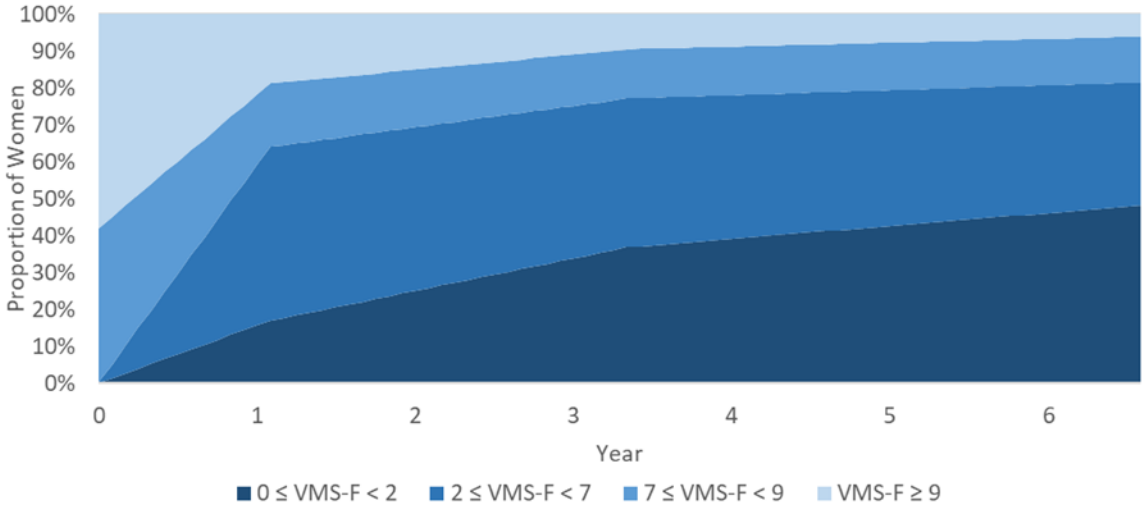
Interquartile range (coloured boxes) and 95% credible interval (error bars).
Abbreviations: VMS: vasomotor symptoms.

As observed in Table 60 and Figure 28, there was considerable uncertainty in the SEE estimates on the natural progression of VMS, highlighting the inherent challenge in estimating the course of

symptoms in postmenopausal women who are receiving no active treatment and are deemed HRT-unsuitable. There may be several reasons for this. Firstly, the experts noted that the trajectory of VMS is highly individual to each menopausal person, which is reflected in the very large credible intervals in the estimates as shown in Figure 28.[115] Additionally, whilst the minimum number of experts (i.e., five experts) required to conduct the SEE was achieved in line with STEER guidance, the relatively low number of experts (six) taking part, compared with the number invited (eight) could have resulted in the discrepancy.[115]

Further discussion with clinical experts suggested that the high uncertainty captured is representative of the highly variable experience of each person experiencing VMS associated with menopause and reflects the counterfactual nature of trying to estimate the VMS frequency in a cohort that would rarely exist in reality – clinicians would not leave people experiencing moderate to severe VMS untreated in this manner. Furthermore, when the three elicited timepoints (1, 3 and 6 years) were applied to the DAYLIGHT baseline distribution, a very rapid rate of VMS frequency decline was observed (Figure 29). As such, on the recommendation of UK clinical expert and patient expert opinion, further input was sought to validate the SEE natural history estimates informing the model base case.

Figure 29: SEE natural history distribution with DAYLIGHT baseline distribution



Year 0 in this figure represented the baseline distribution as observed in the DAYLIGHT trial
Abbreviations: SEE: structured expert elicitation; VMS: vasomotor symptoms.

Clinical validation of SEE estimates

When the original SEE estimates were used in the model base case, as presented in Figure 29, the three UK clinical experts (including one expert involved in the SEE exercise) found the timeline, especially the one-year timepoint, unrealistically fast for the resolution of VMS based on typical clinical practice.[7] Patient experts consulted independently also agreed with this observation, with one expert highlighting that it often takes six months for a patient to realise that VMS is impacting their lives before they seek medical advice, and so it is unlikely that symptoms would resolve within one year.[12]

As such, the following alternative approaches were proposed by the experts to adjust the SEE results to better reflect UK clinical practice:

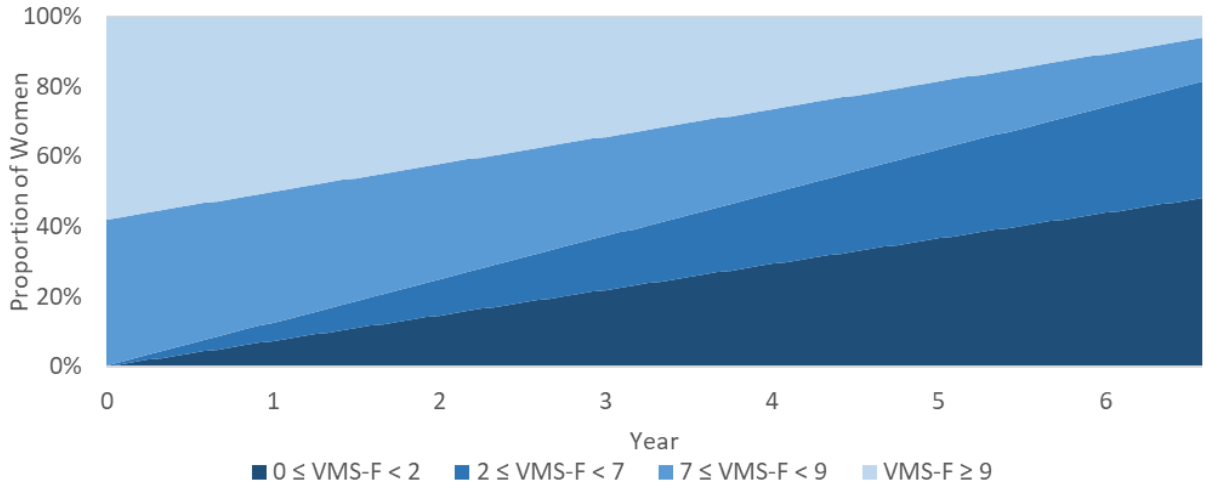
- Only use Year 6 proportions from the SEE

- Using the SEE Year 1 estimates at Year 2
- Only use Year 3 and Year 6 estimates

The option to only use Year 6 proportions from the SEE was chosen as the base case, as the clinical experts agreed that while VMS do gradually improve over time, a significant reduction in symptoms is unlikely to occur within the first year. Instead, symptoms typically start to improve after Year 2, with more noticeable improvements by Year 3. Given this, the experts felt that the Year 6 estimates provided a more realistic reflection of the natural progression of VMS.[7] Patient experts consulted independently of the clinical experts also unanimously chose the option to use only the Year 6 estimates, while acknowledging the challenges in selecting a single distribution to represent all people experiencing moderate to severe VMS. Likewise, the clinical experts acknowledged the challenges in deciding their preferred option; as such, to assess the impact of the other alternative natural history estimates (use Year 1 estimates at Year 2 and only use Year 3 and Year 6 estimates), scenario analyses were conducted (Section B.3.11.3). Details of the estimates used to inform these scenario analyses are presented in Appendix O.

The base case model calculates the distributions at each four-week cycle from Year 0 up to Year 6, assuming a linear change from baseline to Year 6 in each VMS frequency health state (Figure 30). The natural history distribution for Year 6 onwards in each moderate to severe VMS frequency health state was calculated by applying natural waning rate based on the gradient between Year 3 and Year 6 of the SEE distribution. As an alternative approach, a scenario analysis is also provided that resets the modelled cohort to the baseline distribution from the DAYLIGHT trial post-discontinuation rather than following the natural history distribution (Section B.3.11.3).

Figure 30: SEE Year 6 estimates (base case)



Abbreviations: SEE: structure expert elicitation; VMS: vasomotor symptoms.

B.3.3.2.4 Treatment discontinuation

In the base case, upon treatment discontinuation, menopausal people were modelled to revert to the natural history distribution and follow natural history until VMS cessation, death, or the end of the model time horizon.

All-cause discontinuation rates were applied to menopausal people in the on-treatment VMS frequency health states, whether they were receiving fezolinetant or no active treatment, to reflect the proportion of people returning to their natural history distribution in each model cycle.

Discontinuation rates were assumed to be constant over time in the model and are presented in Table 62.

For fezolinetant, the per cycle probability of discontinuation was derived using DAYLIGHT Week 0–24 trial data and was applied in each model cycle up until Week 24. For Week 24 onwards, discontinuation rates were sourced from pooled SKYLIGHT 1 and SKYLIGHT 2 Week 24–52 trial data (HRT-unsuitable subpopulation).

For no active treatment, per cycle discontinuation rates were sourced from DAYLIGHT Week 0–24 placebo trial data and were applied in each model cycle up until Week 12, after which all remaining menopausal people in the on-treatment health state are assumed to revert to the natural history distribution and follow natural history transitions from then onwards until VMS cessation, death, or the end of the model.

Table 62: All-cause treatment discontinuation rates (base case)

Treatment	Timepoint	Per-cycle probability of discontinuation	Source
Fezolinetant 45 mg	Week 0 to 24	2.43%	DAYLIGHT
	Week 24+	[redacted]%	Pooled SKYLIGHT 1 and 2 (HRT-unsuitable subpopulation)
No active treatment	Week 0 to 24	4.17%	DAYLIGHT

B.3.3.2.5 VMS Cessation

The VMS cessation rate was assumed to be constant over the model time horizon, with transitions to the VMS cessation health state in each model cycle based on a per cycle probability of VMS cessation (Table 13).

In the base case, per cycle probability of VMS cessation was derived using the median duration of VMS reported by Avis *et al*., (2015).[98] This value was deemed appropriate to inform the model based on UK clinical opinion which considered the estimate to be aligned with their clinical experience.[7]

Table 63: Approach to VMS cessation state transitions in the model

Source	Median Duration of Symptoms (years)	Duration of Symptoms (SE)	Probability of Symptom Cessation per Cycle
Avis, 2015[98]	7.4	0.74	0.007

Abbreviations: SE: standard error; VMS: vasomotor symptoms.

B.3.3.2.6 Safety

Safety data for fezolinetant are sourced from four phase 3 trials: DAYLIGHT, SKYLIGHT 1, SKYLIGHT 2 and SKYLIGHT 4. Safety information for DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 are presented in Section B.2.10, and SKYLIGHT 4 safety data are presented in Appendix F.

Across all four trials, the proportion of women that experienced at least one TEAE was comparable between the fezolinetant 45 mg group and the placebo group. The frequency of serious TEAEs was also low and similar across both groups in each of the trials. Moreover, no deaths were reported in DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 4, and only one death was reported in SKYLIGHT 2; however, the death was not considered to be related to fezolinetant therapy or to VMS.

Consequently, due to the low incidence of serious adverse events (SAEs; <5%) observed in each of the trials, no adverse events were considered in the model, as it was assumed these would have minimal impact on the model results.

B.3.3.2.7 Mortality

Based on the observed survival of people in DAYLIGHT, SKYLIGHT 1, SKYLIGHT 2 and SKYLIGHT 4, and the published literature,[118] it was concluded that moderate to severe VMS did not have any meaningful impact on mortality risk. As such, the modelled cohort were therefore assumed to have the same mortality rate as for the general female population in the UK.

Age-specific mortality rates for females were derived from the National UK life tables for 2019 (published by the Office for National Statistics [ONS]).[104] Mortality rates from 2019 were used in the base case as these data are the latest UK figures available that were not impacted by COVID-19 mortality rates.

B.3.4 Measurement and valuation of health effects

B.3.4.1 Health-related quality-of-life data from clinical trials

Arithmetic mean utility values for menopausal people with moderate to severe VMS stratified by VMS frequency were derived using all available EQ-5D-5L data collected directly from women from Weeks 0–24 of DAYLIGHT. This is consistent with the NICE reference case.[119]

B.3.4.2 Mapping

In line with the NICE reference case, EQ-5D-5L data collected in DAYLIGHT were mapped to EQ-5D-3L utility values using the mapping algorithm published by Hernández Alava *et al.* (2023).[120]

B.3.4.3 Health-related quality-of-life studies

As described in Section B.3.1, an economic SLR was conducted to identify relevant literature published on previous economic evaluations, utility values and key model inputs to inform the cost effectiveness model for fezolinetant. Utility values from the identified study were not used in the cost-effectiveness model, as EQ-5D data were directly available from the phase 3 DAYLIGHT trial in the population of relevance to the decision problem.

B.3.4.4 Adverse reactions

As noted in Section B.3.3.2.6., due to the low incidence of TEAEs and SAEs observed for fezolinetant and placebo during the DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 trials, it is not expected that a significant detriment in HRQoL would be associated with these events. As such, disutilities associated with adverse events were not included in the model.

B.3.4.5 Health-related quality-of-life data used in the cost-effectiveness analysis

The utility values for each moderate to severe VMS frequency state were estimated using a GEE linear mixed effects regression model, which accounted for the correlation within women's repeated assessments over time. In the regression model, the dependent variable was the utility index score derived from EQ-5D-5L data mapped to EQ-5D-3L, and the independent variable was the VMS frequency health states. Due to the inclusion criteria of the DAYLIGHT trial, only people with a mean daily frequency of VMS of at least 7 episodes per day were included, therefore utility values were not estimable for the 0 ≤ VMS Frequency < 2 and 2 ≤ VMS Frequency < 7 health states at baseline. Hence, post-baseline measures were used as an approximation of the utility values for the 0 ≤ VMS Frequency < 2 and 2 ≤ VMS Frequency < 7 and health states.

Utility values for VMS cessation were calculated from the average EQ-5D-5L data whenever people reported a VMS frequency of zero in DAYLIGHT. Subsequently, these values were mapped to EQ-5D-3L utilities using the algorithm provided by Hernandez Alava *et al*., (2023).[120] No further adjustment was made in the analysis to account for repeated measurements.

While the EQ-5D measure is widely used for assessing overall HRQoL, studies have shown that generic measures such as this are insufficiently sensitive to the improvements brought by pharmacological treatments in menopausal women.[105, 106] In a study involving 3,721 women aged 50–69, HRT statistically significantly alleviated VMS, improved sexual functioning, and reduced sleep problems.[105] However, these benefits did not translate into statistically significant differences in overall HRQoL as measured by EQ-5D.[105] The study demonstrated that while HRT reduced hot flushes, night sweats, aching joints, insomnia, and vaginal dryness, the EQ-5D/VAS scores did not reflect these improvements.[105] This suggests that EQ-5D may not capture the broad range of health impacts associated with menopause and the treatment of these symptoms. Given the extensive impact of HRT on various aspects of menopause beyond VMS alone, it is unsurprising that fezolinetant also struggles to show substantial improvement in EQ-5D/VAS scores.

Given these challenges in capturing HRQoL benefits in menopause, in the base case, the DAYLIGHT utility values were adjusted based on feedback from UK clinical experts (Table 64).[7] Clinician feedback suggested that the difference between VMS cessation and 0 ≤ VMS Frequency < 2 would be greater in UK clinical practice than seen in the DAYLIGHT trial. Experts estimated that the utility of the 0 ≤ VMS Frequency < 2 health state would be 0.810 and suggested that the utilities for the 7 ≤ VMS Frequency < 9 and VMS Frequency ≥ 9 health states would be lower; these were therefore adjusted assuming a 5% reduction in utility compared with the utility reported in the DAYLIGHT trial for the corresponding VMS frequency health state. During validation by patient experts, the experts raised that the adjusted values were conservative and in reality they would expect lower utility values for the higher frequency states, and suggested that these adjusted values could be lowered further.[12]

Table 64: Health-state utilities

Health state	UK clinician adjusted DAYLIGHT data (base case)		DAYLIGHT (scenario)		UK clinician adjusted DAYLIGHT data excluding HRT-averse (scenario)	
	Mean	SE	Mean	SE	Mean	SE
Cessation of VMS	0.843	0.012	0.843	0.012	0.849	0.016
0 ≤ VMS Frequency < 2	**0.810**	0.010	0.833	0.010	0.828	0.014
2 ≤ VMS Frequency < 7	0.793	0.010	0.793	0.010	0.776	0.013
7 ≤ VMS Frequency < 9	**0.746**	0.011	0.785	0.011	**0.733**	0.016
VMS Frequency ≥ 9	**0.710**	0.012	0.747	0.012	**0.700**	0.015

Estimates in bold denote where estimates have been adjusted based on UK clinician estimates compared with the mean values in DAYLIGHT and DAYLIGHT (excluding HRT-averse population). Clinicians suggested the utilities for the 7 ≤ VMS-F < 9 and VMS-F ≥ 9 health states would be lower; these were therefore adjusted assuming a 5% reduction in utility compared to the utility reported in the DAYLIGHT trial (excluding HRT-averse population) for the corresponding VMS frequency health state. Utilities apply to both treatment arms of the model.
Abbreviations: SE: standard error; UK: United Kingdom; VMS: vasomotor symptoms.

Alternative sources of utility values from DAYLIGHT data and SKYLIGHT and DAYLIGHT pooled analysis data were selected as scenario analyses (Section B.3.11.3).

Utility age adjustment

In line with the NICE reference case, the model applies an age-dependent annual adjustment factor to account for the expected decline in health utility with increasing age, using UK data from Hernandez-Alava *et al*., (2022).[121]

B.3.5 Cost and healthcare resource use identification, measurement and valuation

An SLR was conducted to identify any relevant economic evaluations for VMS associated with menopause. The original SLR was conducted in June 2023 and was updated in April 2024. Full details of the SLR search strategy, study selection process and results are reported in Appendix I.

In total, 24 publications reporting on 23 unique studies were included in the SLR. However, due to the health state definitions used in the model and issues with the generalisability of the included studies to UK clinical practice, none of these studies were considered suitable to inform the model base case. As such, UK clinical expert opinion was sought by Astellas to inform healthcare resource use rates by health state in the model.[7]

The following cost categories were considered in the base case model:

- Treatment acquisition costs
- Costs associated with healthcare resource use.

The cost perspective adopted was that of the NHS and PSS. As such, the base-case economic analysis included only costs that would be incurred by the NHS and PSS.

B.3.5.1 Intervention and comparators' costs and resource use

Drug acquisition costs included in the model are summarised in Table 65.

The dosing schedule for fezolinetant 45 mg was in line with the licensed posology for fezolinetant in moderate to severe VMS. The dosing regimen used was also consistent with the fezolinetant 45 mg arm of DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2.

No drug acquisition costs were considered for no active treatment, given that people are receiving no active treatment in this arm of the model.

No administration costs were considered for fezolinetant as it is self-administered orally by patients.

Table 65: Drug cost inputs (base case)

Treatment	Route of administration	Dosing schedule	% receiving treatment	Cost per cycle (£)	Source
Fezolinetant	Oral	45 mg QD	100%	£44.80	BNF[14]
No active treatment	N/A	N/A	100%	£0.00	Assumption

Abbreviations: BNF: British National Formulary; QD: once daily; N/A: not applicable.

B.3.5.2 Health-state unit costs and resource use

A summary of the estimated resource use rates and associated unit costs considered in the base case model are presented in Table 66 and Table 67 respectively.

Healthcare resource use (HCRU) inputs in the base case were derived using estimates from UK clinical experts, which included the use of emergency room (ER) visits and specialist visits as seen in UK clinical practice. The model assumed resource use rates to be health-state dependent and independent of treatment received. These estimates were validated by patient experts, who were in broad agreement with the estimates presented, but suggested adding at least one specialist follow-up visit within a year for patients in the higher VMS frequency health states.[12]

Alternative healthcare resource use rates sourced from Whiteley *et al.,* (2013) were included as a scenario analysis in the model (see Section B.3.11.3).[27] This source was based on a US study due to limited UK studies reporting on resource use rates for VMS.

The unit cost for procedures and monitoring related to the management of VMS were obtained using the Personal Social Services Research Unit (PSSRU) Costs of Health and Social Care 2022 and the NHS Schedule of Costs 2021/22.[96, 122]

Table 66: Healthcare resource use by health state (per year, base case)

Res	Cessation of VMS[a]	0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
Physician visits (first visit)	1.00	1.00	1.00	1.00	1.00
Physician visits (follow-up visit)	0.00	0.00	1.00	2.00	2.00

ER visits	0.00	0.00	0.00	1.00	1.00
Specialist visit (first visit)	0.00	0.00	0.00	0.00	1.00
Specialist visit (follow-up visit)	0.00	0.00	0.00	0.00	0.00

Abbreviations: VMS: vasomotor symptoms.
Source: UK clinician estimates.[7]

Table 67: Healthcare resource unit costs (base case)

Resource use	Unit Cost (£)	Source
Physician visits (first visit)	£55.00	PSSRU, 2023. Cost per GP surgery consultation (10 minutes)
Physician visits (follow-up visit)	£55.00	
ER visit	£158.47	National schedule of NHS costs 2021/22. Code: VB09Z
Specialist visit (first visit)	£215.19	National schedule of NHS costs 2021/22: WF01A (Consultant led, non-admitted face-to-face attendance)
Specialist visit (follow-up visit)	£181.26	National schedule of NHS costs 2021/22: WF01B (Consultant led, non-admitted face-to-face attendance)

Abbreviations: ER: emergency room; NHS: national health service; PSSRU: personal and social services research unit.

Costs of healthcare resource use were applied on a per cycle basis, converted from annual costs, using the respective health state inputs; a summary of the base case resource costs per health state are presented in Table 68.

Table 68: Resource use costs per health state (base case)

Health state	Resource utilisation cost (yearly)	Cost per cycle (£)
Cessation of VMS	£55.00	£4.22
0 ≤ VMS Frequency < 2	£55.00	£4.22
2 ≤ VMS Frequency < 7	£110.00	£8.43
7 ≤ VMS Frequency < 9	£323.47	£24.80
VMS Frequency ≥ 9	£538.66	£41.29

Costs apply to both treatment arms.
Abbreviations: VMS: vasomotor symptoms.
Source: UK clinician estimates.[7]

B.3.5.3 Adverse reaction unit costs and resource use

As noted in Section B.3.3.2.6., due to the low incidence of TEAEs and SAEs observed for fezolinetant and placebo during the DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 trials, costs and resource use associated with adverse reactions were not included in the model.

B.3.5.4 Miscellaneous unit costs and resource use

No additional miscellaneous unit costs and resource use were included in the model.

B.3.6 Severity

This technology does not meet the criteria for a severity weight.

B.3.7 Uncertainty

Neither the nature of moderate to severe VMS nor fezolinetant impacts the ability to generate high-quality evidence. However, there is inherent uncertainty in the natural history distribution estimates that inform post-discontinuation transitions in the model, due to heterogeneity in individual menopausal people's experiences in the trajectory of their VMS and insufficient existing data. In order to address this, clinical insight was sought in order to generate SEE natural history estimates and validated to reflect what is seen in UK clinical practice. This uncertainty was explored through a scenario analysis where no natural history was used post-discontinuation to remove this uncertainty.

B.3.8 Managed access proposal

It is anticipated that the appraisal will result in routine commissioning and therefore no managed access proposal is required.

B.3.9 Summary of base-case analysis inputs and assumptions

B.3.9.1 Summary of base-case analysis inputs

A summary of inputs for the base case analysis is presented in Table 69.

Table 69: Summary of variables applied in the economic model (base case)

Variable	Value	Measurement of uncertainty and distribution: SE (distribution)	Reference to section in submission
Model characteristics			
Time horizon	10 years	N/A	Section B.3.2.2
Cycle length	4 weeks	N/A	Section B.3.2.2
Half-cycle correction	Yes	N/A	Section B.3.2.2
Discount rate (costs)	3.5%	N/A	Section B.3.2.2
Discount rate (benefits)	3.5%	N/A	Section B.3.2.2
Perspective on cost	NHS and PSS	N/A	Section B.3.2.2
Perspective on outcomes	All relevant health effects	N/A	Section B.3.2.2
Age-adjusted utility	Yes	N/A	Section B.3.4.5

Mortality risk	General population mortality	N/A	Section B.3.3.2.7
Patient Characteristics			
Mean starting age, years	51.0	5.10 (Normal)	Section B.3.3.1.1
Baseline VMS frequency distribution (DAYLIGHT)			
0 ≤ VMS Frequency < 2	0.00%	The Dirichlet (Gamma) distribution uses total patient numbers to parameterise a gamma distribution which is used to sample patient numbers which are then converted back to health state distributions	Section B.3.3.1.2
2 ≤ VMS Frequency < 7	0.22%		
7 ≤ VMS Frequency < 9	41.81%		
VMS Frequency ≥ 9	57.96%		
Fezolinetant transition probabilities			
Weeks 0–12	Per cycle transition probabilities from the DAYLIGHT trial (Weeks 0 to 12)	The Dirichlet (Gamma) distribution uses total patient numbers to parameterise a gamma distribution which is used to sample patient numbers which are converted back to transition probabilities	Section B.3.3.2.1
Weeks 12–24	Per cycle transition probabilities from the DAYLIGHT trial (Weeks 12 to 24)		
Weeks 24–52	Per cycle transition probabilities from the pooled SKYLIGHT 1 and 2 trials (Weeks 24 to 52)		
Week 52+	Average transition probabilities from the pooled SKYLIGHT 1 and 2 trials (Weeks 24 to 52)	The variations for 'average' transition probability inputs have been directly sourced from the iterative simulation (CODA) parameters that were used in their generation	
No active treatment transition probabilities			
Weeks 0 – 12	Per cycle transition probabilities from the DAYLIGHT trial (Weeks 0 to 12)	The Dirichlet (Gamma) distribution uses total patient numbers to parameterise a gamma distribution which is used to sample patient numbers which are converted back to transition probabilities	Section B.3.3.2.2
Weeks 12 – 24	Natural history	The Dirichlet (Gamma) distribution uses total patient numbers to parameterise a gamma distribution which is used to sample patient numbers which are then	Section B.3.3.2.3
Weeks 24 – 52			
Week 52+			

		converted back to health state distributions	
SEE natural history health state distribution estimates			
Year 1 - 0 ≤ VMS Frequency < 2	16.91%	The Dirichlet (Gamma) distribution uses total patient numbers to parameterise a gamma distribution which is used to sample patient numbers which are then converted back to health state distributions	Section B.3.3.2.3
Year 1 - 2 ≤ VMS Frequency < 7	47.13%		
Year 1 - 7 ≤ VMS Frequency < 9	17.16%		
Year 1 - VMS Frequency ≥ 9	18.80%		
Year 3 - 0 ≤ VMS Frequency < 2	36.77%		
Year 3 - 2 ≤ VMS Frequency < 7	40.34%		
Year 3 - 7 ≤ VMS Frequency < 9	13.36%		
Year 3 - VMS Frequency ≥ 9	9.53%		
Year 6 - 0 ≤ VMS Frequency < 2	48.12%		
Year 6 - 2 ≤ VMS Frequency < 7	33.33%		
Year 6 - 7 ≤ VMS Frequency < 9	12.39%		
Year 6 - VMS Frequency ≥ 9	6.15%		
All-cause per-cycle treatment discontinuation rates			
Fezolinetant 45 mg - Week 0 to 24	2.43%	0.06% (Beta)	Section B.3.3.2.4
Fezolinetant 45 mg - Week 24 onwards	[redacted]	[redacted]	
No active treatment - Week 0 to 24	4.17%	0.07% (Beta)	
VMS cessation state transitions			
Median duration of symptoms (years)	7.4	0.74 (Normal)	Section B.3.3.2.5
Health state utilities			
Cessation of VMS	0.843	0.012 (Beta)	Section B.3.4.5
0 ≤ VMS Frequency < 2	0.810	0.010 (Beta)	
2 ≤ VMS Frequency < 7	0.793	0.010 (Beta)	
7 ≤ VMS Frequency < 9	0.746	0.011 (Beta)	
VMS Frequency ≥ 9	0.710	0.012 (Beta)	
Treatment acquisition cost inputs			
Fezolinetant – Cost per cycle	£44.80	N/A	Section B.3.5.1
Fezolinetant – % receiving treatment	100%	N/A	

No active treatment – Cost per cycle	£0.00	N/A	
No active treatment – % receiving treatment	100%	N/A	
Resource costs per health state (per cycle)			
Fezolinetant – Cessation of VMS	£4.22	£0.42 (Gamma)	Section B.3.5.2
Fezolinetant – 0 ≤ VMS Frequency < 2	£4.22	£ 0.42 (Gamma)	
Fezolinetant – 2 ≤ VMS Frequency < 7	£8.43	£0.84 (Gamma)	
Fezolinetant – 7 ≤ VMS Frequency < 9	£24.80	£2.48 (Gamma)	
Fezolinetant – VMS Frequency ≥ 9	£41.29	£4.13 (Gamma)	
No active treatment – Cessation of VMS	£4.22	£0.42 (Gamma)	
No active treatment – 0 ≤ VMS Frequency < 2	£4.22	£0.42 (Gamma)	
No active treatment – 2 ≤ VMS Frequency < 7	£8.43	£0.84 (Gamma)	
No active treatment – 7 ≤ VMS Frequency < 9	£24.80	£2.48 (Gamma)	
No active treatment – VMS Frequency ≥ 9	£41.29	£4.13 (Gamma)	

Abbreviations: N/A: not applicable; SE: standard error; SEE: structured expert elicitation; VMS: vasomotor symptoms.

B.3.9.2 Assumptions

A list of the assumptions made in the base case analysis and their justifications are provided in Table 70.

Table 70: Model assumptions (base case)

Parameters	Description of base case assumption(s)	Justification
Model structure	Markov cohort model	The use of a granular VMS frequency health state model was chosen to best capture the health effects and costs of fezolinetant versus comparators in moderate to severe VMS. This model type incorporates probabilities of transitioning between health states, offering a detailed and realistic projection of long-term costs and health outcomes. By capturing these dynamic changes, a Markov cohort model enhances the robustness and reliability of the cost-effectiveness analysis for fezolinetant. The model structure and the use of VMS frequency to define health states were validated by UK clinical experts, who deemed the approach appropriate for evaluating the cost-effectiveness of fezolinetant.[7] The experts emphasised that moderate to severe VMS frequency was the most objective measure of the health effects of treatment and confirmed the method used to define these health states was appropriate.[7] The model structure based on VMS frequency health states was further deemed appropriate by four patient experts.[12]
Time horizon	10 years	A 10-year time horizon was considered appropriate, given that there is no differential mortality effect between technologies and the differences in costs and clinical outcomes relate to a relatively short period. The median treatment duration for fezolinetant is modelled to be approximately [redacted] months ([redacted] years). As such, in line with the NICE reference case, a time horizon shorter than a lifetime was justified. This approach was further validated by UK clinical experts who indicated that a 10-year time horizon was appropriate, given that they expect the vast majority of people to have discontinued treatment by the 10-year timepoint.[7] UK clinical experts further confirmed that the modelled treatment duration was in line with the expected treatment duration for fezolinetant in UK clinical practice.[7]

Discount rate	Costs and health effects will be discounted at an annual rate of 3.5%	A discount rate of 3.5% was used in alignment with the NICE reference case.
Cycle length	4-week cycles	A cycle length of four weeks was used in the model and was in line with the clinical assessment timepoints observed in the DAYLIGHT and SKYLIGHT 1 & 2 trials.
Efficacy	Four-weekly transition probability matrices are applied to the model. For Weeks 0–24, these are derived from DAYLIGHT Week 0–24 per cycle efficacy data for fezolinetant. For Weeks 24–52, they are derived from pooled SKYLIGHT 1 and SKYLIGHT 2 Week 24–52 per cycle data (HRT-unsuitable subpopulation). From Week 52 onwards, the matrices use the pooled SKYLIGHT 1 and SKYLIGHT 2 Week 24–52 average data. Menopausal people who discontinue fezolinetant transition to the off-treatment health state, in which they revert to the natural history distribution of moderate to severe VMS and follow natural history transitions.	Reflects available data from the DAYLIGHT and pooled SKYLIGHT 1 and SKYLIGHT 2 trials. It was deemed appropriate to inform efficacy inputs beyond Week 24 using pooled SKYLIGHT 1 and SKYLIGHT 2 (HRT-unsuitable) efficacy data, given that the trial period for DAYLIGHT lasted up to Week 24, with no further follow-up efficacy data beyond this timepoint. This approach made use of the best available data to inform the efficacy inputs for fezolinetant in the model. Menopausal people who discontinue fezolinetant revert to the natural history distribution in line with clinical expert opinion,[7] as the natural history distribution is more representative of the natural trajectory of VMS frequency post-discontinuation. It is also worth noting that reversion to the DAYLIGHT baseline distribution was not considered appropriate, as this distribution is skewed by the inclusion criterion of a minimum average of 7 moderate to severe VMS per day mandated in the trial by the FDA.[109]
	For no active treatment, moderate to severe VMS frequency health state transitions for menopausal people on treatment up to Week 12 were applied using per cycle transition probabilities derived using Week 0–12 placebo data from DAYLIGHT. From Week 12 onwards these patients modelled to revert to the natural history distribution of moderate to severe VMS and follow natural history transitions.	The approach of using placebo data up to Week 12 for modelling moderate to severe VMS frequency health state transitions in menopausal people receiving no active treatment was conservative for several reasons. Despite receiving pharmacologically inert tablets, people in the placebo arm experienced notable improvements in moderate to severe VMS frequency from baseline through Week 12 in DAYLIGHT, exceeding expectations from sampling variation. UK clinical experts did not expect a long-term placebo effect to persist in clinical practice. As such, in line with this feedback, the model assumed menopausal people revert to their natural history distribution and follow natural history transitions from Week 12 onwards. This was also in alignment with the opinion of the four patient experts.[12]

Treatment discontinuation	For Weeks 0–24, treatment discontinuation rates were modelled using a constant per cycle treatment discontinuation (all-cause) probability derived from Week 0–24 DAYLIGHT data. Beyond Week 24, discontinuation was applied using a constant per cycle treatment discontinuation based on pooled Week 24–52 data from SKYLIGHT 1 and SKYLIGHT 2 (HRT-unsuitable subpopulation). For menopausal people in the no active treatment arm, a constant per cycle treatment discontinuation (any reason) is applied from Week 0–12 DAYLIGHT data, after which they revert to their SEE natural history distribution and no discontinuation rates are applied.	Loss of treatment response in people with moderate to severe VMS is clinically measurable and will therefore lead to treatment discontinuation. As such, the discontinuation rate accounts for loss of response, as well as discontinuation for other reasons. A per cycle treatment discontinuation was applied in the no active treatment arm in order to capture the loss of placebo effect in the model, given that a conservative assumption was made whereby placebo data from DAYLIGHT was considered suitable to be applied to the no active treatment arm, despite the fact that in clinical practice menopausal people would not be receiving any pharmacologically inert tablets or any additional follow-up associated with partaking in clinical trials.
Baseline moderate to severe VMS frequency distributions	The baseline distribution of moderate to severe VMS frequency was based on observed data in the DAYLIGHT trial.	Reflects the inclusion criteria of the trial.
VMS cessation	The model assumed that there will be a natural resolving and eventual cessation of VMS. People from any of the moderate to severe VMS frequency defined states could transition to the VMS cessation state. Upon entry, menopausal people would remain in this state until they transitioned to the death state. Menopausal people in this health state have a separate, higher utility applied to them compared with those in the VMS frequency health states and do not incur any VMS treatment costs. The cessation of VMS state was based on the median duration of symptoms reported in Avis *et al.*, (2015)[98], which was used to estimate a probability of symptom cessation per cycle using an exponential model. The VMS cessation rate was assumed to be constant over the time horizon of the model.	Reflects the natural progression of moderate to severe VMS. The source used to derive per cycle VMS cessation rates was aligned with UK clinical expert opinion.[7]
Health states utilities	Utilities values were health-state dependent and therefore reflective of the burden of moderate to severe VMS experience. However, they were not treatment specific. This assumes that any	The EQ-5D measure, while widely used for assessing HRQoL, is not sensitive enough to detect the benefits of pharmacological treatments in menopausal women.[105, 106] In a study of 3,721 women

	difference in utility in menopausal people is driven solely due to the frequency of their VMS symptoms, and not directly by which treatment is received (it is assumed all treatment effect on QoL is captured by the treatment effect observed on the health state transitions). These utility values were then adjusted following feedback from UK clinicians with expertise in menopause.	aged 50–69, HRT significantly alleviated symptoms such as hot flushes, night sweats, and insomnia, and improved sexual functioning, but these improvements were insufficiently reflected in EQ-5D scores.[105] This suggests that condition-specific instruments are needed to capture the range of health impacts associated with menopause treatments. Consequently, the DAYLIGHT utility values were adjusted based on feedback from UK clinical experts to better reflect HRQoL benefits in menopause. These clinician-adjusted estimates were deemed appropriate by four patient experts.[12]
Adverse events	Adverse events were not considered in the model.	As noted in Section B.3.3.2.6, across all four trials, the proportion of people that experienced at least one TEAE was comparable between the fezolinetant 45 mg group and the placebo group. The frequency of serious TEAEs was also low and similar across both groups in each of the trials. Moreover, no deaths were reported in DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 4, and only one death was reported in SKYLIGHT 2; however, the death was not considered to be related to fezolinetant therapy or to VMS. Consequently, due to the low incidence of SAEs; <5% observed in each of the trials, no adverse events were considered in the model, as their impact on the ICER was expected to be minimal over the model time horizon.
Mortality risk	Patients were assumed to be at risk of death throughout the model time horizon, irrespective of health state.	Based on safety results from DAYLIGHT, SKYLIGHT 1, SKYLIGHT 2 and SKYLIGHT 4 clinical trials and findings from the published literature,[118] it was concluded that VMS associated with menopause did not have any meaningful impact on mortality risk and therefore age-specific mortality rates for females is applied to all health states over the model time horizon.
Costs	Treatment administration costs were not considered in the model. No treatment acquisition costs were considered for the no active treatment comparator. Costs associated with healthcare resource use were assumed to be dependent on health state and were not treatment specific.	As menopausal people in the no active treatment were not expected to receive any active pharmacological treatments for moderate to severe VMS, it was assumed that no active treatment administration or acquisition costs would be incurred. Fezolinetant can be self-administrated orally by patients, which is assumed to be associated with no administration costs. Healthcare resource use rates were aligned with UK clinical expert and patient expert opinion.[7]

	Patients in the same health state are assumed to have a consistent level of health on average. Therefore, patients in the same health state are also assumed to have similar healthcare resource utilisation on average.	

Abbreviations: EQ-5D: EuroQoL Quality of Life Questionnaire, five dimensions; FDA: US Food and Drug Administration; HRQoL: health-related quality of life; HRT: hormone replacement therapy; ICER: incremental cost-effectiveness ratio; NICE: National Institute for Health and Care Excellence; SAE: serious adverse event; TEAE: treatment emergent adverse event; UK: United Kingdom; US: United States; VMS: vasomotor symptoms.

B.3.10 Base-case results

B.3.10.1 Base-case incremental cost-effectiveness analysis results

A summary of results in the probabilistic base-case analysis and incremental net health benefits (INHB) are presented in Table 71 and Table 72, respectively.

The base case probabilistic ICER for fezolinetant 45 mg versus no active treatment was £10,355 per QALY gained, which is below the £30,000 threshold considered cost effective. The probabilities of cost-effectiveness at cost-effectiveness thresholds of £20,000 and £30,000 are presented in Table 73.

These results demonstrate that fezolinetant is a cost-effective option for the treatment of moderate to severe VMS versus no active treatment, the comparator relevant to UK clinical practice.

Disaggregated deterministic results of the base case incremental cost-effectiveness analysis are presented in Appendix J.

Table 71: Probabilistic base-case results

Technologies	Total			Incremental			ICER	
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (£/QALY)	Incremental (£/QALY)
Fezolinetant 45 mg	£2,796.94	8.434	6.659	-	-	-	-	
No Treatment	£1,601.95	8.434	6.543	£1,194.98	0.000	0.115	£10,355	

Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life years gained; QALYs: quality-adjusted life years.

Table 72: Incremental net health benefit

Technologies	Total		Incremental		INHB	
	Costs (£)	QALYs	Costs (£)	QALYs	£20,000	£30,000
Fezolinetant 45 mg	£2,796.94	6.659	-	-	-	-
No Treatment	£1,601.95	6.543	£1,194.98	0.115	0.06	0.08

Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life years gained; QALYs: quality-adjusted life years; INHB: incremental net health benefit.

Table 73: Probability of cost-effective

Technologies	Probability of cost-effective at £20,000/QALY gained	Probability of cost-effective at £30,000/QALY gained
Fezolinetant 45 mg	100.00%	100.00%
No active treatment	0.00%	0.00%

Abbreviations: QALYs: quality-adjusted life years.

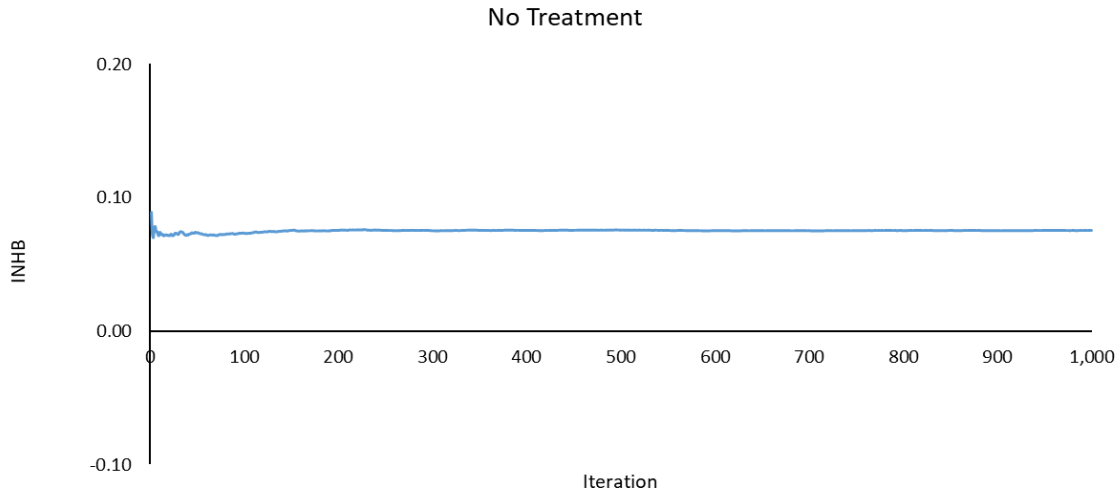
B.3.11 Exploring uncertainty

To address uncertainty in model inputs and assumptions, several sensitivity analyses were conducted, including PSA (Section B.3.11.1), deterministic sensitivity analysis (DSA; Section B.3.11.2) and scenario analyses (Section B.3.11.3). The PSA was performed to analyse the joint uncertainty of the model parameters, and the DSA and scenario analyses were used to identify model drivers or test alternative data sources.

B.3.11.1 Probabilistic sensitivity analysis

As reflected in the base case results presented in Section B.3.10.1, probabilistic sensitivity analyses with 1,000 iterations were performed to assess the uncertainty associated with model input parameters. Use of 1,000 iterations was deemed appropriate based on the results of an INHB convergence tests, as shown in Figure 31.

Figure 31: Convergence plot for INHB at £30,000/QALY



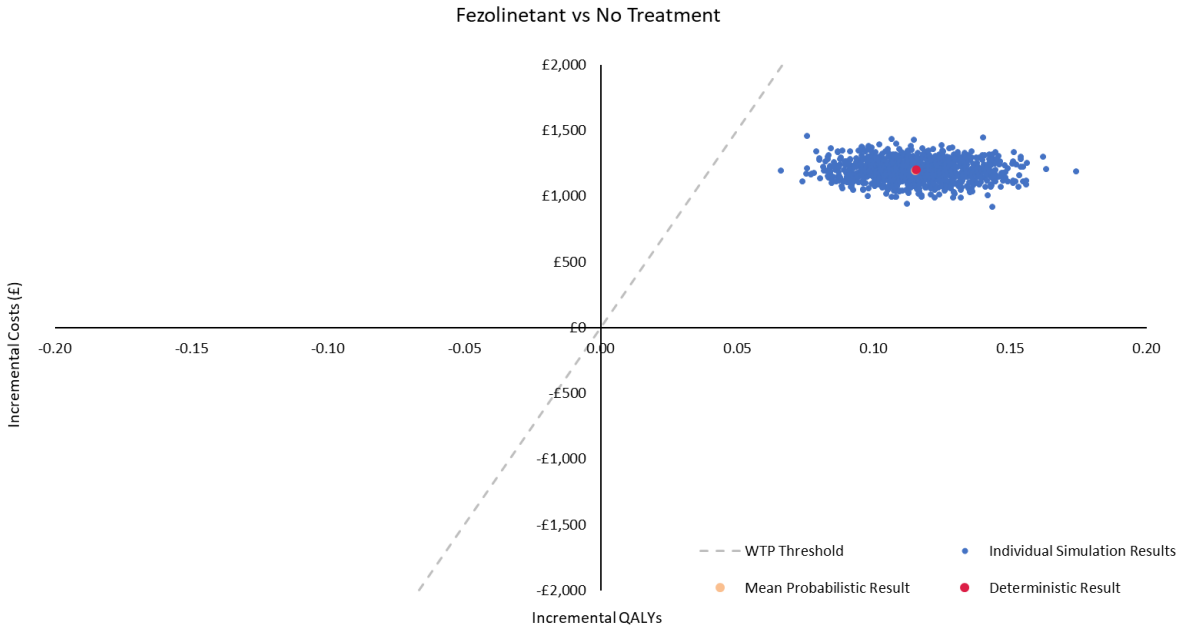
Abbreviations: INHB: incremental net health benefit; QALY: quality-adjusted life year.

The probabilistic base case results are presented in Table 71 and the cost-effectiveness plane scatter plot and cost-effectiveness acceptability curve in Figure 32 and Figure 33, respectively.

The results of the cost-effectiveness plane show little variation with regard to incremental costs in each iteration, as most iterations generally fell between £1,000 and £1,500. On the effectiveness side, all iterations fell within the northeast quadrant and below the £30,000 cost-effectiveness threshold, with modest variation in incremental QALYs gained.

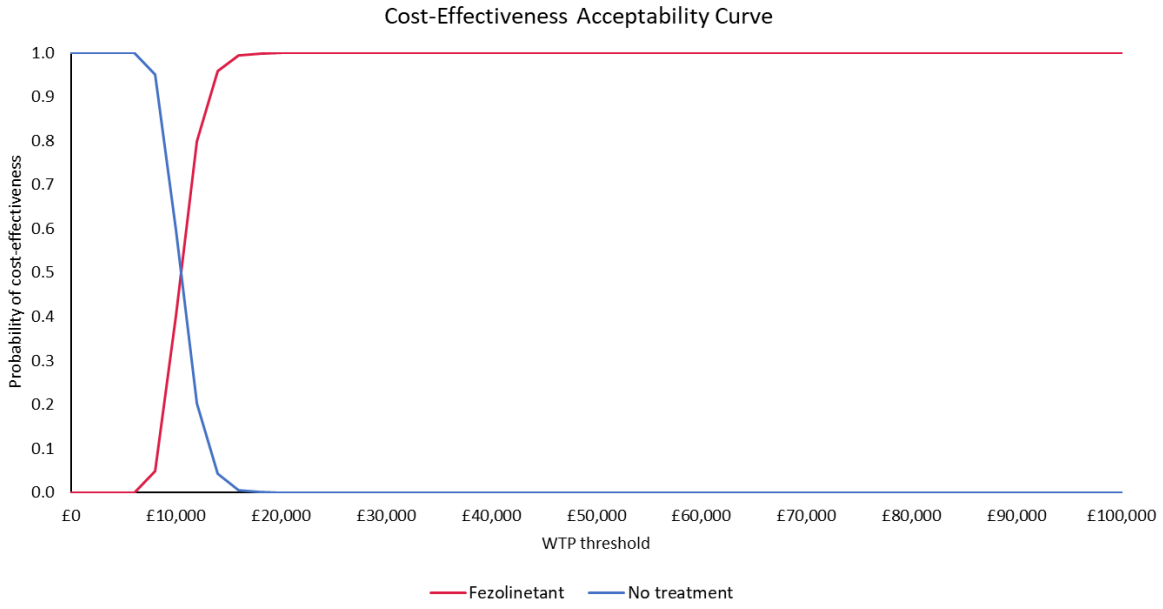
The cost-effectiveness acceptability curve shows a 100% probability of fezolinetant being cost-effective compared with no active treatment at both cost-effectiveness thresholds of £20,000/QALY and £30,000/QALY.

Figure 32: Cost-effectiveness plane for fezolinetant 45 mg versus no active treatment



Abbreviations: ICER: incremental cost-effectiveness ratio; WTP: willingness-to-pay.

Figure 33: Cost-effectiveness acceptability curve for fezolinetant 45 mg versus no active treatment



Abbreviations: WTP: willingness to pay.

B.3.11.2 Deterministic sensitivity analysis

In the DSA, each parameter of interest was changed independently while all others remained at their default base-case values. The following key model inputs were evaluated in the DSA:

- Natural VMS cessation
- Resource use
- Treatment discontinuation

- Utilities
- Patient characteristics

A summary of results in the deterministic base-case analysis and INHB are presented in Table 74 and Table 75, respectively. The deterministic base case results are in close alignment with the probabilistic base case results in Section B.3.10.1.

Table 74: Deterministic base-case results

Technologies	Total			Incremental			ICER	
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	Versus baseline (£/QALY)	Incremental (£/QALY)
Fezolinetant 45 mg	£2,803.18	8.449	6.684	-	-	-	-	
No Treatment	£1,603.69	8.449	6.568	£1,199.49	0.000	0.116	£10,364	

Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life years gained; QALYs: quality-adjusted life years.

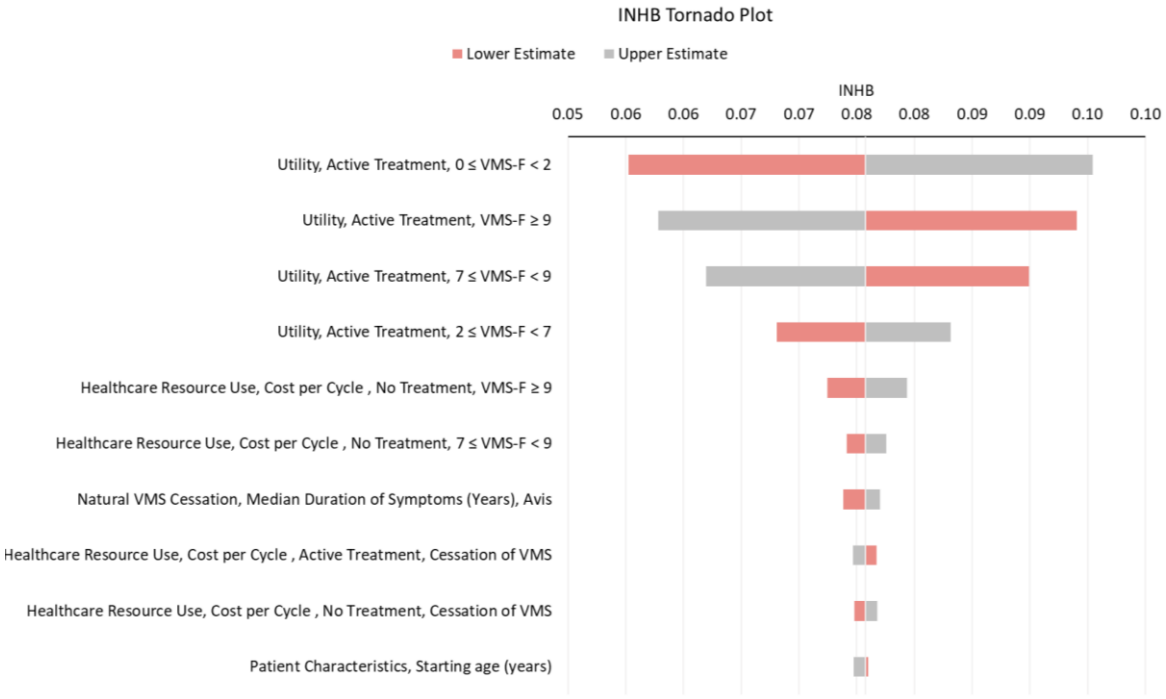
Table 75: Incremental net health benefit

Technologies	Total		Incremental		INHB	
	Costs (£)	QALYs	Costs (£)	QALYs	At £20,000	At £30,000
Fezolinetant	£2,803.18	6.684	-	-	-	-
No Treatment	£1,603.69	6.568	£1,199.49	0.116	0.06	0.08

Abbreviations: QALYs: quality-adjusted life years; INHB: incremental net health benefit.

The ten most influential variables in the DSA for the analysis of fezolinetant versus no active treatment are presented as tornado plots in Figure 34. The DSA results showed that the INHB was most sensitive to health state utility values for VMS frequency health states and was also sensitive, to a lesser extent, to healthcare resource use. None of the ten variables crossed the point of indifference (i.e., when INHB is zero) for either their upper bound or lower bound values or had any substantial impact on the cost-effectiveness estimates, showing the model to be largely robust to individual parameter uncertainty.

Figure 34: Tornado diagrams depicting results of fezolinetant versus no active treatment (INHB at £30,000/QALY)



Abbreviations: INHB: incremental net health benefit; VMS: vasomotor symptoms.

B.3.11.3 Scenario analysis

A list of the scenario analyses considered in the model, and the results from these analyses are presented in Table 76.

In scenario analyses, specific combinations of various parameters are varied at the same time, or where specific alternate assumptions are possible. Several scenario analyses were conducted to explore the impact of alternative data sources and assumptions on the base case results. In particular, the 'Post-discontinuation – No natural history' scenario explores the uncertainty with the post-discontinuation transitions, by removing the uncertainty in the natural history distribution estimates (Section B.3.3.2.3) and reducing the time horizon to the expected length of treatment, resulting in an ICER of £4,439. Additionally, the scenario 'HRT-averse population - Excludes HRT-averse population' explores the uncertainty in the HRT-unsuitable subpopulation, resulting in an ICER of £8,545.

The results of the scenario analyses show that in all scenarios (18 out of 18), the ICERs were below the £30,000/QALY cost-effectiveness threshold. These results provide reassurance that the results of the base-case analysis are robust to uncertainties in both model inputs and assumptions.

Table 76: Summary of scenario analyses and results (probabilistic)

Scenario Analysis	Base-case input	Versus no active treatment		
		Incremental costs	Incremental QALYs	ICER (£/QALY)
Base-case (deterministic)	**-**	**£1,199**	**0.12**	**£10,364**
Pooled SKYLIGHT & DAYLIGHT scenario				
Baseline patient distribution Pooled SKYLIGHT & DAYLIGHT	DAYLIGHT	£1,212	0.075	£16,128
Baseline age Pooled SKYLIGHT & DAYLIGHT	Mean age of menopause onset in the UK			
Fezolinetant efficacy Week 0–12 Pooled SKYLIGHT & DAYLIGHT 0–12 per cycle	Week 0–12 - DAYLIGHT 0–12 per cycle Week 12–24 - DAYLIGHT 12–24 per cycle Week 24–52 - SKYLIGHT 24–52 per cycle Week 52+ - SKYLIGHT 24–52 average			
'No active treatment' efficacy Pooled SKYLIGHT & DAYLIGHT 0–12 per cycle	Week 0–12 - DAYLIGHT 0–12 per cycle			
Fezolinetant treatment discontinuation Week 0–24 Pooled SKYLIGHT & DAYLIGHT	Week 0-24 - DAYLIGHT 0–24 Week 24+ - SKYLIGHT 24–52			
'No active treatment' discontinuation Week 0–24 Pooled SKYLIGHT & DAYLIGHT	Week 0–12 - DAYLIGHT 0–24			
Utility values source	Clinician adjusted DAYLIGHT estimates			

Pooled SKYLIGHT & DAYLIGHT (treatment non-specific)				
Baseline patient distribution				
Pooled SKYLIGHT & DAYLIGHT	DAYLIGHT	£1,183	0.12	£10,083
Baseline age				
DAYLIGHT	Mean age of menopause onset in the UK	£1,192	0.12	£10,349
Pooled SKYLIGHT & DAYLIGHT		£1,193	0.12	£10,330
True 'no active treatment' scenarios (i.e., no placebo effect)				
Natural history	Week 0–12 - DAYLIGHT 0–12 per cycle	£1,145	0.13	£9,105
No transitions		£1,143	0.13	£9,068
Automatic reversion to 'Post discontinuation' efficacy for no active treatment				
Week 24	Week 12	£1,238	0.11	£11,619
Post-discontinuation				
No natural history: • Time horizon: 3 years • Reset after discontinuation: reset to baseline distribution (DAYLIGHT) • Post discontinuation transitions: no transitions	Natural history informed by SEE	£530	0.12	£4,439
Natural history application method				
Use Year 1 distribution at Year 2	Only use Year 6 distribution	£1,384	0.08	£17,914
Only use Year 3 and Year 6 distributions		£1,354	0.08	£16,304
Natural history treatment waning				
No waning applied	Waning applied	£1,191	0.12	£10,239
Fezolinetant treatment discontinuation				
Week 0–24 Pooled SKYLIGHT & DAYLIGHT	Week 0–24	£1,223	0.12	£10,331

Week 24+ User-defined (1% per cycle, based on clinician feedback)	- DAYLIGHT Week 24+ - SKYLIGHT 1 & 2	£1,353	0.12	£11,036
No active treatment discontinuation				
Week 0–12 - Pooled SKYLIGHT & DAYLIGHT	Week 0–12 - DAYLIGHT	£1,195	0.12	£10,351
Utility values source				
DAYLIGHT	Clinician adjusted DAYLIGHT estimates	£1,195	0.08	£14,164
Pooled SKYLIGHT & DAYLIGHT		£1,195	0.07	£16,339
Healthcare resource use				
Whiteley et al. (2013) by treatment arm	UK clinician estimates	£1,588	0.12	£13,757
HRT-averse population				
Excludes HRT-averse population	Includes HRT-averse population	£1,185	0.14	£8,545

Abbreviations: AE: adverse events; EQ-5D: EuroQoL 5 dimensions questionnaire; HCRU: healthcare resource use; HRT: hormone replacement therapy; ICER: incremental cost-effectiveness ratio; NMA: network meta-analysis; QALY: quality-adjusted life years; SEE: structured expert elicitation; SSR: selective serotonin reuptake inhibitor; UK: United Kingdom; VMS: vasomotor symptoms.

B.3.12 Subgroup analysis

No subgroup analyses were included in the model.

B.3.13 Benefits not captured in the QALY calculation

Untreated moderate to severe vasomotor-predominant menopausal symptoms are associated with societal costs due to lost work productivity, with one study estimating a cost of approximately $1.8 billion USD per year, which are not captured in the QALY calculation in the NICE reference case.[123] Fezolinetant provides substantial benefits to work productivity and reducing activity impairment (as measured by the WPAI) which should be taken into consideration.

The unmet need for effective treatment in people with moderate to severe vasomotor-predominant menopausal symptoms who are deemed unsuitable for HRT is widely acknowledged by the UK clinical expert community. Experts have further highlighted location-dependent disparities in access to specialist menopause clinics and the length of appointment waiting lists. These disparities in access would be exacerbated if fezolinetant were not available for prescription in a primary care setting. Additionally, as specialist menopause clinics are usually run in the primary care setting by GPs with a special interest in menopause, many of these GPs will not be able to provide their patients access to fezolinetant if it is restricted to secondary care only.

As such, elements such as productivity and equity are critical benefits of fezolinetant that are not fully captured in the economic analysis.

B.3.14 Validation

B.3.14.1 Validation of cost-effectiveness analysis

External Validity

Expert input from health economists and clinicians was sought during the development of the cost-effectiveness model, and subsequently validated by patient experts, to ensure that the inputs and assumptions used in the analysis were relevant to UK clinical practice.[7]

Internal model validity

In alignment with best practice, validation of the economic model was conducted by an independent health economist prior to the submission. These quality-control procedures made use of a checklist[111] incorporating key elements of the TECH-VER checklist[124] (for technical and stress test checks) to ensure that the model generated accurate results which were consistent with input data and robust to extreme values. A technical cell by cell verification of formulae, functions and coding was performed as part of this process, as was review of all model calculations, including standalone formulae, equations and Excel macros programmed in Visual Basic for Applications. The correct functioning of the sensitivity and scenario analyses was also reviewed. The stress test ensured that the expected effect is observed when key inputs are varied in the model (e.g. when utilities for all health states for AEs are set to 0, all QALYS should result equal to 0).

B.3.15 Interpretation and conclusions of economic evidence

Summary of economic evidence for fezolinetant versus no active treatment

The cost-effectiveness of fezolinetant for treatment of moderate to severe vasomotor-predominant menopausal symptoms in people who are deemed unsuitable for HRT for medical reasons was evaluated versus no active treatment. The base case probabilistic ICER was £10,355 per QALY gained and did not differ meaningfully from the deterministic ICER (£10,364 per QALY gained). The PSA results indicated that the base-case results showed little variation with regard to incremental costs in each iteration, as most iterations generally fell between £1,000 and £1,500. On the effectiveness side, all iterations fell within the northeast quadrant and below the £30,000 cost-effectiveness threshold, with a modest variation in incremental QALYs gained. All probabilistic scenario analyses results were all below the cost-effectiveness threshold used by NICE (£30,000/QALY). Overall, the results indicate fezolinetant to be a cost-effective option for the treatment of moderate to severe vasomotor-predominant menopausal symptoms in menopausal people who are deemed unsuitable for HRT for medical reasons.

Strengths of the cost-effectiveness analysis

There are several strengths to this economic evaluation. The economic evaluation is primarily informed by the population of patients with moderate to severe VMS and deemed unsuitable for HRT enrolled in the phase 3 DAYLIGHT trial, which is aligned with the target population addressed in the decision problem of this submission.

The implementation of a Markov structure with granular VMS frequency health states and a 10-year time horizon allowed the model to adequately capture the long-term health effects and costs of fezolinetant versus no active treatment on moderate to severe VMS frequency. The model structure and model inputs were validated by health economic experts, UK clinical experts with expertise in the menopause, as well as by patient experts, ensuring the model is fit-for-purpose and accurately reflects current UK clinical practice.

Evidence sources and model settings aligned with the NICE reference case, adopting an NHS and PSS perspective.

Limitations

As discussed in B.3.3.2, efficacy inputs for no active treatment were sourced from the DAYLIGHT placebo arm. However, there are limitations to modelling no active treatment using placebo data, given that the target population are unlikely to be receiving placebo in UK clinical practice for ethical reasons and additionally, they are also unlikely to be receiving any additional follow up or GP interactions associated with partaking in a clinical trial. Consequently, the use of placebo efficacy data to inform the no active treatment arm in the model is conservative and therefore the cost-effectiveness estimate is likely to be underestimated.

While longer-term follow-up data up to Week 52 from SKYLIGHT 1 and SKYLIGHT 2 demonstrate robust evidence for the sustained benefits of fezolinetant, the lack of efficacy data beyond this timepoint means that there is uncertainty about the long-term efficacy of fezolinetant. As such, clinically plausible assumptions regarding the long-term efficacy and discontinuation of fezolinetant beyond the trial period was made in the model to address this issue, and validated with UK clinicians.

Due to the inherent uncertainty in the natural history distribution estimates that inform post-discontinuation transitions in the model, SEE estimates were generated from clinical input and further validated in order to best reflect what is seen in UK clinical practice. This was further addressed through a scenario analysis where no natural history was used post-discontinuation in order to explore the removal of this uncertainty.

Overall, DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 are high-quality, robust RCTs, thus reducing uncertainty in the clinical data that are available.

Conclusion

Due to the limited number of licensed and effective treatments for menopausal people with moderate to severe VMS for whom HRT is deemed unsuitable for medical reasons, a clear unmet need exists for novel and effective therapies that specifically target VMS. Based on the modelling estimates fezolinetant is expected to provide an increase of 0.115 QALYs versus no active treatment, and an ICER of £10,355 per QALY gained, which is below the £30,000 threshold considered cost effective. The recommendation of fezolinetant for use in this patient population would therefore address the current unmet need for an effective novel treatment option which can better control VMS and provide more sustained responses, ultimately leading to improvements in quality of life for menopausal people who are affected by the challenging symptoms of VMS.

References

1. Rance NE, Dacks PA, Mittelman-Smith MA, et al. Modulation of body temperature and LH secretion by hypothalamic KNDy (kisspeptin, neurokinin B and dynorphin) neurons: a novel hypothesis on the mechanism of hot flushes. Front Neuroendocrinol 2013;34:211-27.
2. Medicines and Healthcare Products Regulatory Agency (MHRA). Veoza 45 mg film-coated tablets. Summary of Product Characteristics. Available at: https://mhraproducts4853.blob.core.windows.net/docs/b105452ec51f5f4197a8c9f2d2ef29c79710077a [Accessed: June 2024], 2023.
3. National Institute for Health and Care Excellence (NICE). Menopause: diagnosis and management [NG23]. Available at: https://www.nice.org.uk/guidance/NG23 [Accessed September 2024].
4. Medicines and Healthcare Products Regulatory Agency (MHRA). Public Assessment Report. Clinidine Hydrochloride 25 microgram Tablets. PL 20117/0393. Available at: https://mhraproducts4853.blob.core.windows.net/docs/95a87be06e8009fa3bab7100ee1a47b8d0b850d4 [Accessed: September 2024], 2021.
5. The 2023 nonhormone therapy position statement of The North American Menopause Society. Menopause 2023;30:573-590.
6. National Institute for Health and Care Excellence (NICE). Fezolinetant for treating vasomotor symptoms associated with the menopause (ID5071). Response to stakeholder organisation comments on the draft remit and draft scope. Available at: https://www.nice.org.uk/guidance/indevelopment/gid-ta11058 [Accessed 5 August 2024].
7. Astellas Data on File. UK Individual Expert Validation Meetings Report. 2024.
8. Da Silva AS, Baines G, Araklitis G, et al. Modern management of genitourinary syndrome of menopause. Fac Rev 2021;10:25.
9. Astellas. UK HRT-unsuitable Consensus Statement.
10. Mehta J, Kling JM, Manson JE. Risks, Benefits, and Treatment Modalities of Menopausal Hormone Therapy: Current Concepts. Front Endocrinol (Lausanne) 2021;12:564781.
11. Sharma A, Davies R, Kapoor A, et al. The effect of hormone replacement therapy on cognition and mood. Clinical Endocrinology 2023;98:285-295.
12. Astellas Data on File. PAG Engagement Meetings: Fezolinetant (Veoza™) HTA appraisals.
13. Depypere H, Lademacher C, Siddiqui E, et al. Fezolinetant in the treatment of vasomotor symptoms associated with menopause. Expert Opinion on Investigational Drugs 2021;30:681-694.
14. British National Formulary (BNF). Fezolinetant. Medicinal Forms. Available at: https://bnf.nice.org.uk/drugs/fezolinetant/medicinal-forms/ [Accessed: 24 June 2024].
15. SIGN Filters. Volume 03.10.2023.
16. Whiteley J, DiBonaventura M, Wagner JS, et al. The impact of menopausal symptoms on quality of life, productivity, and economic outcomes. J Womens Health (Larchmt) 2013;22:983-90.
17. Monteleone P, Mascagni G, Giannini A, et al. Symptoms of menopause - global prevalence, physiology and implications. Nat Rev Endocrinol 2018;14:199-215.
18. Dennerstein L, Dudley EC, Hopper JL, et al. A prospective population-based study of menopausal symptoms. Obstetrics & Gynecology 2000;96:351-358.
19. Hickey M, Szabo RA, Hunter MS. Non-hormonal treatments for menopausal symptoms. Bmj 2017;359:j5101.
20. Thurston RC, Bromberger JT, Joffe H, et al. Beyond frequency: who is most bothered by vasomotor symptoms? Menopause 2008;15:841-7.
21. Hunter M, Gentry-Maharaj A, Ryan A, et al. Prevalence, frequency and problem rating of hot flushes persist in older postmenopausal women: impact of age, body mass index, hysterectomy, hormone therapy use, lifestyle and mood in a cross-sectional cohort study of 10 418 British women aged 54–65. BJOG: An International Journal of Obstetrics & Gynaecology 2012;119:40-50.
22. Wyrwich KW, Yu H. Validation of POMS questionnaire in postmenopausal women. Quality of Life Research 2011;20:1111-1121.
23. Thurston RC, Joffe H. Vasomotor symptoms and menopause: findings from the Study of Women's Health across the Nation. Obstet Gynecol Clin North Am 2011;38:489-501.

24. Archer DF, Sturdee DW, Baber R, et al. Menopausal hot flushes and night sweats: where are we now? Climacteric 2011;14:515-28.
25. Clinical Knowledge Summaries (CKS). Daignosis of Menopause and Perimenopause. Available at: https://cks.nice.org.uk/topics/menopause/diagnosis/diagnosis-of-menopause-perimenopause/ [Accessed: September 2024].
26. Hunter MS, Gentry-Maharaj A, Ryan A, et al. Prevalence, frequency and problem rating of hot flushes persist in older postmenopausal women: impact of age, body mass index, hysterectomy, hormone therapy use, lifestyle and mood in a cross-sectional cohort study of 10,418 British women aged 54-65. BJOG 2012;119:40-50.
27. Whiteley J, Wagner JS, Bushmakin A, et al. Impact of the severity of vasomotor symptoms on health status, resource use, and productivity. Menopause 2013;20:518-24.
28. Soares CN. Can depression be a menopause-associated risk? BMC Med 2010;8:79.
29. Natari RB, Clavarino AM, McGuire TM, et al. The bidirectional relationship between vasomotor symptoms and depression across the menopausal transition: a systematic review of longitudinal studies. Menopause 2018;25:109-120.
30. Fawcett. Menopause and the Workplace. Available at: https://www.fawcettsociety.org.uk/menopauseandtheworkplace [Accessed: September 2024].
31. Woods NF, Mitchell ES. Symptoms during the perimenopause: prevalence, severity, trajectory, and significance in women's lives. Am J Med 2005;118 Suppl 12B:14-24.
32. National Institute for Health and Care Excellence (NICE). Early and locally advanced breast cancer: diagnosis and management [NG101]. Available at: https://www.nice.org.uk/guidance/NG101 [Accessed September 2024].
33. Umland EM, Falconieri L. Treatment options for vasomotor symptoms in menopause: focus on desvenlafaxine. International Journal of Women's Health 2012:305-319.
34. Shapiro C.M. M NPG, Stute P, Thurston RC, Wolfman W, English M , Zhao J, Ottery FD Early Response with Fezolinetant Treatment of Moderate-to-Severe Vasomotor Symptoms Associated with Menopause: Pooled Data from Two Phase 3 Studies. 14th European Congress on Menopause and Andropause. Florence, Italy, 2023.
35. Clinical Knowledge Summaries (CKS). Menopause Scenario: Managing women with menopause, perimenopause, or premature ovarian insufficiency – Non-hormonal Treatments. Available at: https://cks.nice.org.uk/topics/menopause/management/management-of-menopause-perimenopause-or-premature-ovarian-insufficiency/#non-hormonal-treatments [Accessed: September 2024].
36. Deecher DC, Dorries K. Understanding the pathophysiology of vasomotor symptoms (hot flushes and night sweats) that occur in perimenopause, menopause, and postmenopause life stages. Arch Womens Ment Health 2007;10:247-57.
37. Freeman EW, Sammel MD, Sanders RJ. Risk of long-term hot flashes after natural menopause: evidence from the Penn Ovarian Aging Study cohort. Menopause 2014;21:924-32.
38. Gold EB, Colvin A, Avis N, et al. Longitudinal analysis of the association between vasomotor symptoms and race/ethnicity across the menopausal transition: study of women’s health across the nation. American journal of public health 2006;96:1226-1235.
39. Avis NE, Crawford SL, Green R. Vasomotor symptoms across the menopause transition: differences among women. Obstetrics and Gynecology Clinics 2018;45:629-640.
40. Kiran A, Schultz NM, Siddiqui E, et al. Epidemiology and treatment patterns of UK women diagnosed with vasomotor symptoms: Findings from the Clinical Practice Research Datalink GOLD database. Maturitas 2022;164:1-8.
41. Mittelman-Smith MA, Williams H, Krajewski-Hall SJ, et al. Role for kisspeptin/neurokinin B/dynorphin (KNDy) neurons in cutaneous vasodilatation and the estrogen modulation of body temperature. Proc Natl Acad Sci U S A 2012;109:19846-51.
42. Rossmanith WG, Ruebberdt W. What causes hot flushes? The neuroendocrine origin of vasomotor symptoms in the menopause. Gynecol Endocrinol 2009;25:303-14.
43. Jayasena CN, Comninos AN, Stefanopoulou E, et al. Neurokinin B Administration Induces Hot Flushes in Women. Scientific Reports 2015;5:8466.
44. Tepper PG, Brooks MM, Randolph Jr JF, et al. Characterizing the trajectories of vasomotor symptoms across the menopausal transition. Menopause (New York, NY) 2016;23:1067.
45. Kwon R, Chang Y, Kim Y, et al. Alcohol Consumption Patterns and Risk of Early-Onset Vasomotor Symptoms in Premenopausal Women. Nutrients 2022;14.

46. British Menopause Society (BMS). Menopause in ethnic minority women. Available at: https://thebms.org.uk/wp-content/uploads/2023/07/20-BMS-TfC-Menopause-in-ethnic-minority-women-JULY2023-B.pdf [Accessed: September 2024].
47. Iliodromiti S, Wang W, Lumsden MA, et al. Variation in menopausal vasomotor symptoms outcomes in clinical trials: a systematic review. Bjog 2020;127:320-333.
48. European Medicines Agency (EMA). GUIDELINE ON CLINICAL INVESTIGATION OF MEDICINAL PRODUCTS FOR HORMONE REPLACEMENT THERAPY OF OESTROGEN DEFICIENCY SYMPTOMS IN POSTMENOPAUSAL WOMEN. Available at: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-hormone-replacement-therapy-oestrogen-deficiency_en.pdf [Accessed: September 2024].
49. Avis NE, Crawford SL, Greendale G, et al. Duration of menopausal vasomotor symptoms over the menopause transition. JAMA Intern Med 2015;175:531-9.
50. Kumari M, Stafford M, Marmot M. The menopausal transition was associated in a prospective study with decreased health functioning in women who report menopausal symptoms. J Clin Epidemiol 2005;58:719-27.
51. Nappi RE SE, Todorova L, Rea C, Gemmen E, Schultz NM. Prevalence and quality-of-life burden of vasomotor symptoms associated with menopause: A European cross-sectional survey. Maturitas 2023;167.
52. Makara-Studzinska MT, Krys-Noszczyk KM, Jakiel G. Epidemiology of the symptoms of menopause - an intercontinental review. Prz Menopauzalny 2014;13:203-11.
53. MR-VMS Patient Landscape: Patient journey. Data on file. Astellas, 2017.
54. Lins L, Carvalho FM. SF-36 total score as a single measure of health-related quality of life: Scoping review. SAGE Open Med 2016;4:2050312116671725.
55. Kiltz U, Kiefer D, Boonen A. (Health-Related) Quality of Life as an Outcome in Studies of Axial Spondyloarthritis. Rheumatic Disease Clinics of North America 2020;46:379-393.
56. Baker FC, de Zambotti M, Colrain IM, et al. Sleep problems during the menopausal transition: prevalence, impact, and management challenges. Nat Sci Sleep 2018;10:73-95.
57. Berecki-Gisolf J, Begum N, Dobson AJ. Symptoms reported by women in midlife: menopausal transition or aging? Menopause 2009;16:1021-9.
58. Freeman EW, Sammel MD, Gross SA, et al. Poor sleep in relation to natural menopause: a population-based 14-year follow-up of midlife women. Menopause 2015;22:719-26.
59. Williams RE, Levine KB, Kalilani L, et al. Menopause-specific questionnaire assessment in US population-based study shows negative impact on health-related quality of life. Maturitas 2009;62:153-9.
60. DePree B, Shiozawa A, King D, et al. Association of menopausal vasomotor symptom severity with sleep and work impairments: a US survey. Menopause 2023;30:887-897.
61. Freeman EW, Sammel MD, Lin H, et al. Associations of hormones and menopausal status with depressed mood in women with no history of depression. Arch Gen Psychiatry 2006;63:375-82.
62. Reed SD, Ludman EJ, Newton KM, et al. Depressive symptoms and menopausal burden in the midlife. Maturitas 2009;62:306-10.
63. DiBonaventura Md, Chandran A, Hsu M-A, et al. Burden of vasomotor symptoms in France, Germany, Italy, Spain, and the United Kingdom. International Journal of Women's Health 2013:261-269.
64. Griffiths A, MacLennan SJ, Hassard J. Menopause and work: an electronic survey of employees' attitudes in the UK. Maturitas 2013;76:155-9.
65. Jack G, Riach K, Bariola E, et al. Menopause in the workplace: What employers should be doing. Maturitas 2016;85:88-95.
66. Kopenhager T, Guidozzi F. Working women and the menopause. Climacteric 2015;18:372-5.
67. Hardy C, Thorne E, Griffiths A, et al. Work outcomes in midlife women: the impact of menopause, work stress and working environment. Women's Midlife Health 2018;4.
68. University of York. UK Population Norms for EQ-5D. Available at: https://www.york.ac.uk/che/pdf/DP172.pdf [Accessed 25 July 2024].
69. Kontodimopoulos N, Pappa E, Papadopoulos AA, et al. Comparing SF-6D and EQ-5D utilities across groups differing in health status. Qual Life Res 2009;18:87-97.
70. English M, Stoykova B, Slota C, et al. Qualitative study: burden of menopause-associated vasomotor symptoms (VMS) and validation of PROMIS Sleep Disturbance and Sleep-Related

Impairment measures for assessment of VMS impact on sleep. J Patient Rep Outcomes 2021;5:37.
71. National Institute for Health and Care Excellence (NICE). Menopause: diagnosis and management [GID-NG10241] – In Development. Available at: https://www.nice.org.uk/guidance/indevelopment/gid-ng10241 [Accessed: September 2024].
72. National Institute for Health and Care Excellence (NICE). STA Fezolinetant for treating vasomotor symptoms associated with the menopause - Equality impact assessment. Available at: https://www.nice.org.uk/guidance/gid-ta11058/documents/equality-impact-assessment-scoping. [Last accessed: August 2024].
73. Green R, Santoro N. Menopausal symptoms and ethnicity: the Study of Women's Health Across the Nation. Women's Health 2009;5:127-133.
74. Harlow SD, Burnett-Bowie S-AM, Greendale GA, et al. Disparities in reproductive aging and midlife health between Black and White women: the Study of Women's Health Across the Nation (SWAN). Women's midlife health 2022;8:1-17.
75. ClinicalTrials.gov. A Study of Fezolinetant to Treat Hot Flashes in Women Going Through Menopause (Daylight). Available at: https://clinicaltrials.gov/study/NCT05033886 [Accessed: September 2024].
76. ClinicalTrials.gov. A Study to Find Out if Fezolinetant Helps Reduce Moderate to Severe Hot Flashes in Women Going Through Menopause (Skylight 1). Available at: https://clinicaltrials.gov/study/NCT04003155 [Accessed: September 2024].
77. ClinicalTrials.gov. A Study to Find Out if Fezolinetant Helps Reduce Moderate to Severe Hot Flashes in Women Going Through Menopause - 2 (Skylight 2). Available at: https://clinicaltrials.gov/study/NCT04003142 [Accessed: September 2024].
78. Astellas Data on File. Pre-specified Analysis Report for Studies 2693-CL-0301 and 2693-CL-0302 for the 12-week Double-Blind Period.
79. Astellas Data on File. DAYLIGHT Study Protocol.
80. Astellas Data on File. SKYLIGHT 1 Study Protocol.
81. Astellas Data on File. SKYLIGHT 2 Study Protocol.
82. Astellas Data on File. DAYLIGHT Clinical Study Report.
83. Astellas Data on File. SKYLIGHT 1 Clinical Study Report.
84. Astellas Data on File. SKYLIGHT 2 Clinical Study Report.
85. Astellas Data on File. SKYLIGHT 1 and SKYLIGHT 2 Integrated Summary of Efficacy.
86. Food and Drug Administration (FDA). Guidance for Industry. Estrogen and Estrogen/Progestin Drug Products to Treat Vasomotor Symptoms and Vulvar and Vaginal Atrophy Symptoms — Recommendations for Clinical Evaluation. Available at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/estrogen-and-estrogenprogestin-drug-products-treat-vasomotor-symptoms-and-vulvar-and-vaginal-atrophy [Accesssed 15 September 2023], 2003.
87. Temple R. Hy's law: predicting serious hepatotoxicity. Pharmacoepidemiology and Drug Safety 2006;15:241-243.
88. Office for National Statistics (ONS). Ethnicity facts and figures. Available at: https://www.ethnicity-facts-figures.service.gov.uk/ [Accessed 29 July 2024].
89. Avis NE, Coeytaux RR, Isom S, et al. Acupuncture in Menopause (AIM) study: a pragmatic, randomized controlled trial. Menopause 2016;23:626-37.
90. Johnson KA, Martin N, Nappi RE, et al. Efficacy and safety of fezolinetant in moderate to severe vasomotor symptoms associated with menopause: a phase 3 RCT. The Journal of Clinical Endocrinology & Metabolism 2023.
91. Cano A, Nappi RE, Santoro N, et al. Fezolinetant impact on health-related quality of life for vasomotor symptoms due to the menopause: Pooled data from SKYLIGHT 1 and SKYLIGHT 2 randomised controlled trials. BJOG: An International Journal of Obstetrics & Gynaecology 2024;131:1296-1305.
92. Astellas Data on File. SKYLIGHT 1 and SKYLIGHT 2 Pre-specified Analysis Report.
93. Astellas Data on File. SKYLIGHT 1 and SKYLIGHT 2 Pooled Analysis Table 9.3.2.3.1.
94. ClinicalTrials.gov. A Study to Confirm if Fezolinetant Helps Reduce Hot Flashes in Women With Breast Cancer Who Are Having Hormone Therapy. Available at: https://clinicaltrials.gov/study/NCT06440967?intr=fezolinetant&viewType=Table&page=1&limit=100&rank=18 [Accessed: 4 July 2024].

95. Lensen S, Archer D, Bell RJ, et al. A core outcome set for vasomotor symptoms associated with menopause: the COMMA (Core Outcomes in Menopause) global initiative. Menopause 2021;28:852-858.
96. Personal Social Services Research Unit (PSSRU). Unit Costs of Health and Social Care.
97. Beaudoin F, McQueen R, Wright A. Fezolinetant for Moderate to Severe Vasomotor Symptoms Associated with Menopause: Effectiveness and Value; Final Evidence Report. Institute for Clinical and Economic Review 2023.
98. Avis NE, Crawford SL, Greendale G, et al. Duration of menopausal vasomotor symptoms over the menopause transition. JAMA Intern Med 2015;175:531-9.
99. Coon C, Bushmakin A, Tatlock S, et al. Evaluation of a crosswalk between the European Quality of Life Five Dimension Five Level and the Menopause-Specific Quality of Life questionnaire. Climacteric 2018;21:566-573.
100. Sarrel P, Portman D, Lefebvre P, et al. Incremental direct and indirect costs of untreated vasomotor symptoms. Menopause 2015;22:260-6.
101. Institute for Clinical Economic Review (ICER). Final Evidence Report: Fezolinetant for Moderate to Severe Vasomotor Symptoms Associated with Menopause: Effectiveness and Value 2023.
102. Coyle D, Cranney A, Tugwell P. Economic evaluation of norethisterone acetate/ethinylestradiol (FemHRT) for women with menopausal symptoms. Pharmacoeconomics 2003;21:661-9.
103. Brown A, Coyle D, Chen S, et al. Transdermal Hormone Replacement Therapy Patches for Women with Postmenopausal Symptoms: Economic Analysis of Short Term Use. Technology report no 61. Ottawa: Canadian Coordinating Office for Health Technology Assessment, 2006.
104. Office for National Statistics (ONS). National life tables: UK. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables [Accessed on 6 June 2024].
105. Welton AJ, Vickers MR, Kim J, et al. Health related quality of life after combined hormone replacement therapy: randomised controlled trial. BMJ 2008;337:a1190.
106. Brazier JE, Roberts J, Platts M, et al. Estimating a preference-based index for a menopause specific health quality of life questionnaire. Health and Quality of Life Outcomes 2005;3:13.
107. ClinicalTrials.gov. Trial to Investigate the Effect of ESN364 in Early Postmenopausal Women Suffering From Hot Flashes.
108. European Medicines Agency (EMA). Guideline on clinical investigation of medicinal products for hormone replacement therapy of oestrogen deficiency symptoms in postmenopausal women.
109. Food and Drug Administration (FDA). Guidance for Industry. Estrogen and Estrogen/Progestin Drug Products to Treat Vasomotor Symptoms and Vulvar and Vaginal Atrophy Symptoms – Recommendations for Clinical Evaluation., 2003.
110. Morga A, Ajmera M, Gao E, et al. Systematic review and network meta-analysis comparing the efficacy of fezolinetant with hormone and nonhormone therapies for treatment of vasomotor symptoms due to menopause. Menopause 2023:10.1097/GME.0000000000002281.
111. Astellas Data on File. Model QC Checklist.
112. NHS inform. Menopause. Available at: https://www.nhsinform.scot/healthy-living/womens-health/later-years-around-50-years-and-over/menopause-and-post-menopause-health/menopause. [Accessed: September 2024], 2022.
113. Freeman EW, Ensrud KE, Larson JC, et al. Placebo improvement in pharmacologic treatment of menopausal hot flashes: time course, duration, and predictors. Psychosom Med 2015;77:167-75.
114. Miyazaki K, Kaneko M, Narukawa M. Factors associated with high placebo response in clinical studies of hot flashes: a meta-analysis. Menopause 2022;29:239-246.
115. Astellas. (Data on File). Structured Expert Elicitation on Vasomotor Symptoms in Menopause to Inform Health Technology Assessment in the UK.
116. University of York. Structured expert elicitation resources (STEER). Available at: https://www.york.ac.uk/che/research/teehta/elicitation/steer [Accessed on 20 August 2024].
117. Politi MC, Schleinitz MD, Col NF. Revisiting the duration of vasomotor symptoms of menopause: a meta-analysis. J Gen Intern Med 2008;23:1507-13.

118. Svartberg J, von Mühlen D, Kritz-Silverstein D, et al. Vasomotor symptoms and mortality: the Rancho Bernardo Study. Menopause 2009;16:888-91.
119. National Institute for Health and Care Excellence (NICE). NICE health technology evaluations: the manual. NICE process and methods [PMG36].
120. 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;41:199-207.
121. M Hernández Alava SP, A Wailo,. Estimating EQ-5D by age and sex for the UK., 2022.
122. National Health Service (NHS). 2021/22 National Cost Collection Data Publication.
123. Sarrel P, Portman D, Lefebvre P, et al. Incremental direct and indirect costs of untreated vasomotor symptoms. Menopause 2015;22:260-266.
124. Büyükkaramikli NC, Rutten-van Mölken MPMH, Severens JL, et al. TECH-VER: A Verification Checklist to Reduce Errors in Models and Improve Their Credibility. PharmacoEconomics 2019;37:1391-1408.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Fezolinetant for treating vasomotor symptoms associated with the menopause [ID5071]

Summary of Information for Patients (SIP)

October 2024

Template version	Date amended	Changes since previous version
2.0	Dec 2023	Clarifications made to guidance notes in section 3i regarding inclusion of statements on cost effectiveness.

File name	Version	Contains confidential information	Date
ID5071_Fezolinetant_NICE_SIP_Final_2.0 [No CON].docx	2.0	No	20th November 2024

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

1a) Name of the medicine (generic and brand name):

Generic name: fezolinetant; **Brand name:** Veoza™

1b) Population this treatment will be used by. Please outline the main patient population that is being appraised by NICE:

This submission focusses on the use of fezolinetant in the treatment of moderate to severe **vasomotor-predominant menopausal symptoms**, in people who are deemed unsuitable for treatment with **hormone replacement therapy** (**HRT**) for medical reasons.

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.

Fezolinetant has been approved by the **Medicines and Healthcare products Regulatory Agency** (**MHRA**) in the UK and is indicated for the "treatment of moderate to severe vasomotor symptoms associated with menopause". This **marketing authorisation** (**MA**) for fezolinetant was granted on 14th December 2023. Further information on the MA status for fezolinetant is available at https://products.mhra.gov.uk/product/?product=veoza%2045%20mg%20film-coated%20tablets

Please note: Further explanations for the words and phrases highlighted in **black bold text** are provided in the glossary (Section 4b). Cross-references to other sections or documents are highlighted in orange.

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:

The table below shows support from Astellas to relevant patient advocacy groups in the United Kingdom (UK), and how the company engages or supports these charities and/or patients who use them.

Patient group:	Engagement/activity with each group:	Financial support provided:
Menopause Support UK	Speaker and participant in Policy Roundtable Meeting discussing Women's Health Hubs	£935.00
Menopause Support UK	Review of survey data (YouGov survey on people with lived experience of pre-, peri-, menopause and post-menopause) and collaboration in developing outputs to share data externally	£1,275.00
Menopause Support UK	Meeting with Astellas HEOR team to validate model assumptions from patient perspective	£282.00
Menopause Alliance	Participation in disease awareness programme providing insights on the disparity of care in the UK for women from ethnic backgrounds	£170.00
Menopause Alliance	Meeting with Astellas HEOR team to validate model assumptions from patient perspective	£282.00
Menopause & Cancer	Grant funding to support core costs	£40,369.00
Menopause & Cancer	Meeting with Astellas HEOR team to validate model assumptions from patient perspective	£282.00

SECTION 2: Current landscape

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.

Please note: Most people going through menopause are biological females who identify as women. But Astellas recognises that some people experiencing menopause may not identify as women. To be inclusive, Astellas use the term 'menopausal people' in this submission. Astellas only use 'women' when the data or context specifically refers to women.

The condition that fezolinetant plans to treat is VMS

What are VMS?

Vasomotor symptoms (VMS) are the umbrella term for **hot flushes** (or hot flashes) and **night sweats**.[1] Although they may be caused by a range of different conditions, they are most commonly associated with the changes that occur to a person's body during the menopause.[1]

The experience of VMS varies widely between menopausal people, and may be affected by factors such as your race, ethnicity, weight, etc. However, the most commonly reported symptoms include:[1]

- A feeling of heat in the upper part of your body. The intensity of this may be anywhere from mild to severe,[2] and is usually most noticeable in the face, neck and upper chest
- Alongside a feeling of heat, you may also have a red, flushed face
- You may notice an increase (**tachycardia**) or irregularity (**palpitations**) in your pulse
- Your body may sweat to try to cool you down
- A feeling of cold may occur as the hot flush subsides
- You may feel dizzy (**vertigo**) or even anxious

What causes VMS?

Menopause is a natural process that usually affects biological females over the age of 45 years. It is associated with a decline in the ability to bear children due to reduced activity of the ovaries.

Human reproduction is regulated by chemical messengers (**hormones**), including **luteinising hormone** (**LH**) and **follicle-stimulating hormone** (**FSH**).These hormones are released by the pituitary gland, a small pea-sized gland located at the base of your brain.[3] As you age, your ovaries become less sensitive to these hormones and, in turn, produce less of the reproductive hormones such as **oestradiol** and **oestrogen**. As the production of these hormones decreases further, the function of your ovaries will also decrease, eventually leading to a complete stop in your menstrual cycle (period).[4] This is also known as **amenorrhoea**.[4] Your final menstrual period is known as the menopause, and menopausal people are considered **postmenopausal** following 12 months of amenorrhoea or 6 months of amenorrhoea and a blood test showing high levels of FSH (over 40 IU/L).[4, 5]

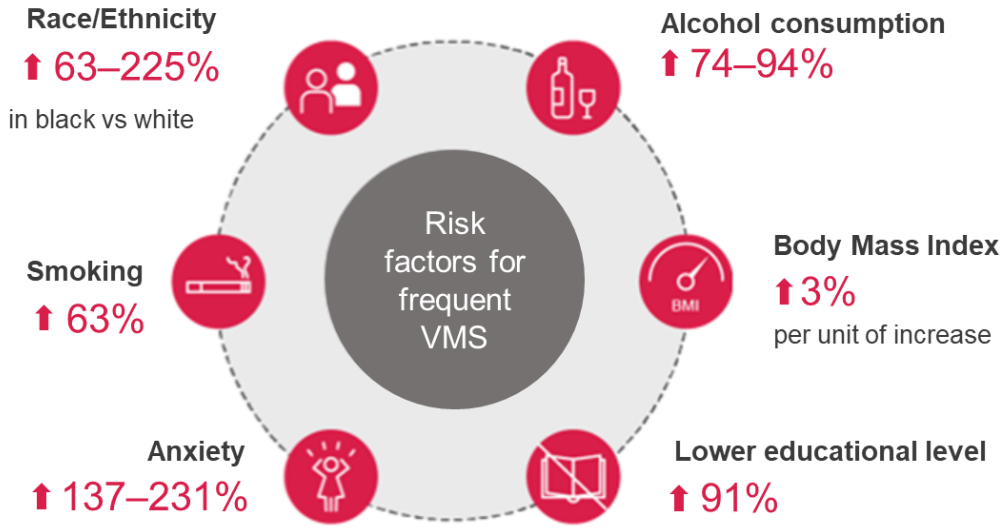
While the exact causes of VMS are not fully understood,[6] it is thought that the large changes in hormone levels that occur during the menopause (particularly of oestrogen) are what causes the appearance of various menopause-related symptoms. These symptoms include VMS, problems with sleep, low moods and problems when urinating.[7-10]

What are the risk factors for VMS?

Your individual experience of VMS may vary based on a combination of genetic background, psychological factors, and other health conditions or treatments you may be receiving.[11] However, on average, postmenopausal women with VMS have been observed to experience around 17 hot flushes and 11 night sweats per week.[12, 13]

While there is variation in the reported **frequency** of VMS, certain risk factors – such as demographic factors (including race or ethnicity), health conditions, and lifestyle factors like smoking, obesity and alcohol consumption – may increase the likelihood of experiencing more symptoms, as shown in Figure 1.[14, 15]

Figure 1: Risk factors for VMS frequency



Percentages show the increased risk of VMS for each risk factor, based on two studies involving 3,198 and 1,455 women.[14, 15]
Abbreviations: VMS: vasomotor symptoms.

How does VMS progress over time?

VMS are usually mild and infrequent in the early stages of the menopause transition (also known as **perimenopause**).[16] As time progresses you may notice that VMS are occurring more frequently and with a greater severity.[14, 16, 17] This may continue throughout the perimenopausal period.

Duration of VMS

On average, from the time you first experience VMS, you may continue to experience VMS for 7.4 years, including 4.5 years after your final period.[18] However, the overall duration of VMS is influenced by the menopausal stage during which your symptoms first appear.[18] People who first report VMS during the early stages of menopause (**premenopause** or early perimenopause) have VMS for the longest time, lasting up to 12 years, with symptoms persisting for over 9 years after their final menstrual period.[18]

How many menopausal people get VMS?

Researchers estimate that VMS are experienced by up to 80% of women during menopause.[9, 14]

A study in the UK found that the **prevalence** of moderate to severe VMS remained stable between 2009–2018, averaging 21.1% in women aged 40–65 years.[19]

What is the impact of VMS?

Physical impact and effect on quality of life

The occurrence of VMS can cause physical discomfort, anxiety, stress, self-consciousness, embarrassment, and sleep loss, which can ultimately have a considerable negative impact on your quality of life (QoL).[13, 20, 21] When asked, over 8 out of 10 women experiencing VMS reported their symptoms to significantly affect their ability to take care of themselves.[22] The unpredictable nature of VMS flare-ups were also reported as a source of anxiety, and caused these women to avoid social or leisure situations.[13]

VMS were also associated with lower **Short Form-36** (**SF-36**) scores, indicating poorer health and functioning. The SF-36 is a widely used, and respected, questionnaire designed to assess the impact of health conditions on people's QoL. It assesses QoL across a range of domains including, physical functioning, emotional wellbeing, social functioning and pain. Women with VMS had three times lower physical functioning scores than women without VMS.[21]

Women also report their sleep as being very frequently affected during menopause.[20] Women transitioning through menopause, and those who are postmenopausal are 60–67% more likely to experience sleep difficulties than premenopausal women.[23, 24] As VMS can occur during both the day and night, VMS specifically may contribute to a lack of sleep (**insomnia**) and sleep disruption in people undergoing menopause. Researchers noted that, in a group of 255 women, moderate to severe hot flushes were **statistically significantly** associated with poor sleep.[25] Considering the impact that VMS may have on sleep, treatments that lower VMS frequency and severity could in turn improve your sleep quality, which may have associated benefits, such as increased energy, concentration and work productivity.

Emotional impact and effect on mental health

The menopause is also associated with an impact on mental health and has been reported to be linked to symptoms of both anxiety and depression. In a survey of 4,014 women in the UK, 69% said they had experienced difficulties with anxiety or depression due to menopause.[26] Although the relationship between the menopause and depression has been widely debated, VMS may play a key role.[27, 28] Among 1,358 women in a US survey, VMS, particularly more severe VMS, were statistically significantly associated with depressive symptoms.[29]

Socioeconomic impact

It is believed that the impact of VMS on sleep, depression, and QoL, as well as other areas of daily life, are likely to also have an impact on work productivity. As more and more menopausal-aged people occupy senior positions in the work force, the impact of VMS can have considerable societal costs. In a survey of 4,014 women in the UK, 44% reported that menopausal symptoms affected their ability to work, while 61% said they experienced a loss of motivation at work.[26] In the same UK-based survey, 10% of women who had been employed during the menopause reported that they had left work due to menopausal symptoms, with a further 13% of women reporting that they had considered leaving their job.[26] Another large survey of over 3,800 women across France,

Germany, Italy, Spain, and the UK reported that more intense and frequent bouts of VMS were associated with a greater effect on women's productivity at work.[30] Despite the considerable productivity losses and associated costs due to VMS, menopause-related challenges are rarely discussed in the workplace.[31-33] This may further worsen the impact of VMS, as the challenges of coping with symptoms at work may lead to reduced performance, greater perceived workplace stress, and a desire to exit the workplace.[34]

Furthermore, there are considerable costs to the healthcare system associated with the treatment of VMS, especially given the large number of affected menopausal people. A large survey of women from the UK reported a substantial burden on postmenopausal women with severe VMS, which is associated with increased healthcare resource use and costs.[30] A statistically significant effect of VMS was also observed with respect to the number of physician visits, and the number of visits further increased with severity of VMS.[30]

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?

How are VMS diagnosed?

Your doctor can usually diagnose VMS based on a description of your symptoms, your age, and how far along the menopausal transition you are.[1] Your doctor will not usually require you to undergo any tests to confirm VMS if you are otherwise healthy, and over 45 years old.[4]

However, if you are below the age of 45, and the doctor suspects you may be showing symptoms of VMS or menopause, they may ask you to perform a blood test to help confirm your diagnosis. This test usually measures the levels of FSH in your blood.[4]

During a visit to your doctor, a healthcare professional will take a blood sample from a vein in your arm, using a small needle. After the needle is inserted, a small amount of blood will be collected into a test tube or vial. Taking the blood sample usually takes less than five minutes.[35]

Classification of VMS

Your doctor may also classify the severity of your VMS based on the symptoms you describe to them. The classification of VMS helps to direct the doctor towards the best possible treatment options for managing your VMS.

Severity is not easily defined due to its subjective nature, and because the experience of VMS varies widely among menopausal people. Therefore, VMS severity can be defined in various ways. However, for the purpose of the DAYLIGHT trial, we used definitions provided by the **European Medicines Agency** (**EMA**), as shown in Table 1.[2]

Table 1: Classification of VMS[2]

Severity of VMS	Definition
Mild	A sensation of heat, without sweating
Moderate	A sensation of heat with sweating. You are able to continue with normal activities.
Severe	A sensation of heat with sweating. You are not able to continue with normal activities, due to the severity of your symptoms

Abbreviations: VMS: vasomotor symptoms.

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.

What are the current treatment options for VMS?

Recommended treatment options for VMS in the UK are set out in the guidelines published by the National Institute for Health and Care Excellence (NICE) [NG23] and [NG101].[4, 36] Your doctor will usually refer to these guidelines when deciding on the best course of treatment for you, and your personal management plan will usually be dependent on a combination of your own personal and medical history, and the symptoms you present with, at the time of your appointment.

A breakdown of the treatments available in the UK for treating VMS, are presented below, correct at the time of submission.

Hormonal Therapy

As discussed in **Section 2a**, it is thought that VMS are linked with the hormonal changes that occur within the body during menopause. As you go through the menopause, the levels of certain hormones in your body, such as oestrogen, decrease, and therefore lead to physical changes in the body, ultimately presenting as hot flushes or night sweats. To correct these hormonal imbalances, your doctor may prescribe for you a course of HRT.

HRT offers effective relief of a broad range of menopausal symptoms including VMS, mood disturbances, sleep disturbances, cognitive changes, and bone health.[4] HRT includes oestrogens and, in some cases, **progestogens**. The effectiveness of HRT in treating a range of menopausal symptoms underscores its use as the initial treatment option for VMS in menopausal people who are suitable for HRT. According to NICE NG23, menopausal people can be offered HRT for VMS after discussing with your doctor the short-term (up to 5 years) and longer-term benefits and risks.[4]

If prescribed for you, your HRT treatment plan may depend on whether you have a uterus or not;

- Oestrogen, and progestogen, if you have a uterus
- Oestrogen only if you do not have a uterus

This difference is because oestrogen alone can increase the risk of uterine cancer, so adding progestogen helps protect the uterus from this risk.

Despite the notable improvements to health and QoL, long-term HRT treatments are also linked with a number of potential side effects including:

- An increased risk of blood clots in the “deep veins” or lungs (**venous thromboembolism** [**VTE**])
- A small increase in the risk of stroke
- An increase in the risk of breast cancer (when both oestrogen and progestogen are given)

These side effects are dependent on individual characteristics, such as age and risk factors. Before prescribing HRT, your doctor will weigh these risks against the benefits of treatment. However, not every person can or prefers to use HRT. Some menopausal people may have medical conditions

that make HRT unsafe (**HRT-contraindicated**), or their doctor may advise against it due to potential risk factors (**HRT-caution**). Others may have already tried HRT and found it did not work or caused unwanted side effects (**HRT-stoppers**). These individuals may be deemed unsuitable for HRT for medical reasons. Menopausal people who prefer not to use HRT (**HRT-averse**) are considered by the company to still be suitable for HRT.

For menopausal people deemed unsuitable for HRT, your doctor will further assess your medical and family history to identify the best course of action, going forward.

Other pharmacological treatments

If you are deemed unsuitable for HRT, other **pharmacological** treatments are sometimes used in UK clinical practice, including:

Therapies **licensed** for VMS:

- Clonidine

Therapies not licensed for VMS:

- Antidepressants, such as selective serotonin reuptake inhibitors (SSRIs) and serotonin and norepinephrine reuptake inhibitors (SNRIs)
- Anti-convulsant, such as gabapentin and pregabalin

However, NICE NG23 does not recommend their use as the initial treatment option for VMS alone because they have limited effectiveness and unpleasant side effects.[4] The unpleasant side effects often led people to stop treatment. NICE NG101 suggests that SSRI antidepressants can be considered for those with breast cancer, but not for those taking tamoxifen, as some SSRI antidepressants may reduce its anti-cancer effect.[36]

Non-pharmaceutical treatments

NICE NG23 also discusses the place of non-pharmaceutical treatments, such as complementary or herbal remedies (**isoflavones** and **black cohosh**), in managing VMS.[4] Due to the lack of licensed and effective pharmacological treatments for menopausal people deemed unsuitable for HRT, many menopausal people choose to use non-pharmaceutical treatments. Multiple preparations of non-pharmaceutical treatments are available. Despite evidence that isoflavones and black cohosh may relieve VMS, the safety of these non-pharmaceutical treatments across preparations is unclear. Interactions between non-pharmaceutical treatments and other medicines have also been reported.[4]

Because **cognitive behaviour therapy** (**CBT**) is being appraised in the partial update to the NICE NG23 draft gudiance,[37] CBT is a possible treatment option for VMS. The draft guidance suggested that CBT should be an option rather than a routine treatment for all.[37]

Where does fezolinetant fit in the treatment pathway for VMS?

As mentioned earlier, because HRT treats menopausal symptoms beyond VMS alone, HRT remains the initial treatment option for VMS in menopausal people who are suitable for HRT. Because of this, UK clinical experts, who were consulted by Astellas for their insights, believe that fezolinetant will be used in menopausal people who are deemed unsuitable for HRT for medical reasons in NHS primary care.[38] This includes people who are HRT-contraindicated, HRT-caution and HRT-stoppers, but not those who are c as they are considered to still be suitable for HRT. Experts consulted by Astellas anticipate fezolinetant will be used ahead of other non-hormonal

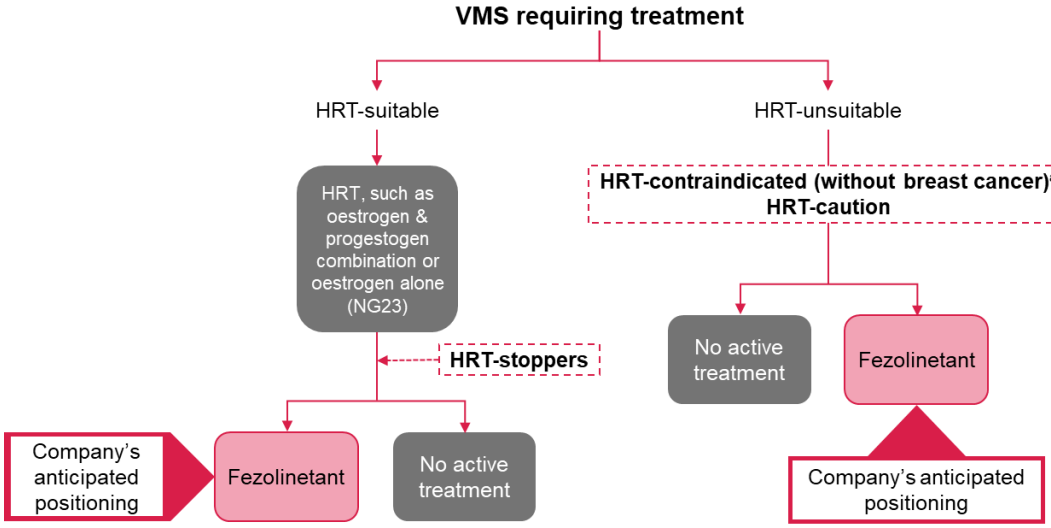
pharmacological treatments because of their limited effectiveness in treating VMS and unpleasant side effects.[38]

For non-pharmaceutical treatments, because CBT is viewed as a long-term strategy for treating VMS rather than an immediate relief intervention, UK clinical expert opinion gathered by Astellas is that CBT will be used alongside pharmacological treatments as supportive care for VMS. Complementary or herbal remedies are not recommended by NICE for the treatment of VMS and are not typically funded by the NHS. As such, they are not used in NHS clinical practice.

Given these considerations, Astellas believe that no active treatment is the only relevant comparator to fezolinetant in its anticipated place in the current treatment pathway, as presented in Figure 2.

For further details related to the place of fezolinetant in clinical practice, please refer to Section B.1.3.3. of the main Company submission.

Figure 2: Anticipated place of fezolinetant in the UK clinical pathway



[a] Fezolinetant should not be used by women who have breast cancer or are currently being treated for it. If a woman has finished her cancer treatment, her doctor should carefully weigh the benefits and risks before prescribing fezolinetant to manage.[39]

Abbreviations: HRT: hormonal therapy; VMS: vasomotor symptoms.

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.

Researchers have performed several studies to better describe women's experience of VMS, and to identify their needs. As such, the effect of VMS on daily living are well documented. A few examples of patient descriptions of VMS are presented in **Figure 3**.

Figure 3: Patient experiences of daily life with VMS[40, 41]

Abbreviations: VMS: vasomotor symptoms.

SECTION 3: The treatment

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.

How does fezolinetant work?

As discussed in **Section 2a**, oestrogens are hormones made by the ovaries. The ovaries normally stop making oestrogens during menopause. When oestrogen levels begin dropping, some menopausal people get the very bothersome VMS, also discussed in **Section 2a** earlier. Fezolinetant works to reduce these symptoms.

Unlike other non-hormonal treatments currently available in NHS clinical practice, fezolinetant directly acts on the area of the brain (hypothalamus) responsible for controlling body temperature. Fezolinetant is the first and only licensed treatment designed to specifically block neurokinin B (NKB). NKB is a chemical that binds to a receptor in your brain, triggering VMS.

Further details relevant to how fezolinetant works are available in the Patient Information Leaflet available at https://mhraproducts4853.blob.core.windows.net/docs/ebabff48caf8f7d55593db2d825e9f3a49765efd

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.

Fezolinetant is not intended to be used in combination with any other treatment for the management of moderate to severe VMS.

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?

How is fezolinetant administered?

Fezolinetant is available in tablet form and is to be taken orally (by mouth). You should be able to take fezolinetant from the comfort of your home and you should not need to see your doctor or go to a hospital for help with taking your medication. You should, however, take care to take the medication exactly as described by your doctor or pharmacist.

How should I take fezolinetant?

Each tablet of fezolinetant contains 45 mg of the active drug. You should take **one tablet of fezolinetant per day**, and at roughly the same time every day. The tablet can be taken with or without food and should be swallowed whole. You should not break, crush or chew the tablet.

You should have regular discussions with your doctor to discuss how well fezolinetant is working for you, however, you will usually be required to take fezolinetant for as long as you need it to control your symptoms. You should not stop taking this medicine unless your doctor tells you to do so.

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.

A number of clinical trials have been conducted to investigate how well fezolinetant works (its **efficacy**) in reducing the number, and severity, of moderate and severe VMS experienced by postmenopausal women. These included the **DAYLIGHT**, **SKYLIGHT 1** and **SKYLIGHT 2** trials. These trials also investigated the safety of fezolinetant, and the impact of fezolinetant on QoL. To investigate how well the body tolerates the prolonged use of fezolinetant, a long-term safety study, **SKYLIGHT 4**, was also performed.

Across all studies, fezolinetant was compared against a dummy treatment (**placebo**). The allocation of fezolinetant or placebo was random. Participants therefore received one of the following treatment options:

1. **Fezolinetant 45 mg**
2. **Placebo**

* Fezolinetant 30 mg – *fezolinetant 30 mg was a potential treatment option in the SKYLIGHT trials, however as it does not represent the licenced dose for fezolinetant to treat VMS, any results for these groups of participants are not relevant for this submission*

The main clinical evidence for fezolinetant as a treatment for participants with moderate or severe VMS is from the DAYLIGHT trial. In this study, participants were selected for inclusion if they:

- were adult women, between the ages of 40 and 65
- were confirmed as being menopausal, and had to be seeking treatment for VMS
- had a minimum of seven moderate to severe VMS per day
- were unsuitable to receive HRT therapy

The SKYLIGHT 1, SKYLIGHT 2 and SKYLIGHT 4 trials included a similar patient population, however, they did not include a requirement for participants to be unsuitable for HRT.

A summary of the key information about each trial is provided in Table 2.

Table 2: Trials investigating fezolinetant

Trial name and number	Location	Participants included	Completion date

DAYLIGHT (NCT05033886)	International (Belgium, Canada, Czechia, Denmark, Finland, France, Germany, Hungary, Italy, Netherlands, Norway, Poland, Spain, Sweden, Turkey, United Kingdom)	453	April 2023
SKYLIGHT 1 (NCT04003155)	International (Canada, Czechia, Hungary, Poland, Spain, United Kingdom, United States)	527	August 2021
SKYLIGHT 2 (NCT04003142)	International (Canada, Czechia, Latvia, Poland, Spain, United Kingdom, United States)	501	April 2021
SKYLIGHT 4 (NCT04003389)	International (Canada, Czechia, Latvia, Poland, Spain, Ukraine, United Kingdom, United States)	1831	January 2022

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.

How was the effectiveness of fezolinetant measured?

In the DAYLIGHT trial, the effectiveness of fezolinetant was measured by comparing how well fezolinetant reduced the frequency and severity of VMS experienced by the participants in the trial, against placebo. This was mainly measured by investigating:

- the average change in the frequency of moderate to severe VMS from the beginning of the trial (**baseline**) to Week 24
- the average change in the severity of moderate to severe VMS from baseline to Week 24

Full details of the **treatment-policy estimand** and other statistical methods are presented in **Section B.2.4** of the **main Company Submission**.

Results from the DAYLIGHT trial

Table 3 shows the key results from the DAYLIGHT trial after 24 Weeks of treatment with fezolinetant or placebo. These results show that treatment with fezolinetant was linked with a

greater reduction compared with placebo in the frequency and severity of moderate or severe events of VMS, in the studied participants.

In both sets of results, a "**p value**" of less than 0.001 was also reported. The p value describes the likelihood of any trial results being a result of chance. For trial results to be scientifically accepted, a p value less than 0.05 is usually expected, i.e., there would only be a 5% chance that the results were observed if fezolinetant had no effect. With a p value less than 0.001, it can be concluded that the results from the DAYLIGHT trial had a less than 0.1% of being observed if fezolinetant had no effect i.e., they were statistically significant.

Further effectiveness results for the DAYLIGHT trial, and the results of the analyses of the SKYLIGHT 1 and SKYLIGHT 2 trials, are presented in **Section B.2.6** of **Document B** of the **main Company submission**.

Table 3: Key effectiveness results from the DAYLIGHT trial after 24 weeks

Parameter	Placebo	Fezolinetant 45 mg	p value
Average change in the daily **frequency** of moderate to severe VMS from baseline to Week 24	−6.09	−8.15	<0.001
Average change in the **severity** of moderate to severe VMS from baseline to Week 24	−0.54	−0.99	<0.001

Abbreviations: VMS: vasomotor symptoms.

Indirect treatment comparison

As fezolinetant specifically targets VMS and has fewer side effects than other non-hormonal therapies described in **Section 2c**, UK clinical experts believe fezolinetant would be used as an initial treatment option for menopausal people who are deemed unsuitable for HRT for medical reasons. The main comparator to fezolinetant is expected to be no active treatment, which is represented by the placebo arm of DAYLIGHT. Because of this, the trial data from DAYLIGHT is sufficient to inform the economic analysis and no further analyses are needed.

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.

How was QoL measured?

A number of questionnaires were used to measure the impact of fezolinetant on patient QoL in the DAYLIGHT trial:

- **Patient-Reported Outcomes Measurement Information System Sleep Disturbance – Short Form 8b** (**PROMIS SD SF 8b**) questionnaire: sleep disturbance and deprivation are key complaints of participants with VMS. The PROMIS SD SF 8b questionnaire was used to assess how well fezolinetant managed to reduce the perceptions of being sleep-deprived in the menopausal people included in the study, and compared this against those who received placebo

- **Patient-Reported Outcomes Measurement Information System Sleep-Related Impairment – Short Form 8a** (**PROMIS SRI SF 8a**) questionnaire: similar to above, the PROMIS SRI SF 8a questionnaire was used to assess how well fezolinetant managed to reduce sleep impairments in the menopausal people included in the study, and compared this against those who received placebo

- **Menopause Quality of Life** questionnaire: the MENQOL questionnaire assesses how well fezolinetant improves the quality of life of menopausal people included in the study, and compared this against those who received placebo

- The **EuroQoL 5-dimensions, 5-levels** (**EQ-5D-5L**) and EQ-5D-5L **visual analogue scales** (**VAS**) questionnaires: the EQ-5D questionnaires are used to assess the general health-related QoL of participants. In the DAYLIGHT the EQ-5D questionnaire was used to quantify the impact of fezolinetant and placebo on patient QoL. This questionnaire assessed topics such as mobility, self-care, usual activities, pain and discomfort and anxiety and depression

- The **Work Productivity and Activity Impairment questionnaire for VMS** (**WPAI-VMS**): to assess the impact of fezolinetant on a participant's ability to continue working and living a productive life, the WPAI-VMS questionnaire was used in the DAYLIGHT trial. Among others, this questionnaire assessed the domains of productivity loss and activity impairment

Participants in the DAYLIGHT trial completed these questionnaires at baseline, and at each subsequent visit until they had received fezolinetant or placebo for 24 weeks. Comparing the questionnaire scores at the start and the end of the trials showed whether participants thought their physical functioning and quality of life had improved.

Impact of fezolinetant on participants' QoL

Compared with the baseline reading at the start of the trial, menopausal people who received fezolinetant reported a better score for the sleep disturbance domain of the PROMIS SD SF 8b questionnaire after 24 weeks of treatment, compared with those treated with placebo. This difference was statistically significant.

Participants in the DAYLIGHT trial also demonstrated an improvement in their health-related QoL, as measured by the EQ-5D questionnaires, regardless of whether they received fezolinetant or placebo. However, the improvement to QoL was nearly twice as great in the participants who were treated with fezolinetant.

Finally, participants who were treated with fezolinetant were also able to stay more productive and active, than those who were treated with placebo, as indicated by the results of the WPAI-VMS questionnaires.

Further QoL results for the DAYLIGHT trial, and the results of the QoL analyses of the SKYLIGHT 1 and SKYLIGHT 2 trials, are presented in **Section B.2.6** of **Document B** of the **main Company submission**.

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.

Like all medicines, this medicine can cause side effects, although not everybody gets them. As mentioned earlier, the DAYLIGHT trial was also designed to capture the safety profile of fezolinetant when used to treat moderate to severe VMS. Fezolinetant was generally well tolerated by those who received it, and overall had a similar safety profile to placebo. A summary of the most common side effects experienced by participants in the DAYLIGHT trial are given in **Table 4**, below.

Table 4: Summary of the most common side effects experienced by participants during the DAYLIGHT trial

Side effect	Placebo	Fezolinetant 45 mg
Participants experiencing any side effect	**61.1%**	**65.0%**
COVID-19	12.8%	13.3%
Headache	9.3%	8.8%
Fatigue	0.4%	5.8%
Cold	4.9%	4.0%
Nausea	1.8%	2.7%
Diarrhoea	1.3%	2.7%
Upper respiratory tract infection	1.8%	2.7%
Flu	2.7%	2.7%
Dizziness	2.2%	2.7%
Sleeping difficulties	0.4%	2.7%
Anxiety	0.9%	2.7%
Urinary tract infection	2.2%	2.2%
Bronchitis	1.3%	2.2%
Weight increased	0.4%	2.2%
Throat pain	0.9%	2.2%

The DAYLIGHT trial was conducted during the COVID-19 pandemic. As expected, COVID-19 infections were therefore reported as the most common side effect observed across both groups of participants. However, the investigators in the DAYLIGHT trial did not consider any of these cases of COVID-19 to be related to the study treatments.

Other common side effects reported by participants during the 24-week treatment period included headaches, fatigue, the common cold. In the vast majority of cases, these side effects were not considered a serious concern, and could be treated by following the advice of a doctor. Less than 1 in 20 participants had to stop taking fezolinetant due to the severity of their symptoms, with slightly more participants withdrawing from the study if they received the placebo. The proportion of participants who experienced a more serious side effect or stopped their treatment because of side effects during the DAYLIGHT trial is shown in **Table 5**.

Table 5: Summary of serious side effects experienced and treatment discontinuation during the DAYLIGHT trial

	Placebo	Fezolinetant 45 mg
Serious side effects	3.5%	4.4%
Side effects leading to the discontinuation of the study treatment	6.2%	4.9%

A more detailed description of the side effects observed during the DAYLIGHT trial, and the SKYLIGHT 1 and SKYLIGHT 2 trials, are presented in **Section B.2.6** of **Document B** of the main Company submission. The long-term safety of fezolinetant after 52 weeks of continuous treatment, was assessed in the SKYLIGHT 4 study. The results from this clinical trial are presented in **Appendix F** of the **Appendices to Document B**. Broadly, the results from the SKYLIGHT 4 trial are supportive of the results of the DAYLIGHT trial and indicate the ability of fezolinetant to reduce VMS symptom frequency and severity, with a low adverse risk to the health of those taking it.

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

The key benefits of fezolinetant to menopausal people with moderate to severe VMS include:

Reduced severity and frequency of moderate or severe VMS events

- Fezolinetant significantly reduces the frequency of experiencing burdensome hot flushes and/or night sweats, compared to placebo
- The reductions in frequency were paired with reductions in severity. Compared to those who received placebo, any episodes of VMS in menopausal people who received fezolinetant were reported as being less intense

Manageable safety profile

- Fezolinetant is generally well tolerated
- The side effects are likely to be familiar and readily managed by the healthcare community

Positive impact on QoL and productivity

- Fezolinetant significantly reduces sleep deprivation compared to placebo
- Fezolinetant leads to greater improvements in quality of life and physical functioning compared with placebo
- People who receive fezolinetant are more productive and are far less impaired in their ability to work and perform daily activities, than those who receive placebo

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

While fezolinetant is usually well tolerated and helps reduce the symptom burden of menopausal people living with VMS, there are some things which you may want to consider before starting treatment.

Efficacy

Fezolinetant may not work for everyone, and some menopausal people might not experience any improvement in disease symptoms. Menopausal people for whom fezolinetant does not work may still experience side effects, which are detailed further below.

Side effects

Like all medicines, some menopausal people may experience side effects while they are taking fezolinetant. The most common side effects experienced in the DAYLIGHT trial include headache, fatigue nausea and the common cold. These are usually manageable, and most menopausal people do not need to stop treatment because of side effects.

Administration

Fezolinetant should be taken every day for as long as your symptoms continue, however your doctor will have regular discussions with you to check how well it is managing your symptoms.

3j) 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.

Healthcare administrators need to get the best value from their limited budgets. To do this, they want to know whether a new medicine provides 'good value for money' compared with existing alternative courses of action. They will look at the costs of the new medicine and how the health of a person is likely to improve if they use it. The pharmaceutical company that develops the medicines provides this information to healthcare administrators using a **health economic model**. The pharmaceutical company uses the health economic model to perform an analysis, which compares the costs and benefits of the new medicine (fezolinetant) with the relevant comparator(s), which in this submission is no active treatment.

How the model reflects VMS associated with menopause

The economic model was designed to best capture the key features of VMS experienced by people as they transition through the menopause, and beyond. The goal of the model was to compare the costs and benefit of menopausal people treated with fezolinetant or no active treatment.

To do this a **Markov cohort model** structure was chosen. This type of model was considered most suitable to model VMS after the company had discussions with health-economic experts and clinical experts experienced in treating menopausal people with VMS. If fezolinetant demonstrates an improvement to health and quality of life over no active treatment, at a reasonable cost, it is considered a good use of NHS resources.

Modelling how much fezolinetant improves the health of menopausal people with moderate or severe VMS

In a Markov model, the symptoms and experiences of people as they progress through the disease are categorised into several "**health states**", which reflect the different stages of disease a patient, in this case menopausal people with VMS, may go through during the menopausal transition.

To categorise health states for VMS, the company referred to the number of VMS experienced by menopausal people in the DAYLIGHT trial (see **Section B.3.3** of the **main Company submission**

for details). The reduction in the number of VMS observed in the DAYLIGHT trial was a key result considered in the model along with SKYLIGHT 1 and SKYLIGHT 2.

The results of the DAYLIGHT trial cover a total of 24 weeks and SKYLIGHT 1 and SKYLIGHT 2 for 52 weeks. However, the economic model simulates menopausal people for 10 years, to adequately capture all costs and benefits of treatment with fezolinetant. This is a much longer period of time than the length of all three trials. As such, assumptions had to be made on how best to use the data available in the model.

Modelling how much fezolinetant improves quality of life

Participants from the DAYLIGHT trial were asked about their QoL at the start of the trial and whilst they were on treatment. Their responses were collected using EQ-5D-5L. This was then used to inform the model.

Modelling how the costs of the comparator differ with the new treatment

Various different costs are included in the company's model for the different treatments for VMS. These costs include:

- The cost of the medicine itself
- The costs associated with visits to your GP, a specialist, or any potential visits to an emergency room

In the company's economic model, fezolinetant is more costly than no active treatment; however, treatment with fezolinetant will yield greater health benefits, in terms of **quality adjusted life years** (QALYs) for menopausal people with moderate to severe VMS. The company's model shows that overall resource use will be lower for patients treated with fezolinetant, as more patients treated with fezolinetant will move to lower VMS frequency health states in the model, which are associated with lower resource use due to needing fewer visits to GPs or specialists.

Uncertainty

As is common practice in economic analyses, there are various assumptions that were made in the model. Information on these assumptions can be found in **Document B, Section B.3.9.2**. These assumptions were varied in tests which were conducted to see the impact on the cost-effectiveness results. The smaller the difference in the results before and after changing an assumption, the more reassured we are about the robustness of the model. Further details of these tests can be found in **Document B, Section B.3.11**.

One of the key uncertainties was the way in which no active treatment was modelled in the company's model. Because placebo does not reflect a treatment that would be received in clinical practice, the effects of placebo observed in the trial for fezolinetant are not expected to reflect clinical practice. Therefore, assumptions had to be made about the natural history of VMS for menopausal people receiving no active treatment in the company's model. Two sources of data were used: placebo arm of the DAYLIGHT trial up to Week 12 of the company's model, after which natural history data sourced by Astellas from UK clinical experts were applied for the remainder of the model. Therefore, the use of placebo data is likely to undervalue the benefits of fezolinetant compared with no active treatment in the model.

Benefits of fezolinetant not captured in the economic analysis

As described in **Section 2a**, VMS can cause menopausal people to miss work or be less productive, negatively impacting society. Fezolinetant has been shown to improve work productivity and reduce activity impairments. This improvement is not considered in the economic analyses.

Additionally, experts have highlighted that access to specialist menopausal clinics and appointment wait times can vary depending on where you live. If fezolinetant is not available for prescription by GPs, this may add pressures to specialist clinics, making regional differences and wait times worse. Many of these specialist clinics are run by GPs with a special interest in menopause, and so fezolinetant may also not be available from specialist clinics if availability is restricted to secondary care.

As such, important impacts of fezolinetant to work productivity and equality are not fully captured in the economic model.

3k) 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)

Fezolinetant is a first-in-class nonhormonal therapy that is proven to significantly reduce both the frequency and severity of moderate to severe VMS, which translates to improved outcomes in terms of sleep, quality of life, daily functioning and mental health.

The way fezolinetant works means that it specifically targets VMS. Therefore, the use of fezolinetant in menopausal people with moderate to severe VMS and for whom HRT is deemed unsuitable for medical reasons would represent a step change in the treatment care pathway, given the current unmet need for non-hormonal alternatives to HRT and the current lack of robust, licensed and evidenced-based therapies for treating the VMS.

3l) 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

There is a substantial unmet need for treatment options for menopausal people with VMS, and historically there has been a lack of innovation in this area of women's health, particularly for non-hormonal treatment options. New non-hormonal treatment options would therefore be welcome by this group of people.

VMS variations between ethnicities

There are differences in the number of cases (**prevalence**) and severity of VMS between menopausal people of different ethnicities, which should be considered. A higher prevalence and longer duration of VMS has been reported in women of African and Caribbean descent, and symptoms are also more severe.[42] These women are also more likely to report problems with sleep disturbance. Findings from the SWAN study further support that VMS are more prevalent in black and Hispanic women than other ethnicities.[43, 44]

Socioeconomic differences

Differences in VMS severity between socioeconomic levels also exist. Women in a household where the main earner works in semi-skilled or unskilled manual work have been reported to be more likely to have difficulty with physical and psychological postmenopausal symptoms than those where the main earner works in a managerial role.[26] Severe symptoms were also reported to be worse for working class women, with 44% of women describing three or more symptoms as 'very difficult'.[26]

Location-dependent differences

There are also differences in access to specialist menopause clinics and appointment waiting lists depending on where patients live.[38] Such disparities in access would be exacerbated should fezolinetant not be available for prescription in a primary care setting.[45] Additionally, many specialist menopause clinics are in reality run by GPs with a special interest in menopause. If the availability of fezolinetant is restricted to secondary care only, many of these specialist menopause clinics will not be able to provide their patients access to fezolinetant.

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.

Further information on VMS and menopause:

- British Menopause Society: https://thebms.org.uk/
- NHS website: https://www.nhs.uk/conditions/menopause/
- NICE guidelines: https://www.nice.org.uk/guidance/ng23
- NICE guidelines: https://www.nice.org.uk/guidance/ng101/

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

This glossary explains certain terms in this summary of information for patients. At times, an explanation for a term might mean you need to read other terms to understand the original terms.

Amenorrhoea. A complete stop in your menstrual cycle (period).

Baseline. The initial set of measurements or data collected from participants in a clinical trial before they begin the treatment or intervention being studied. These baseline measurements serve as a reference point, allowing researchers to compare changes over time and determine the effects of the treatment.

Black cohosh. A plant used by some menopausal people as a non-pharmaceutical treatment for VMS.

Cognitive behavioural therapy (CBT). A type of talking therapy aimed challenge negative thoughts and develop healthier ways of thinking and acting.

Efficacy. The ability of a drug to produce the desired beneficial effect on your disease or illness in a clinical trial.

European Medicines Agency (**EMA).** The regulatory agency of the European Union responsible for the scientific evaluation, supervision, and safety monitoring of medicines in the EU. The EMA ensures that all medicines available on the EU market are safe, effective, and of high quality.

Follicle-stimulating hormone (FSH). A hormone which controls the menstrual cycle and egg production.

Frequency. How often something happens.

Health economic model. A way to predict the costs and effects of a technology over time or in patient groups not covered in a clinical trial.

Health state. A description of a person's health.

Hormones. Chemical substances that carry messages within the body to help coordinate different bodily functions.

Hot flushes. Sudden feelings of intense heat that can make you feel very warm and cause sweating.

Hormone replacement therapy (HRT). A treatment used to relieve symptoms of menopause by replacing the hormones which are no longer produced by the body.

HRT-averse. Menopausal people who prefer not to use HRT.

HRT-caution. Menopausal people advised by their doctor against taking HRT due to the potential risks.

HRT-contraindicated. Menopausal people who have medical conditions that make HRT unsafe for them.

HRT-stoppers. Menopausal people who have tried HRT and found it not to be effective or to cause unwanted side effects.

Indirect treatment comparison. An analysis that compares medicines that have not been compared directly in a clinical trial.

Insomnia. Trouble falling asleep, staying asleep or getting good-quality sleep.

Isoflavones. A chemical found in high concentrations in beans such as soyabeans.

Licensed: Treatments that have been officially approved by the MHRA for use in treating specific medical conditions.

Luteinising hormone (LH). A hormone which controls the menstrual cycle and triggers the release of an egg from the ovary.

Marketing authorisation (MA). The legal approval by a regulatory body that allows a medicine to be given to patients in a particular country.

Markov cohort model. A type of health economic model.

Medicines and Healthcare products Regulatory Agency (MHRA). The regulatory body that evaluates, approves and supervises medicines in the UK.

Menopause. When you stop having menstrual periods.

Night sweats. Intense sweating which happens during sleep causing you to wake up.

Oestradiol. A type of oestrogen which plays a key role in regulating the menstrual cycle.

Oestrogen. A hormone primarily produced in the ovaries which plays a crucial role in regulating the menstrual cycle.

P value. A number which reflects how likely results occurred due to chance or a real effect.

Palpitations. The sensation that your heart is racing.

Patient reported outcomes (PROs). Measures of a person's quality of life.

Patient-based evidence. Information gathered directly from patients about their experiences.

Perimenopause. The phase leading up to menopause where the body starts to transition and hormone levels begin to decrease.

Pharmacological. Therapies involving medications or drugs used to manage or treat medical conditions. These treatments work by affecting biological processes in the body to alleviate symptoms.

Placebo. A placebo is a substance or treatment with no active therapeutic effect. It is often used in clinical trials as a control to test the effectiveness of new treatments. People taking a placebo may experience changes in their condition simply because they believe they are receiving a real treatment, an effect known as the placebo effect.

Postmenopausal. The time following 12 months without a menstrual period.

Premenopause. The time before a person starts experiencing perimenopause and menopause, during which their menstrual cycle is regular and hormone levels are stable.

Prevalence. The percentage of a population that has a particular condition at a given time.

Progestogens. Hormones that play a key role in regulating the menstrual cycle.

Quality adjusted life years (QALYs). A measure of the state of health of a person, where the length of life is adjusted to reflect the quality of life. One quality-adjusted life year (QALY) is equal to 1 year of life in perfect health. QALYs are calculated by estimating the years of life remaining for a patient following a particular treatment or intervention and weighting each year with a quality-of-life score (on a 0 to 1 scale). It is often measured in terms of the person’s ability to carry out the activities of daily life, and freedom from pain and mental disturbance.

Severity. How serious or intense a symptom is.

Statistically significant. Results which are unlikely to occur due to chance.

Tachycardia. A heart rate of 100 beats per minute or higher.

Vasomotor symptoms (VMS). Hot flushes and night sweats.

Venous thromboembolism (VTE). A condition where blood clots form in the veins.

Vertigo. A condition where you feel like you or your surroundings are spinning or moving. This can make you dizzy, nauseous or unsteady.

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:

1. EverydayHealth. What Are Vasomotor Symptoms of Menopause? Symptoms, Causes, Diagnosis, Treatment, and Prevention. Available at: https://www.everydayhealth.com/menopause/vasomotor-symptoms/guide/ [Accessed: September 2024], 2022.
2. European Medicines Agency (EMA). GUIDELINE ON CLINICAL INVESTIGATION OF MEDICINAL PRODUCTS FOR HORMONE REPLACEMENT THERAPY OF OESTROGEN DEFICIENCY SYMPTOMS IN POSTMENOPAUSAL WOMEN. Available at: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-hormone-replacement-therapy-oestrogen-deficiency_en.pdf [Accessed: September 2024].
3. Cleveland Clinic. Pituitary Gland: What It Is, function & Anatomy. Available at: https://my.clevelandclinic.org/health/body/21459-pituitary-gland [Accessed: September 2024], 2022.
4. National Institute for Health and Care Excellence (NICE). Menopause: diagnosis and management [NG23]. Available at: https://www.nice.org.uk/guidance/NG23 [Accessed September 2024].
5. Clinical Knowledge Summaries (CKS). Menopause Scenario: Managing women with menopause, perimenopause, or premature ovarian insufficiency – Non-hormonal Treatments. Available at: https://cks.nice.org.uk/topics/menopause/management/management-of-menopause-perimenopause-or-premature-ovarian-insufficiency/#non-hormonal-treatments [Accessed: September 2024].
6. Thurston RC, Joffe H. Vasomotor symptoms and menopause: findings from the Study of Women's Health across the Nation. Obstet Gynecol Clin North Am 2011;38:489-501.
7. Whiteley J, DiBonaventura M, Wagner JS, et al. The impact of menopausal symptoms on quality of life, productivity, and economic outcomes. J Womens Health (Larchmt) 2013;22:983-90.
8. Monteleone P, Mascagni G, Giannini A, et al. Symptoms of menopause - global prevalence, physiology and implications. Nat Rev Endocrinol 2018;14:199-215.
9. Dennerstein L, Dudley EC, Hopper JL, et al. A prospective population-based study of menopausal symptoms. Obstetrics & Gynecology 2000;96:351-358.
10. Hickey M, Szabo RA, Hunter MS. Non-hormonal treatments for menopausal symptoms. Bmj 2017;359:j5101.
11. Archer DF, Sturdee DW, Baber R, et al. Menopausal hot flushes and night sweats: where are we now? Climacteric 2011;14:515-28.
12. Thurston RC, Bromberger JT, Joffe H, et al. Beyond frequency: who is most bothered by vasomotor symptoms? Menopause 2008;15:841-7.
13. Hunter MS, Gentry-Maharaj A, Ryan A, et al. Prevalence, frequency and problem rating of hot flushes persist in older postmenopausal women: impact of age, body mass index, hysterectomy, hormone therapy use, lifestyle and mood in a cross-sectional cohort study of 10,418 British women aged 54-65. BJOG 2012;119:40-50.
14. Gold EB, Colvin A, Avis N, et al. Longitudinal analysis of the association between vasomotor symptoms and race/ethnicity across the menopausal transition: study of women's health across the nation. American journal of public health 2006;96:1226-1235.

15. Tepper PG, Brooks MM, Randolph Jr JF, et al. Characterizing the trajectories of vasomotor symptoms across the menopausal transition. Menopause (New York, NY) 2016;23:1067.

16. Deecher DC, Dorries K. Understanding the pathophysiology of vasomotor symptoms (hot flushes and night sweats) that occur in perimenopause, menopause, and postmenopause life stages. Arch Womens Ment Health 2007;10:247-57.

17. Freeman EW, Sammel MD, Sanders RJ. Risk of long-term hot flashes after natural menopause: evidence from the Penn Ovarian Aging Study cohort. Menopause 2014;21:924-32.

18. Avis NE, Crawford SL, Greendale G, et al. Duration of menopausal vasomotor symptoms over the menopause transition. JAMA Intern Med 2015;175:531-9.

19. Kiran A, Schultz NM, Siddiqui E, et al. Epidemiology and treatment patterns of UK women diagnosed with vasomotor symptoms: Findings from the Clinical Practice Research Datalink GOLD database. Maturitas 2022;164:1-8.

20. Nappi RE SE, Todorova L, Rea C, Gemmen E, Schultz NM. Prevalence and quality-of-life burden of vasomotor symptoms associated with menopause: A European cross-sectional survey. Maturitas 2023;167.

21. Kumari M, Stafford M, Marmot M. The menopausal transition was associated in a prospective study with decreased health functioning in women who report menopausal symptoms. J Clin Epidemiol 2005;58:719-27.

22. Makara-Studzinska MT, Krys-Noszczyk KM, Jakiel G. Epidemiology of the symptoms of menopause - an intercontinental review. Prz Menopauzalny 2014;13:203-11.

23. Baker FC, de Zambotti M, Colrain IM, et al. Sleep problems during the menopausal transition: prevalence, impact, and management challenges. Nat Sci Sleep 2018;10:73-95.

24. Berecki-Gisolf J, Begum N, Dobson AJ. Symptoms reported by women in midlife: menopausal transition or aging? Menopause 2009;16:1021-9.

25. Freeman EW, Sammel MD, Gross SA, et al. Poor sleep in relation to natural menopause: a population-based 14-year follow-up of midlife women. Menopause 2015;22:719-26.

26. Fawcett. Menopause and the Workplace. Available at: https://www.fawcettsociety.org.uk/menopauseandtheworkplace [Accessed: September 2024].

27. Soares CN. Can depression be a menopause-associated risk? BMC Med 2010;8:79.

28. Natari RB, Clavarino AM, McGuire TM, et al. The bidirectional relationship between vasomotor symptoms and depression across the menopausal transition: a systematic review of longitudinal studies. Menopause 2018;25:109-120.

29. Reed SD, Ludman EJ, Newton KM, et al. Depressive symptoms and menopausal burden in the midlife. Maturitas 2009;62:306-10.

30. DiBonaventura Md, Chandran A, Hsu M-A, et al. Burden of vasomotor symptoms in France, Germany, Italy, Spain, and the United Kingdom. International Journal of Women's Health 2013:261-269.

31. Griffiths A, MacLennan SJ, Hassard J. Menopause and work: an electronic survey of employees' attitudes in the UK. Maturitas 2013;76:155-9.

32. Jack G, Riach K, Bariola E, et al. Menopause in the workplace: What employers should be doing. Maturitas 2016;85:88-95.

33. Kopenhager T, Guidozzi F. Working women and the menopause. Climacteric 2015;18:372-5.

34. Hardy C, Thorne E, Griffiths A, et al. Work outcomes in midlife women: the impact of menopause, work stress and working environment. Women's Midlife Health 2018;4.

35. MedlinePlus. Follicle-Stimulating Hormone (FSH) Levels Test. Available at: https://medlineplus.gov/lab-tests/follicle-stimulating-hormone-fsh-levels-test/ [Accessed: September 2024], 2023.

36. National Institute for Health and Care Excellence (NICE). Early and locally advanced breast cancer: diagnosis and management [NG101]. Available at: https://www.nice.org.uk/guidance/NG101 [Accessed September 2024].

37. National Institute for Health and Care Excellence (NICE). Menopause: diagnosis and management [GID-NG10241] – In Development. Available at: https://www.nice.org.uk/guidance/indevelopment/gid-ng10241 [Accessed: September 2024].

38. Astellas Data on File. UK Individual Expert Validation Meetings Report. 2024.

39. Astellas Data on File. Veoza 45 mg film-coated tablets. Draft Summary of Product Characteristics (SmPC). 2023.

40. Samarasiri D, Rajanayake I, Wattegedara P, et al. Women's Experiences on Menopause: A Phenomenological Study. OUSL Journal 2017;12:43.

41. English M, Stoykova B, Slota C, et al. Qualitative study: burden of menopause-associated vasomotor symptoms (VMS) and validation of PROMIS Sleep Disturbance and Sleep-Related Impairment measures for assessment of VMS impact on sleep. J Patient Rep Outcomes 2021;5:37.

42. British Menopause Society (BMS). Menopause in ethnic minority women. Available at: https://thebms.org.uk/wp-content/uploads/2023/07/20-BMS-TfC-Menopause-in-ethnic-minority-women-JULY2023-B.pdf [Accessed: September 2024].

43. Green R, Santoro N. Menopausal symptoms and ethnicity: the Study of Women's Health Across the Nation. Women's Health 2009;5:127-133.

44. Harlow SD, Burnett-Bowie S-AM, Greendale GA, et al. Disparities in reproductive aging and midlife health between Black and White women: the Study of Women's Health Across the Nation (SWAN). Women's midlife health 2022;8:1-17.

45. National Institute for Health and Care Excellence (NICE). Fezolinetant for treating vasomotor symptoms associated with the menopause (ID5071). Response to stakeholder organisation comments on the draft remit and draft scope. Available at: https://www.nice.org.uk/guidance/indevelopment/gid-ta11058 [Accessed 5 August 2024].

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal

Fezolinetant for treating vasomotor symptoms associated with the menopause [ID5071]

Clarification questions

November 2024

File name	Version	Contains confidential information	Date
ID5071_Fezolinetant EAG CQs_2.0 [REDACTED].docx	2.0	Yes	22nd November 2024

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

Decision Problem

A1. PRIORITY: In the decision problem, fezolinetant is only compared to no treatment in people for whom HRT is not considered suitable. Clinical advice to the EAG is that although only clonidine is licensed, all the non-hormonal pharmacological treatments listed in the NICE scope are currently and commonly prescribed in UK primary care for menopausal people for whom HRT is not considered suitable. Therefore, the EAG considers these to be relevant comparators. Please consider including non-hormonal pharmacological comparators in the decision problem and provide the relevant clinical effectiveness and cost-effectiveness evidence comparing fezolinetant to all non-hormonal pharmacological comparators listed in the NICE scope.

Relevant comparators

As noted in Section B.1.1 of Document B, the decision to select no active treatment as the sole relevant comparator to fezolinetant was primarily informed by a total of five UK clinical experts with specialty in menopause consulted by Astellas for this appraisal (n=3)[1] and during early engagement with NICE at the Office for Market Access (OMA) meeting (n=2). Although other non-hormonal alternatives, such as those included in the final scope, exist in NHS clinical practice, UK clinical experts emphasised that these treatments, which are mainly prescribed off-label (except clonidine), are suboptimal in terms of their effectiveness and unpleasant side effects.[1] As such, these treatments are not routinely prescribed in NHS clinical practice, and

there is an unmet need for treatments that target vasomotor-predominant menopausal symptoms. These views are further echoed by the British Menopause Society (BMS) consulted by NICE during the draft scope consultation for this appraisal.[2] The BMS emphasised the importance of highlighting the shortcomings of currently reimbursed non-hormonal treatments in terms of their effectiveness, side effects and lack of licence in this indication.[2] The BMS highlighted that many women still struggle to access effective non-hormonal treatments and so there remains an urgent need for this appraisal.[2] Given the unanimous feedback from the UK clinical expert community, Astellas' position is that no active treatment remains the sole relevant comparator to fezolinetant in this appraisal.

Exploratory network meta-analysis (NMA)[3]

The Company maintains that an NMA is not relevant to the anticipated positioning of fezolinetant. However, an exploratory NMA is presented here to generate comparative clinical evidence for fezolinetant versus non-hormonal alternatives.

The feasibility of an exploratory NMA was considered due to the lack of head-to-head randomised controlled trials (RCTs) for fezolinetant versus other non-hormonal alternatives. To be considered eligible for the NMA, studies identified in the clinical systematic literature review (SLR) had to have met the following criteria:

- **Population:** postmenopausal women with moderate to severe VMS who are unsuitable for hormone replacement therapy (HRT).
- **Intervention and comparators:** fezolinetant and non-hormonal alternatives (including selective serotonin reuptake inhibitors [SSRIs], serotonin noradrenaline reuptake inhibitors [SNRIs], clonidine and anticonvulsants).
- **Outcomes:** at least one of the outcomes of interest (change from baseline in vasomotor symptoms [VMS] frequency or severity) at 12 weeks or 24 weeks.
- Comparators must have a UK marketing authorisation (for any indication, not necessarily VMS) and used at a dose in line with UK clinical practice.

The feasibility assessment identified four (SKYLIGHT 1, SKYLIGHT 2, N30-003 and N30-004) out of the 14 studies included in the clinical SLR (in addition to DAYLIGHT, which was unpublished at the time of the SLR[1]), as eligible for inclusion in the exploratory NMA. Given the eligibility criteria for this network, the only comparison that was feasible was between fezolinetant and one SSRI (paroxetine 7.5 mg). Consistent with the feedback from the Guidance Development Group in the NICE clinical guidelines for menopause (NG23; see Appendix L.1.3.7), it is reasonable for paroxetine to represent SSRIs, given that treatments in the same class are expected to have similar efficacy.[4] As the other 10 studies identified in the clinical SLR did not meet the eligibility criteria for inclusion in the network, comparisons with other non-hormonal alternatives were not feasible.

All 5 studies included in the network evaluated change from baseline in frequency of moderate to severe VMS at Week 12; however, only DAYLIGHT reported on change from baseline in

[1] The DAYLIGHT primary manuscript, **Efficacy and safety of fezolinetant for moderate-severe vasomotor symptoms associated with menopause in individuals unsuitable for hormone therapy: phase 3b randomised controlled trial**, is expected to be published online by *BMJ* on November 13th 2024 (and print published in the November 16th 2024 issue of *BMJ*.

frequency of moderate to severe VMS at Week 24, and therefore no comparative analysis could be performed for this timepoint.

As presented in Table 1, the NMA results demonstrated that fezolinetant 45mg was associated with a statistically significant reduction in the frequency of moderate to severe VMS when compared with paroxetine 7.5 mg in the fixed effects models and a numerical (non-statistically significant) reduction in the random effects models.

Table 1: Pairwise treatment effect comparisons of fezolinetant 45mg versus paroxetine 7.5 mg for change from baseline in frequency of moderate to severe VMS at 12 weeks (fixed effects and random effects)

Comparator	Fezolinetant 45mg versus comparator			
	Fixed effects		Random effects	
	Treatment difference (mean)	95% CrI	Treatment difference (mean)	95% CrI
Paroxetine 7.5mg (oral)	[redacted]	[redacted]	[redacted]	[redacted]

Negative point estimates indicate better efficacy for fezolinetant 45mg. **Bolded** numbers indicate statistically significant results.
Abbreviations: CrI: credible interval.

While the studies included in the exploratory NMA were identified by a robust SLR, underwent thorough feasibility assessment to ensure similarity in trial and baseline characteristics, and had sufficient data to allow for the formation of a fully connected network, the exploratory NMA had limitations:

- The DAYLIGHT trial was the only study that had an explicit inclusion criterion for the HRT-unsuitable population. While SKYLIGHT 1 and SKYLIGHT 2 trials did not have this explicit inclusion criterion, results for the HRT-unsuitable subpopulation were available from both trials. All other trials did not provide enough information regarding the HRT suitability of the patients. That said, HRT unsuitability was not considered to be a treatment effect modifier in the feasibility assessment.
- Race or ethnicity was identified as a potential treatment-effect modifier, and there was evidence of an imbalance in this characteristic across studies. The percentage of white and black patients varies across the trials, with the proportion of black patients varying between 0% (DAYLIGHT) and 32.8% (N30-003). However, additional adjustments for race via meta-regression were not considered feasible given the sparsity of studies in the network.
- Lastly, the dose of paroxetine (7.5 mg) used in the identified studies were lower than the dose (10 mg) recommended for use in NHS clinical practice for VMS, which introduces some uncertainty in the comparison.[5]

Further details of the exploratory NMA are provided as data on file alongside this response for transparency and completeness.[3]

Cost-effectiveness results

The results of the NMA were used to generate transition probabilities for paroxetine which were incorporated in the model, the methodology of which is summarised in the Appendices below.

and topline cost-effectiveness results of the pairwise comparisons between fezolinetant and SSRIs (paroxetine) are presented in Table 2.

These results should be interpreted with caution because while it was feasible to use the exploratory NMA within the model, a singular treatment effect output had to be mapped to the transition probability structure that is considered in the model. The model considers four-by-four transition probabilities matrices and so mapping the singular NMA output to these 16-variable matrices presents a significant source of uncertainty. Additionally, while the derived transition probabilities consider a level of time dependency, the treatment effect from the NMA that informed these calculations were based on only 12 weeks of data and, as such, assumptions had to be made that the relative treatment effects are the same across all timepoints – an assumption that cannot be verified given the lack of efficacy data for included studies beyond Week 12 (except DAYLIGHT). Lastly, the model assumes that from Week 52 onwards, the Week 48–52 transition probabilities for paroxetine are suitable to be applied for the rest of the model time horizon, given the lack of efficacy beyond Week 52.

Table 2: Summary of cost-effectiveness results for fezolinetant versus paroxetine (SSRI)

Technologies	Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	
Fezolinetant 45mg				-
Paroxetine 7.5 mg (SSRI)	£1,366.40	0.000	0.074	£18,553.52

Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life years gained; QALYs: quality-adjusted life years; SSRI: selective serotonin reuptake inhibitor.

A2. PRIORITY: The decision problem population comprises menopausal people for whom HRT is unsuitable for medical reasons. HRT-unsuitability is defined in the company submission (CS) as people who have stopped HRT for any reason, or people who have contraindications or cautions against HRT. The list of contraindications and cautions for the DAYLIGHT trial is provided in Table 6 of the CS. Clinical advice to the EAG suggests that, of these contraindications or cautions, only a history of oestrogen receptor positive tumours would prevent the prescription of HRT in the NHS, and that the other contraindications relate to oral oestrogen and synthetic hormones, but are not relevant for body-identical formulations and natural hormones, e.g. oestrogen delivered through the skin. Please provide outcome data from DAYLIGHT, SKYLIGHT 1, SKYLIGHT 2 and SKYLIGHT 1&2 (pooled) for mean VMS frequency change from baseline, mean VMS severity change from baseline, sleep disturbance (PROMIS SD SF8b, PROMIS SRI SF 8a), psychological

effects (PHQ-4), quality of life / function (MENQUOL, ED-5D-5L, WPAI-VMS) and adverse effects in the following sub-populations:

a) **HRT stoppers and people with a history of oestrogen-receptor positive cancer only**

b) **HRT unsuitable i.e. excluding HRT-naïve but willing participants (only applies to SKYLIGHT 1 and 2). Please clearly define and justify the sample sizes ('N' or 'n') used and why they may differ across outcomes.**

HRT-unsuitable definition

In defining the HRT-unsuitable population within Section B.1.1 of Document B, Astellas relied extensively on clinical feedback from UK experts specialising in menopause across diverse healthcare settings and professions to capture real-world scenarios in which HRT may be deemed unsuitable. Notably, Astellas facilitated the establishment of an independent panel of 10 UK clinical experts in menopause, comprising six consultant gynaecologists and/or endocrinologists, a menopause specialist nurse, a general practitioner (GP) menopause specialist, a consultant oncologist, and a consultant pharmacist, to develop a consensus statement defining populations for whom fezolinetant would be suitable within the NHS when HRT is deemed unsuitable.[6]

In addition to the panel, feedback was sought from three UK clinical experts during individual meetings with Astellas[1] and from two UK clinical experts during early engagement with NICE at the OMA meeting. Across the 15 clinical experts consulted, there was unanimous agreement that the definitions of HRT-unsuitability used in the Company submission align closely with those used in UK clinical practice. Moreover, the Clinical Knowledge Summaries (CKS) for menopause published on the NICE website and the Summary of Product Characteristics (SmPC) for various HRT products clearly outline the list of medical conditions for which HRT is contraindicated or cautioned, which further supports the HRT-unsuitable definitions used in this appraisal.[7, 8] It is also worth noting that while it may be a relatively small number of criteria that explicitly prevent the prescription of HRT, in reality women may present with multiple comorbidities and risk factors; as such, a holistic approach by the treating physician is necessary to assess HRT-suitability or -unsuitability on an individual basis. Consequently, Astellas maintains that the definitions applied to HRT-contraindicated and HRT-caution populations are consistent with NHS clinical practice and are appropriate for the decision problem.

With respect to specific contraindications, our consultation process confirmed that while a history of oestrogen receptor-positive cancer is a primary factor in excluding HRT, additional contraindications and cautions remain clinically relevant when defining HRT unsuitability across various hormone delivery methods, including both oral and non-oral formulations. Thus, our definitions reflect a nuanced understanding of HRT suitability in diverse clinical scenarios.

HRT-unsuitable subpopulation in SKYLIGHT 1 and SKYLIGHT 2

As noted in Table 7 of Document B, the HRT-unsuitable subpopulation of SKYLIGHT 1 and SKYLIGHT 2 comprised women who were HRT-contraindicated, HRT-caution, HRT-stoppers or HRT-averse (mutually exclusive subgroups). As such, Astellas can confirm that the HRT-

unsuitable subpopulation of pooled SKYLIGHT 1 and SKYLIGHT 2 does not include women who are HRT-naïve/willing.

The outcome data and sample sizes for the HRT-unsuitable subpopulation can be found in the following sections of Document B and the Appendices:

- Section B.2.7.2.2 Subgroup analyses in the pooled HRT-unsuitable subpopulation
- Appendix E.2.1 SKYLIGHT 1 and SKYLIGHT 2 (pooled analysis; HRT-unsuitable subgroup)
    - E.2.1.1 Baseline characteristics
    - E.2.1.2 Co-primary efficacy endpoints (frequency and severity of moderate to severe VMS)
    - E.2.1.3 Key secondary endpoint (PROMIS SD SF 8b total score)
    - E.2.1.4 (Responder analysis of frequency of moderate to severe VMS)

Tables of the clinical study report (CSR) where results for the following outcomes can be found are detailed below and provided in the reference pack for completeness:

- PROMIS SRI SF 8a: SKYLIGHT 1 and 2 pooled analysis Table 9.3.5.2.2.3
- MENQOL total score: SKYLIGHT 1 and 2 pooled analysis Table 9.3.5.5.2.4
- EQ-5D-5L VAS score: SKYLIGHT 1 and 2 pooled analysis Table 9.3.5.8.2.4
- WPAI-VMS domain scores: SKYLIGHT 1 and 2 pooled analysis Table 9.3.5.7.2.4
- Adverse effects
    - Overview of treatment emergent adverse events (TEAEs) and death: SKYLIGHT 1 and 2 pooled analysis Table 9.6.1.1.6.1
    - Common TEAEs: SKYLIGHT 1 and 2 pooled analysis Table 9.6.1.2.6
    - Study drug-related TEAEs: SKYLIGHT 1 and 2 pooled analysis Table 9.6.1.3.6
    - Serious TEAEs: SKYLIGHT 1 and 2 pooled analysis Table 9.6.1.4.6
    - TEAEs leading to withdrawal of treatment: SKYLIGHT 1 and 2 pooled analysis Table 9.6.1.6.6

A3. PRIORITY. The company submission (p 14) states that NICE advised during the scoping consultation that 'menopausal people who are HRT-averse i.e., those who made an informed choice not to take HRT following a consultation on the benefits and risks of HRT, have been excluded from the targeted population relevant to the decision problem, as it is expected that HRT remains the relevant comparator for these people'. This is incorrect and was discussed during the decision problem meeting. People who are 'HRT-averse' would be categorised within 'People for whom HRT is not considered suitable' because the people themselves have deemed themselves not

suitable. Please include this population within your submission or provide additional reasons, and supporting evidence, for why this population is not included.

As noted during the EAG/NICE clarification call, Astellas acknowledges that the exclusion of menopausal people who are HRT-averse from the decision problem was a misunderstanding following on from the scoping consultation for this appraisal.

Astellas can confirm that the population of relevance to the decision problem has now been revised to include HRT-averse: Menopausal people with moderate to severe vasomotor-predominant symptoms for whom HRT is deemed unsuitable i.e., HRT-contraindicated, HRT-caution, HRT-stoppers and HRT-averse.

A4: In the DAYLIGHT trial, there were 27 people in the fezolinetant group and 23 people in the placebo group who had contraindications to HRT. How many of these had a history of oestrogen-receptor positive breast cancer?

The number of patients in DAYLIGHT who had a history of oestrogen-receptor positive breast cancer are presented in Table 3.

Table 3: Number of patients with contraindication to HRT: history of known or suspected oestrogen dependent tumours in DAYLIGHT

Trial	Fezolinetant 45 mg	Placebo	Total
DAYLIGHT			
Contraindication to HRT, N	[redacted]	[redacted]	[redacted]
History of known or suspected oestrogen dependent tumours, n (%)	[redacted]	[redacted]	[redacted]

Abbreviations: HRT: hormone replacement therapy.

A5: In the pooled SKYLIGHT trials, there were 57 people in the fezolinetant 45mg group and 54 people in the placebo group who had contraindications to HRT. For each of these trials separately, please indicate how many people in the fezolinetant 45mg and placebo groups had a history of oestrogen-receptor positive breast cancer.

The number of patients in SKYLIGHT 1, SKYLIGHT 2 and the pooled SKYLIGHT 1 and SKYLIGHT 2 analysis who had a history of oestrogen-receptor positive breast cancer are presented in Table 4.

Table 4: Number of patients with contraindication to HRT: history of known or suspected oestrogen dependent tumours in the SKYLIGHT trials

Trial	Fezolinetant 45 mg	Placebo	Total
SKYLIGHT 1			
Contraindication to HRT, N	[redacted]	[redacted]	[redacted]

History of known or suspected oestrogen dependent tumours, n (%)	[redacted]	[redacted]	[redacted]
SKYLIGHT 2			
Contraindication to HRT, N	[redacted]	[redacted]	[redacted]
History of known or suspected oestrogen dependent tumours, n (%)	[redacted]	[redacted]	[redacted]
Pooled SKYLIGHT 1 & SKYLIGHT 2			
Contraindication to HRT, N	[redacted]	[redacted]	[redacted]
History of known or suspected oestrogen dependent tumours, n (%)	[redacted]	[redacted]	[redacted]

Abbreviations: HRT: hormone replacement therapy.

A6. The decision problem population is described as 'menopausal people'. However, Table 4 in the CS mentions perimenopausal people as eligible for fezolinetant.

a) Please clarify whether this submission covers both perimenopausal and post-menopausal people, or just post-menopausal people.

The licensed indication for fezolinetant includes both perimenopausal and postmenopausal people with moderate to severe VMS. As such, Astellas does not intend to restrict the population covered in the decision problem to only postmenopausal people.

b) Please indicate whether both peri- and post-menopausal people are:

i. covered by fezolinetant's marketing authorisation,

As noted above, the licensed indication includes both perimenopausal and postmenopausal people with moderate to severe VMS. Throughout the regulatory process, extensive discussions with the European Medicines Agency (EMA), specifically with the Committee for Medicinal Products for Human Use (CHMP) at Day 120 and Day 180, provided clarity on the indication wording. During the Day 120 review, the CHMP initially considered whether the indication should be limited to postmenopausal women only. In response, Astellas presented substantial evidence from the published literature, demonstrating that the physiological mechanism underlying VMS is consistent across the menopause transition, from perimenopause to postmenopause, and is not exclusive to postmenopausal women.[9-20] This evidence supported the conclusion that there is no physiological rationale to limit the use of fezolinetant solely to postmenopausal people.

Given the shared pathophysiology and the comparable VMS burden experienced across these stages, it is reasonable to expect similar safety and efficacy outcomes for fezolinetant in perimenopausal women, comparable to those observed in postmenopausal populations. This broader indication is further supported by the European Menopause and Andropause Society (EMAS)[21] and International Menopause Society (IMS) guidelines[22], which recommend pharmacological treatment across all stages of the menopause transition. Additionally, NICE NG23 guidelines do not differentiate between menopause phases in their treatment recommendations, focusing instead on managing menopause symptoms irrespective of specific stages.[4] Although most regulatory clinical studies are traditionally conducted with

postmenopausal populations due to regulatory precedents,[23, 24] NICE and other guidelines do not make such distinctions in clinical practice recommendations.

Consequently, as noted in the European Public Assessment Report (EPAR), CHMP ultimately deemed the indication wording to be approvable, finding no major clinical efficacy or safety concerns in the postmenopausal population studied in the pivotal Phase 3 trials for fezolinetant (SKYLIGHT 1 and SKYLIGHT 2).[25] To reflect that the licensed indication encompasses both peri- and postmenopausal individuals while the studied population was postmenopausal, the term "postmenopausal" was introduced in Section 5.1 of the SmPC. Additionally, the indication wording in Section 4.1 explicitly references Section 5.1, where the studied clinical trial population is described.[26]

ii. represented in the RCT evidence and the systematic review.

RCT evidence

The Phase 3 clinical development programme for fezolinetant (DAYLIGHT, SKYLIGHT 1, SKYLIGHT 2, and SKYLIGHT 4) enrolled only postmenopausal women with VMS associated with menopause. This design choice was driven by regulatory requirements from the US Food and Drug Administration (FDA) for evaluating treatments for VMS associated with menopause, as outlined in the FDA's 2003 guidance on estrogen and estrogen/progestin drug products.[23] These guidelines influenced the selection criteria to focus on a homogeneous postmenopausal population to meet regulatory expectations. Consequently, clinical data for the perimenopausal population is not available from the phase 3 RCTs conducted for fezolinetant.

Systematic review

The clinical SLR conducted for this submission included studies containing perimenopausal people at the abstract stage to check if results were available for postmenopausal people in the full text stage. If the full text did not contain postmenopausal people, or did not report separate results for the postmenopausal subgroup, then the record was excluded. The population selection criterion was chosen to ensure included studies were similar to the patient populations enrolled in the DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 trials.

c) If perimenopausal people are also included, please comment on how the evidence presented for post-menopausal people may generalise to peri-menopausal people.

As noted above, the CHMP considered the licensed indication for fezolinetant to be approvable, capturing both perimenopausal and postmenopausal people, as no major efficacy or safety issues were identified in the postmenopausal population studied in the pivotal Phase 3 trials. Additionally, substantial evidence from the published literature demonstrates that the physiological mechanism underlying VMS remains consistent across the menopause transition, from perimenopause to postmenopause, and is not exclusive to postmenopausal women.[9-19] Consequently, there is no reason to expect that the efficacy and safety of fezolinetant in perimenopausal people would significantly differ from that observed in postmenopausal people.

d) Please stratify all baseline characteristics and outcome data for VMS frequency, VMS severity, sleep disturbance (PROMIS SD SF8b, PROMIS SRI SF 8a), psychological effects (PHQ-4), quality of life / function

(MENQUOL, ED-5D-5L, WPAI-VMS) and adverse effects in the DAYLIGHT and SKYLIGHT trials by peri- and post-menopause.

As discussed above, perimenopausal women were not enrolled in the phase 3 trials for fezolinetant; as such, it is not possible to stratify the data from these trials into peri- and post-menopause subgroups.

A7. Clinical advice to the EAG suggests that the decision problem should:

a) **include all of the non-hormonal pharmacological treatments in the NICE scope as comparators (see question A1);**

b) **re-define HRT-unsuitability as people who have stopped HRT for any reason, or people who have a history of oestrogen receptor positive tumours (see question A2);**

c) **clarify whether peri-menopausal people are also included (see question A5).**

Please re-define the decision problem according to the responses to questions A1, A2, A3 and A6 and present clinical and cost-effectiveness evidence relevant to the new decision problem.

In line with the responses to Question A1, A2, A3 and A6, the decision problem for this submission remains unchanged, except for the inclusion of HRT-averse in the target population for this submission i.e., menopausal people with moderate to severe vasomotor-predominant symptoms for whom HRT is deemed unsuitable:

- **HRT-contraindicated:** Menopausal people for whom HRT is contraindicated, including due to venous thromboembolism, cardiovascular disease, metabolic syndrome, severe hypertension, uncontrolled/complex diabetes mellitus, porphyria, etc.
- **HRT-caution:** Menopausal people for whom medical risk assessment of the specific caution has concluded that the risk of HRT outweighs the likely benefit.
- **HRT-stoppers:** Menopausal people who have previously received HRT but no longer take HRT.
- **HRT-averse:** Menopausal people who do not want to take hormones.

A8. PRIORITY. Please compare all the interventions included in the systematic review (i.e. those listed in Table 19 of the appendices) by conducting the following network meta-analyses. Please provide the code and all input data

for all NMAs so that all analyses can be reproduced. Please also present the baseline data and results data for all trials in Word tables.

a) Please analyse the following outcomes reported at (or close to) 12 weeks and 24 weeks: i) change from baseline in mean VMS frequency and ii) change from baseline in mean VMS severity; appendix Table 18 indicates that these data are available for all but one of the studies.

Exploratory NMA

As detailed in the response to Question A1, an exploratory NMA was conducted to assess the efficacy of fezolinetant versus non-hormonal alternatives. Among the studies identified in the clinical SLR, only five were eligible for inclusion, restricting comparisons to fezolinetant and paroxetine (SSRI). Data availability allowed for an analysis of change from baseline in the mean frequency of moderate-to-severe VMS at Week 12 only; thus, a comparison at Week 24 was not feasible. The results of this analysis are presented in Table 1.

The exploratory NMA was conducted using the statistical software R and JAGS, with the JAGS code used for the NMA provided below.

Bayesian continuous outcome, fixed effects (normal distribution)

```
model{
  for(i in 1:ns) {
                    # indexes studies
    mu[i] ~ dnorm(0,.0001)
             # vague priors for all trial baselines
    for (k in 1:na[i]) {
                            # indexes arms
    varr[i,k] <- pow(se[i,k],2)
             # calculate variances
    prec[i,k] <- 1/varr[i,k]
                           # sets precisions
    y[i,k] ~ dnorm(theta[i,k],prec[i,k])
             # normal likelihood
    theta[i,k] <- mu[i] + d[t[i,k]] - d[t[i,1]]
                    # model for linear predictor
    # deviance contribution
    dev[i,k] <- (y[i,k]-theta[i,k])*(y[i,k]-theta[i,k])*prec[i,k]
   }
                            # close arm loop
  resdev[i] <- sum(dev[i,1:na[i]])
                    # summed deviance contribution
  }
                                  # close study loop
  totresdev <- sum(resdev[])
                           # total residual deviance
  d[1]<-0
                            # effect is 0 for reference treatment
  for (k in 2:nt)[27]                                                    # vague
priors for treatment effects
}
                            # close treatment loop
```

Characteristics and outcomes for studies included in the clinical SLR

The baseline characteristics and outcomes for the 14 studies identified in the clinical SLR have been included in the reference pack alongside this document.

b) Please run the analyses for the following populations: i) people for whom HRT is unsuitable for medical reasons as defined in the submission (Table 6) and ii) patients for whom HRT is unsuitable for medical reasons because they are either HRT stoppers or they have a history of oestrogen sensitive tumours.

As noted in the response to Question A2, Astellas maintains that the definition of HRT-unsuitable individuals provided in the company submission accurately reflects NHS clinical practice, based on input from a diverse group of UK clinical experts. Consequently, the target population for this submission aligns with the definition outlined in response to Question A7.

In the exploratory NMA, it was not feasible to conduct a separate analysis for the HRT-unsuitable population across all included studies. This limitation arises because, apart from the DAYLIGHT and SKYLIGHT 1 and SKYLIGHT 2 trials, other included studies did not specify the HRT-suitability status of enrolled patients.

A9. The CS refers to a consensus statement which states that "shared management with secondary/tertiary... care is advised…[for]…people who have had oestrogen dependent tumours, e.g., breast, ovarian, uterine". This suggests that management of these people would not only be in primary care and may incur additional costs and healthcare resource use. Please comment on whether this shared management is standard practice, or an extra level of monitoring associated with taking fezolinetant for VMS.

The referral of complex cases to secondary or tertiary care does not represent an additional monitoring layer specific to fezolinetant but is instead already part of the routine care pathway in the NHS. This is supported by recent insights from a UK advisory board which indicated that referral to secondary or tertiary care is routine practice for managing complex cases of this nature.[28] The advisory board comprised a diverse group of experts (N=10), including two gynaecologists, two specialists/consultants in Sexual and Reproductive Health, four GP menopause specialists, and two menopause nurse specialists.[28]

Clinical efficacy

A10. The submission states that the fezolinetant clinical development programme comprises 19 completed studies (11 phase 1 studies, 4 phase 2 studies and 4 phase 3 studies) to date.

a) Please provide a list of the phase 2 and phase 3 completed studies, with clinical trial registry identifiers. Please also provide references, PDFs

and clinical study reports / protocols which have not already been provided.

A list of the phase 2 and phase 3 completed studies are presented in Table 5. Please note that there are now 5 completed phase 2 studies and 6 completed phase 3 studies within the clinical development programme for fezolinetant.

The CSRs, protocols and publications that have not yet been provided have been included within the reference pack supplied alongside the Company response.

Table 5: Completed phase 2 and phase 3 trials

Type of Study	Study Number (Location)	Clinical trial registry identifier	Publication	Protocol	CSR
Phase 2a 12-week proof-of-concept study in female participants with heavy menstrual bleeding due to uterine fibroids	ESN364-UF-02 (Western Europe)	EudraCT No.: 2014-004425-41	No	Yes (included in CSR)	Yes
Phase 2a 12-week proof-of-concept study in female participants with polycystic ovary syndrome	ESN364-PCO-201 (Western Europe, Georgia)	EudraCT No: 2014-004409-34	Fraser *et al.,* (2021)[29]	Yes (included in CSR)	Yes
Phase 2a 12-week proof-of-concept study in female participants with moderate to severe VMS	ESN364_HF_204 (Belgium)	NCT05419908	Depypere *et al.,* (2019)[30]	Yes (included in CSR)	Yes
Phase 2b 12-week randomized, placebo-controlled, double-blind, dose-ranging study in female participants with VMS associated with menopause	ESN364_HF_205 (US)	NCT03192176	Fraser *et al.,* (2020)[31]	Yes	Yes
Phase 2 randomized, placebo-controlled, 12-week, double-blind, phase 2 study to select the optimal dose(s) in Japanese women with mild to severe VMS **(STARLIGHT)**	2693-CL-0206 (Japan)	NCT05034042	Takamatsu *et al.,* (2024)[32]	Yes	Yes
Phase 3 randomized, placebo-controlled, 12-week double-blind study, followed by a non-controlled extension treatment period in participants suffering from moderate to severe VMS associated with menopause **(SKYLIGHT 1)**	2693-CL-0301 (US/Canada/Europe)	NCT04003155	Lederman *et al.,* (2023)[33]	Yes	Yes

Phase 3 randomized, placebo-controlled, 12-week double-blind study, followed by a non-controlled extension treatment period in participants suffering from moderate to severe VMS associated with menopause **(SKYLIGHT 2)**	2693-CL-0302 (US/Canada/Europe)	NCT04003142	Johnson *et al.*, (2023)[34]	Yes	Yes
Phase 3 randomized, placebo-controlled, double-blind long-term (52-week) safety study in participants suffering from VMS associated with menopause **(SKYLIGHT 4)**	2693-CL-0304 (US/Canada/Europe)	NCT04003389	Neal-Perry *et al.*, (2023)[35]	Yes	Yes
Phase 3 randomized, placebo-controlled, 12-week double-blind study followed by a non-controlled extension treatment period to evaluate the efficacy and safety of fezolinetant in women with moderate to severe VMS **(MOONLIGHT 1)**	2693-CL-0305 (Mainland China/South Korea/Taiwan)	NCT04234204	Ruan *et al.*, (2024)[36]	Yes	Yes
Phase 3 single-arm, 52-week study to evaluate the long-term safety and tolerability of fezolinetant in women with VMS associated with menopause **(MOONLIGHT 3)**	2693-CL-0307 (Mainland China)	NCT04451226	Yu *et al.*, (2024)[37]	Yes	Yes
Phase 3b randomized, double-blind, placebo-controlled, 24-week study to assess the efficacy and safety of fezolinetant in menopausal women with moderate to severe VMS and considered unsuitable for HRT **(DAYLIGHT)**	2693-CL-0312 (Canada/Europe)	NCT05033886	Unpublished [a]	Yes	Yes

[a] The DAYLIGHT primary manuscript, **Efficacy and safety of fezolinetant for moderate-severe vasomotor symptoms associated with menopause in individuals unsuitable for hormone therapy: phase 3b randomised controlled trial**, is expected to be published online by *BMJ* on November 13th 2024 (and print published in the November 16th 2024 issue of *BMJ*.

Abbreviations: CSR: clinical study report; HRT: hormone replacement therapy; VMS: vasomotor symptoms.

b) Please provide a full list of ongoing studies (with trial registry identifiers). Please state whether the STARLIGHT, HIGHLIGHT and

MOONLIGHT trials are completed or ongoing and describe their relevance to the submission.

A list of the ongoing phase 3 studies is presented in Table 6. Astellas can confirm that the MOONLIGHT trial has been completed as shown in Table 5, and that there is no MOONLIGHT 2 trial in the overall trial programme.

MOONLIGHT, STARLIGHT 2 and STARLIGHT 3 are not considered relevant to this submission as these trials are conducted in an exclusively Asian population that is not generalisable to the UK population. HIGHLIGHT 1 is not considered relevant to this appraisal as the trial is conducted in a specific subpopulation of women with hormone-receptor positive breast cancer, that is contraindicated in the current license for fezolinetant.

Table 6: Ongoing phase 3 trials

Type of Study	Study Number (Location)	Clinical trial registry identifier	Publication	Protocol	CSR
A Phase 3, Randomized, Placebo-controlled, Double-blind Study to Assess the Efficacy and Safety of Fezolinetant in Japanese Women Experiencing Vasomotor Symptoms (Hot Flashes) Associated with Menopause **(STARLIGHT 2)**	2693-CL-0310 (Japan)	NCT06206408	No	No	No
A Phase 3, Randomized, Placebo-controlled, Double-blind Study to Investigate the Long Term Safety of Fezolinetant in Japanese Women Experiencing Vasomotor Symptoms (Hot Flashes) Associated with Menopause (**STARLIGHT 3)**	2693-CL-0311 (Japan)	NCT06206421	No	No	No
A Randomized, Placebo-controlled, Double-blind, Phase 3 Clinical Study to Investigate the Efficacy and Safety of Fezolinetant for Treatment of Moderate to Severe Vasomotor Symptoms (Hot Flashes) in Women with Stage 0 to 3 Hormone Receptor-positive Breast Cancer Who Are Receiving Adjuvant Endocrine Therapy **(HIGHLIGHT 1)**	2693-CL-1303 (Canada/Europe)	NCT06440967	No	No	No

c) **Please also explain whether there is a SKYLIGHT 3 trial ongoing or planned, or completed but not published.**

Astellas can confirm that there is no SKYLIGHT 3 trial in the overall programme.

A11. Please provide data on the characteristics of the UK target population (people with moderate/severe VMS who are either HRT stoppers or who have a history of oestrogen sensitive tumours) using the same variables as in Table 10 of the CS (including BMI, age, smoking history, race, baseline VMS frequency, alcohol history, previous HRT, concurrent use of SSRIs, global region and time since amenorrhea onset, variables that are suggested by the sub-group analyses to be potential outcome modifiers), so that comparability can be checked with the DAYLIGHT trial characteristics.

As outlined in responses to Questions A2 and A7, the target population of relevance to the decision problem is menopausal people with moderate to severe vasomotor-predominant symptoms for whom HRT is deemed unsuitable (HRT-contraindicated, HRT-caution, HRT-stoppers and HRT-averse). As such, baseline characteristics data have already been presented in the Company submission for this population.

A12. Please provide

a) **the likely numbers of people in the UK who are either HRT stoppers or who have a history of oestrogen sensitive tumours**

Estimating the likely number of menopausal people in the UK with a history of oestrogen-sensitive tumours who might be eligible for fezolinetant presents several challenges. For many patients with conditions like breast cancer, long-term management often includes active or periodic treatments, such as hormone therapies (e.g., tamoxifen or aromatase inhibitors; including extended hormone therapy for longer than 5 years[38]), which may continue for several years to reduce recurrence risk, even when the patient is in remission. Due to this treatment history and contraindications outlined in the SmPC for fezolinetant[26] (not licensed for use in women with breast cancer or those being treated for breast cancer), it is expected that only a minimal number of patients with a history of oestrogen-sensitive tumours would be deemed suitable for fezolinetant.

Astellas is also not aware of any published data that could reliably estimate patient numbers for this group within the UK. As such, Astellas is unable to provide the likely number of people in the UK who have a history of oestrogen sensitive tumours.

Regarding HRT stoppers, three UK clinical experts have estimated that, on average, [redacted]% of menopausal people with moderate to severe VMS are HRT-stoppers.[1]

b) the number of people who are HRT-unsuitable, according to the definitions in the CS

In line with responses to Questions A2 and A7, Astellas has revised the target population of the decision problem to include menopausal people who are HRT-averse.

Astellas is aware of a fairly recent publication estimating the HRT-averse population in the UK at 17.28% (n=70/405).[39] However, this cross-sectional online survey, conducted between December 2019 and February 2020, is now outdated. Feedback from UK clinical experts at the OMA meeting indicated that since 2020, most women are now HRT willing and actively requesting HRT from their NHS GPs, driven by menopause initiatives, celebrity endorsement and an improved understanding of HRT risks. As such, the estimate from Nappi *et al.*, (2023)[39] is expected to overestimate the current HRT-averse population in the UK due to changes in perceptions and prescribing practices around HRT.

Due to the lack of up-to-date data sources from the literature, Astellas consulted three UK clinical experts for estimates of the HRT-averse population. On average, experts estimated that [redacted]% (range: [redacted]%) of menopausal people who are deemed unsuitable for HRT are HRT-averse. Notably, since two of the three clinicians work in secondary care (the other working in primary care), this estimate may be skewed toward the higher end of the range, suggesting that the actual figure is likely lower than the average estimate, given that fezolinetant will primarily be used in NHS primary care.

Overall, the proportion of menopausal people who are deemed HRT-unsuitable (HRT-contraindicated, HRT-caution, HRT-stoppers or HRT-averse) is estimated at [redacted]% in line with feedback from UK clinical experts.[1, 40]

A13. Please provide information, where known, on the clinically important thresholds for all outcomes used in the DAYLIGHT and SKYLIGHT trials.

Clinically important thresholds for change from baseline in moderate to severe VMS frequency, PROMIS SD SF 8b and MENQOL were derived from pooled SKYLIGHT 1 and SKYLIGHT 2 data trials and are described below. These thresholds were subsequently applied to the outcomes of the DAYLIGHT trial.

VMS

In the pooled analysis of SKYLIGHT 1 and SKYLIGHT 2, the mean (standard deviation [SD]) thresholds for a meaningful within-patient change in moderate to severe VMS frequency were estimated to be a reduction of 5.73 (3.47) VMS episodes per day at Week 4 and a reduction of 6.20 (5.18) VMS episodes per day at Week 12. When these thresholds were applied, greater proportions of responders were observed in the fezolinetant 45 mg group compared with placebo at Week 4 (odds ratio [OR]: 2.91, 95% CI: 2.09–4.07; p<0.001) and Week 12 (OR: 2.68; 95% CI: 1.94–3.74; p<0.001).[41]

When these same thresholds were applied to the DAYLIGHT trial, the proportion of responders with ≥50%, ≥75%, and 100% reductions in moderate to severe VMS frequency was higher in the fezolinetant 45 mg group than placebo group at Week 24 (≥50% reduction: OR 1.82, 95% CI: 1.25–2.65; p=0.002; ≥75% reduction: OR: 2.10, 95% CI: 1.43–3.10; p<0.001; 100% reduction: OR: 2.39, 95% CI: 1.42–4.10, p=0.001).[42]

PROMIS SD SF 8b

In the pooled analysis of SKYLIGHT 1 and SKYLIGHT 2, a PROMIS SD-SF-8b total score of 8 points (range: 6 to 12 points) was identified as the within-patient threshold to use to confirm a meaningful change in sleep disturbance.[43]

MENQOL

In the pooled analysis of SKYLIGHT 1 and SKYLIGHT 2, a reduction in MENQOL overall score of ≥ 0.9 points was identified as the clinically important threshold of response. Thresholds of 2.0 points for the vasomotor domain and 0.9 for the psychosocial domain were estimated, in addition to distribution-based threshold estimates of 0.8 and 1.2 for the physical and sexual domains, respectively.[44]

A14. PRIORITY: The CS claims that there were no noticeable differences between fezolinetant and placebo groups in rates of cancer in the DAYLIGHT or SKYLIGHT trials. However, a recent Lancet publication[45] has drawn attention to the significantly higher risk of neoplasm in the fezolinetant groups than the placebo group. Please comment on this risk and provide any additional relevant evidence.

Astellas is committed to patient safety and the continual monitoring of its medicines. The totality of data assessed by Astellas including nonclinical and clinical safety data from the fezolinetant development programme, the structural properties of fezolinetant, literature on key characteristics of carcinogens and epidemiological data, and post-marketing data, do not support a signal for an increased risk of neoplasms.

The potential risk of neoplasms with fezolinetant has been a topic of regulatory review by all regulatory authorities that have granted marketing authorisations for fezolinetant. These authorities have concluded that an association between the use of fezolinetant and the occurrence of malignancies could not be established.[25, 46, 47] As stated in the FDA Clinical Review file published on the FDA website, the FDA concluded that 50% of malignancy events in fezolinetant treatment arms were likely due to pre-existing malignancy undiagnosed at the time of enrolment, bringing the rate of events to within the normal background rate of cancer for the age group.[46] In the EPAR published on the EMA website, the EMA noted the imbalance to be a chance finding.[25]

Douxfils *et al.,* (2023)[45] published a meta-analysis of SKYLIGHT 1, SKYLIGHT 2 and SKYLIGHT 4 in *The Lancet* in October 2023 using the Peto method, which is not an appropriate method for investigating rare events and is not an advocated default approach for meta-analyses.

Neurokinin receptor (NKR; NK1R, NK2R, NK3R) agonism in neoplastic development (proliferation and migration of tumour cells) is reported, and many tumour cells express NKRs.[48, 49] NKR antagonism may play a role as a potential antineoplastic target.[48, 50] The literature does not support the role of NK3R antagonism in neoplasm development.

In SKYLIGHT 1 and SKYLIGHT 2, there were no neoplasm events related to the study drug. In SKYLIGHT 4, a numeric imbalance was observed in the incidence of TEAEs for neoplasms between the fezolinetant and placebo groups.[35] No malignant neoplasms were considered

related to fezolinetant in the completed phase 2 STARLIGHT and phase 3 MOONLIGHT 1, MOONLIGHT 3 and DAYLIGHT trials.

Based on the imbalance seen in SKYLIGHT 4, a thorough *post hoc* review of all reported malignant neoplasm events in phase 2 and 3 fezolinetant studies was conducted. The *post hoc* review included 2,555 individuals who received fezolinetant treatment and 736 who received placebo. Clinical observations included a short latency period to diagnosis, and a drug effect was not supported considering tumour type heterogeneity, prior neoplastic/risk factor history and presence of alternative baseline aetiologies (i.e., pre-existing conditions). There was no evidence of genotoxicity or carcinogenicity in the fezolinetant nonclinical programme. There is also no known agent that could be singularly responsible for the infrequent and diverse pathophysiology of observed events.

Cumulatively to the data lock point of 11 May 2024, 4,744 patients have been enrolled in the fezolinetant clinical trial programme, of which 3,643 were exposed to fezolinetant. The exposure from marketing experience at the data lock point of 11 May 2024 is estimated to be 19,416 patient-years. From post marketing data since first launch to date, there are no safety signals of neoplasm identified.

A15. SSRI/SNRI medication, clonidine and gabapentinoids do not appear to have been prohibited during therapy in the DAYLIGHT and SKYLIGHT 1 and 2 trials, but no information is given on the number of patients in the fezolinetant and placebo groups in each of the three trials who were using these medications before or during the trials. Please provide the numbers of patients using 1) antidepressants, 2) clonidine 3) gabapentinoids and 4) any non-hormonal pharmacologic treatment in each arm of the DAYLIGHT and SKYLIGHT 1 and 2 trials.

The numbers of patients using antidepressants (SSRI/SNRI), clonidine, gabapentinoids and any non-hormonal pharmacologic treatment across the DAYLIGHT and SKYLIGHT 1 and SKYLIGHT 2 trials are presented in Table 7 and Table 8 below.

Table 7: Number of patients using select concomitant medications in DAYLIGHT

Medications	Fezolinetant 45 mg	Placebo
Medication use at baseline		
N	[redacted]	[redacted]
Antidepressants (SSRI/SNRI)	[redacted]	[redacted]
Clonidine	[redacted]	[redacted]
Gabapentinoids (pregabalin and gabapentin)	[redacted]	[redacted]
Medication use initiated post-baseline		
N	[redacted]	[redacted]
Antidepressants (SSRI/SNRI)	[redacted]	[redacted]
Clonidine	[redacted]	[redacted]
Gabapentinoids (pregabalin and gabapentin)	[redacted]	[redacted]

Abbreviations: SSRI: Serotonin–norepinephrine reuptake inhibitor; SNRI: serotonin–norepinephrine reuptake inhibitor.

Table 8: Number of patients using select concomitant medications in SKYLIGHT 1 and SKYLIGHT 2

Medications	SKYLIGHT 1		SKYLIGHT 2	
	Fezolinetant 45 mg	Placebo	Fezolinetant 45 mg	Placebo
Medication use at baseline				
N	[redacted]	[redacted]	[redacted]	[redacted]
Antidepressants (SSRI/SNRI)	[redacted]	[redacted]	[redacted]	[redacted]
Clonidine	[redacted]	[redacted]	[redacted]	[redacted]
Gabapentinoids (pregabalin and gabapentin)	[redacted]	[redacted]	[redacted]	[redacted]
Medication use initiated post-baseline				
N	[redacted]	[redacted]	[redacted]	[redacted]
Antidepressants (SSRI/SNRI)	[redacted]	[redacted]	[redacted]	[redacted]
Clonidine	[redacted]	[redacted]	[redacted]	[redacted]
Gabapentinoids (pregabalin and gabapentin)	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: SSRI: Serotonin–norepinephrine reuptake inhibitor; SNRI: serotonin–norepinephrine reuptake inhibitor.

A16. Priority question: Only naïve pooling of outcome data from the DAYLIGHT and SKYLIGHT 1 and 2 trials is presented. Please provide a meta-analysis of the SKYLIGHT 1 and 2 and DAYLIGHT trials for all relevant outcomes (and including appropriate subgroups of SKYLIGHT 4 where possible).

Data for all relevant outcomes (change from baseline VMS frequency, change from baseline VMS severity, PROMIS 8a, PROMIS 8b, EQ5D-5L VAS, MENQOL, WPAI-VMS) from the meta-analysis of 12 week pooled data from DAYLIGHT and HRT-unsuitable populations from SKYLIGHT 1, SKYLIGHT 2 and SKYLIGHT 4 are provided in Table 9 to Table 16 below.

Table 9: Change from baseline to Week 12 in mean frequency of moderate to severe VMS per 24 hours – ITT Analysis Set

Analysis visit	Statistics	Placebo (n=[redacted])	Fezolinetant 45 mg (n=[redacted])
Baseline	n	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]
Week 12	n	[redacted]	[redacted]
	mean (SD)	[redacted]	[redacted]
	Change from baseline[a]		
	n	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]
	Median	[redacted]	[redacted]

SKYLIGHT-1, SKYLIGHT-2, and DAYLIGHT studies are included.
SDs are calculated as a pooled estimate of the within-study subject-to-subject variability.
[a]A negative change indicated a reduction from baseline (that is, a favourable outcome).
Abbreviations: ITT: intention to treat; SD: standard deviation; VMS: vasomotor symptoms.

Table 10: Change from baseline to Week 12 in mean severity of moderate to severe VMS per 24 hours – ITT Analysis Set

Analysis visit	Statistics	Placebo (n=[redacted])	Fezolinetant 45 mg (n=[redacted])
Baseline	n	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]
Week 12	n	[redacted]	[redacted]
	mean (SD)	[redacted]	[redacted]
	Change from baseline[a]		
	n	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]
	Median	[redacted]	[redacted]

SKYLIGHT-1, SKYLIGHT-2, and DAYLIGHT studies are included.
SDs are calculated as a pooled estimate of the within-study subject-to-subject variability.
[a]A negative change indicated a reduction from baseline (that is, a favourable outcome).
Abbreviations: ITT: intention to treat; SD: standard deviation; VMS: vasomotor symptoms.

Table 11: Change from baseline to Week 12 in patient-reported sleep disturbance by the PROMIS SRI SF 8a total – ITT Analysis Set

Analysis visit	Statistics	Placebo (n=[redacted])	Fezolinetant 45 mg (n=[redacted])
Baseline	n	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]
Week 12	n	[redacted]	[redacted]
	mean (SD)	[redacted]	[redacted]
	Change from baseline[a]		
	n	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]
	Median	[redacted]	[redacted]

SKYLIGHT-1 and SKYLIGHT-2 studies are included.
SDs are calculated as a pooled estimate of the within-study subject-to-subject variability.
[a]A negative change indicated a reduction from baseline (that is, a favourable outcome).
Abbreviations: ITT: intention to treat; PROMIS SRI SF 8a: Patient−Reported Outcome Measurement Information System Sleep Related Impairment – Short Form 8a; SD: standard deviation; VMS: vasomotor symptoms.

Table 12: Change from baseline to Week 12 in patient-reported sleep disturbance by the PROMIS SD SF 8b total – ITT Analysis Set

Analysis visit	Statistics	Placebo (n=[redacted])	Fezolinetant 45 mg (n=[redacted])
Baseline	n	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]
Week 12	n	[redacted]	[redacted]
	mean (SD)	[redacted]	[redacted]
	Change from baseline[a]		
	n	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]
	Median	[redacted]	[redacted]

SKYLIGHT-1, SKYLIGHT-2, and DAYLIGHT studies are included.

SDs are calculated as a pooled estimate of the within-study subject-to-subject variability.
[a]A negative change indicated a reduction from baseline (that is, a favourable outcome).
Abbreviations: ITT: intention to treat; PROMIS SD SF 8b: Patient-reported Outcomes Measurement Information System Sleep Disturbance – Short Form 8b; SD: standard deviation; VMS: vasomotor symptoms.

Table 13: Change from baseline to Week 12 in EQ-5D-5L VAS Score – ITT Analysis Set

Analysis visit	Statistics	Placebo (n=[redacted])	Fezolinetant 45 mg (n=[redacted])
Baseline	n	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]
Week 12	n	[redacted]	[redacted]
	mean (SD)	[redacted]	[redacted]
	Change from baseline		
	n	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]
	Median	[redacted]	[redacted]

SKYLIGHT-1, SKYLIGHT-2, SKYLIGHT-4, and DAYLIGHT studies are included.
Note that the inclusion and exclusion criteria of SKYLIGHT 4 is different from SKYLIGHT 1 and SKYLIGHT 2, as SKYLIGHT 4 didn't have the criteria of moderate-to-severe hot flushes of 7+.
SDs are calculated as a pooled estimate of the within-study subject-to-subject variability.
Abbreviations: EQ-5D-5L: EuroQoL five dimensions five level version; ITT: intention to treat; SD: standard deviation; VMS: vasomotor symptoms.

Table 14: Change from baseline to Week 12 in MENQOL vasomotor domain scores – ITT Analysis Set

Analysis visit	Statistics	Placebo (n=[redacted])	Fezolinetant 45 mg (n=[redacted])
Baseline	n	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]
Week 12	n	[redacted]	[redacted]
	mean (SD)	[redacted]	[redacted]
	Change from baseline[a]		
	n	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]
	Median	[redacted]	[redacted]

SKYLIGHT-1, SKYLIGHT-2, SKYLIGHT-4, and DAYLIGHT studies are included.
Note that the inclusion and exclusion criteria of SKYLIGHT 4 is different from SKYLIGHT 1 and SKYLIGHT 2, as SKYLIGHT 4 didn't have the criteria of moderate-to-severe hot flushes of 7+.
SDs are calculated as a pooled estimate of the within-study subject-to-subject variability.
[a]A negative change indicated a reduction from baseline (that is, a favourable outcome).
Abbreviations: ITT: intention to treat; MENQOL: menopause-specific quality-of-life; SD: standard deviation; VMS: vasomotor symptoms.

Table 15: Change from baseline to Week 12 in WPAI−VMS domain (overall productivity loss) scores – ITT Analysis Set

Analysis visit	Statistics	Placebo (n=[redacted])	Fezolinetant 45 mg (n=[redacted])
Baseline	n	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]
Week 12	n	[redacted]	[redacted]
	mean (SD)	[redacted]	[redacted]

	Change from baseline[a]		
	N	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]
	Median	[redacted]	[redacted]

SKYLIGHT-1, SKYLIGHT-2 and DAYLIGHT studies are included.
SDs are calculated as a pooled estimate of the within-study subject-to-subject variability.
[a]A negative change indicated a reduction from baseline (that is, a favourable outcome).
Abbreviations: ITT: intention to treat; SD: standard deviation; VMS: vasomotor symptoms; WPAI−VMS: Work Productivity and Activity Impairment for VMS Questionnaire.

Table 16: Change from baseline to Week 12 in WPAI−VMS domain (activity impairment) scores – ITT Analysis Set

Analysis visit	Statistics	Placebo (n=[redacted])	Fezolinetant 45 mg (n= [redacted])
Baseline	n	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]
Week 12	n	[redacted]	[redacted]
	mean (SD)	[redacted]	[redacted]
	Change from baseline[a]		
	n	[redacted]	[redacted]
	Mean (SD)	[redacted]	[redacted]
	Median	[redacted]	[redacted]

SKYLIGHT-1, SKYLIGHT-2 and DAYLIGHT studies are included.
SDs are calculated as a pooled estimate of the within-study subject-to-subject variability.
[a]A negative change indicated a reduction from baseline (that is, a favourable outcome).
Abbreviations: ITT: intention to treat; SD: standard deviation; VMS: vasomotor symptoms; WPAI−VMS: Work Productivity and Activity Impairment for VMS Questionnaire.

A17. For each fezolinetant RCT please state how many patients were excluded, prior to randomisation, for each of the following reasons:

i) For having a medical condition or chronic disease

ii) Due to blood pressure measurements (e.g. having systolic blood pressure ≥130 mm Hg or diastolic blood pressure ≥80 mm Hg based on an average of two or three readings on at least two different occasions)?

iii) For using a prohibited treatment for vasomotor symptoms e.g. prescription, over the counter, or herbal

The numbers of patients excluded from the DAYLIGHT and SKYLIGHT 1 and SKYLIGHT 2 trials due to having a medical condition or chronic disease, blood pressure measurements and for using a prohibited treatment for vasomotor symptoms are presented in Table 8 below.

Table 17: Number of patients excluded prior to randomisation for select reasons in DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2

	Number of patients

Reason for exclusion	DAYLIGHT	SKYLIGHT 1	SKYLIGHT 2
Medical condition or chronic disease	[redacted]	[redacted]	[redacted]
Blood pressure measurements	[redacted]	[redacted]	[redacted]
Prohibited treatment for VMS	[redacted]	[redacted]	[redacted]

Some participants who met these exclusion criteria were randomised in error; this summary does not include those participants.
Abbreviations: NA: not applicable; VMS: vasomotor symptoms.

A18: Please clarify and justify the definition of 'moderate to severe VMS' used:

a) in the decision problem

b) as an inclusion criterion in the DAYLIGHT trial

c) as an inclusion criterion in the SKYLIGHT trials

d) in NHS clinical practice

with respect to number of events and severity of each event, and whether only one or both are considered.

The decision problem uses the definition of 'moderate to severe VMS' in alignment with the licensed indication of fezolinetant.[26]

The DAYLIGHT and SKYLIGHT trials used definitions of severity presented in guidance documents published by the EMA and FDA, as shown in Table 18.[23, 24]

Table 18: Definitions of the severity of VMS according to the EMA/FDA[23, 24]

Severity of VMS	Definition
Mild	Sensation of heat without sweating
Moderate	Sensation of heat with sweating, able to continue activity
Severe	Sensation of heat with sweating, causing cessation of activity

Abbreviations: EMA: European Medicines Agency; FDA: US Food and Drug Administration; VMS: vasomotor symptoms.

In UK clinical practice the definition of moderate to severe VMS may be somewhat more subjective than the formal EMA and FDA definitions used in the trials (Table 18), but the data are expected to be generalisable to the UK population and no major deviations are expected between clinical practice and trial definitions.

This topic was discussed during the OMA meeting, where UK clinical experts acknowledged that there is no clear-cut definition of severity, and that clinical decisions are primarily based on the individual patient's experience and how their symptoms impact daily life, but were broadly in agreement that definitions used in the trials were reasonable and relevant to NHS practice.

Lastly, it is worth noting that, while moderate to severe refers to the number of each individual episode, the number of events is not used to define moderate to severe VMS in either the decision problem, clinical trials and NHS clinical practice.

A19. The FDA has recently recommended baseline and follow-up hepatic laboratory tests for people taking fezolinetant. Please comment on whether similar recommendations are expected from the MHRA in the near future, and how they would be implemented in the NHS.

Astellas can confirm that discussions around hepatic laboratory tests for patients receiving fezolinetant are ongoing with the EMA and shortly Medicines and Healthcare Products Regulatory Agency (MHRA). The expected result of these discussions is currently unknown. However, outcomes could be expected within the next year.

No changes have occurred in the wording of the marketing authorisation, the base case therefore has remained unchanged.

A20. Please provide test for interaction results for all the subgroup analyses reported in Section B.2.7.2.1.

The test for interaction values for the subgroup analyses (pooled SKYLIGHT 1 and SKYLIGHT 2 analysis) in section B.2.7.2.1 are presented in Table 19 to Table 22 below.

Table 19: Intrinsic factors: Subgroup analysis of difference in mean change in frequency of moderate to severe VMS from baseline to Week 12 compared with placebo (Full Analysis Set); pooled SKYLIGHT 1 and SKYLIGHT 2

Figure 20 of Document B	
Subgroup	**P-value for subgroup treatment interaction at Week 12**
Age category 1	[redacted]
Age category 2	[redacted]
Race	[redacted]
Ethnicity	[redacted]
BMI	[redacted]
Time since amenorrhea onset	[redacted]
VMS duration	[redacted]
Baseline VMS frequency	[redacted]
Baseline VMS severity	[redacted]

p-value assesses the statistical significance of the subgroup-by-treatment interaction term in the MMRM model.
Abbreviations: BMI: body mass index; CI: confidence interval; MMRM: mixed model repeated measures; VMS: vasomotor symptoms.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs and ISE.

Table 20: Extrinsic factors: Subgroup analysis of difference in mean change in frequency of moderate to severe VMS from baseline to Week 12 compared with placebo (Full Analysis Set); pooled SKYLIGHT 1 and SKYLIGHT 2

Figure 21 of Document B	
Subgroup	**P-value for subgroup treatment interaction at Week 12**
Smoking status	[redacted]
Region	[redacted]
Sleep disturbance (PROMIS SD SF 8b)	[redacted]

Medical history of hypertension	[redacted]
Oophorectomy history	[redacted]
Hysterectomy history	[redacted]
Hysterectomy/Oophorectomy history	[redacted]
Previous HRT treatment	[redacted]
Reason for stopping HRT	[redacted]
Concurrent use of SSRIs	[redacted]
Alcohol history	[redacted]

p-value assesses the statistical significance of the subgroup-by-treatment interaction term in the MMRM model.
Abbreviations: HRT: hormone replacement therapy; MMRM: mixed model repeated measures; SSRI: selective serotonin reuptake inhibitor.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs and ISE.

Table 21: Intrinsic factors: Subgroup analysis of difference in mean change in severity of moderate to severe VMS from baseline to Week 12 compared with placebo (Full Analysis Set); pooled SKYLIGHT 1 and SKYLIGHT 2

Figure 22 of Document B	
Subgroup	**p-value for subgroup treatment interaction at Week 12**
Age category 1	[redacted]
Age category 2	[redacted]
Race	[redacted]
Ethnicity	[redacted]
BMI	[redacted]
Time since amenorrhea onset	[redacted]
VMS duration	[redacted]
Baseline VMS frequency	[redacted]
Baseline VMS severity	[redacted]

p-value assesses the statistical significance of the subgroup-by-treatment interaction term in the MMRM model.
Abbreviations: BMI: body mass index; CI: confidence interval; MMRM: mixed model repeated measures; VMS: vasomotor symptoms.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs and ISE.

Table 22: Extrinsic factors: Subgroup analysis of difference in mean change in severity of moderate to severe VMS from baseline to Week 12 compared with placebo (Full Analysis Set); pooled SKYLIGHT 1 and SKYLIGHT 2

Figure 22 of Document B	
Subgroup	**p-value for subgroup treatment interaction at Week 12**
Smoking status	[redacted]
Region	[redacted]
Sleep disturbance (PROMIS SD SF 8b)	[redacted]
Medical history of hypertension	[redacted]
Oophorectomy history	[redacted]
Hysterectomy history	[redacted]
Hysterectomy/Oophorectomy history	[redacted]
Previous HRT treatment	[redacted]

Reason for stopping HRT	[redacted]
Concurrent use of SSRIs	[redacted]
Alcohol history	[redacted]

p-value assesses the statistical significance of the subgroup-by-treatment interaction term in the MMRM model.
Abbreviations: HRT: hormone replacement therapy; MMRM: mixed model repeated measures; SSRI: selective serotonin reuptake inhibitor.
Source: Astellas (Data on File). SKYLIGHT 1 and SKYLIGHT 2 CSRs and ISE.

Use of estimands

A21. Please explain why different estimand approaches were used across the DAYLIGHT and SKYLIGHT trials and how this might affect results and heterogeneity between the trials.

The approved Version 1 of the protocols for SKYLIGHT 1 and SKYLIGHT 2 were finalised on 28 January 2019.[51, 52] This was before the "ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials" was adopted (20 November 2019), and also before the Astellas protocol preparation processes fully incorporated the estimand structure.

Despite this, the SKYLIGHT 1 and SKYLIGHT 2 protocols describe the hypothetical estimand as the primary strategy.[51, 52] This is appropriate in the confirmatory regulatory setting for which these studies were designed, to allow a comparison of treatment arms as though the subjects had continued on the assigned treatments. Although there is no explicit definition of intercurrent events in the protocol, it is mentioned in the Statistical Analysis Plan (SAP) in relation to discontinued subjects (see the response to A22 below), and the statistical methods in the SAP address missing data.[53] There was no attempt to collect efficacy data following treatment discontinuation.

It should be noted that the SKYLIGHT 1 and SKYLIGHT 2 trials are identical in their design, and that the co-primary endpoints were assessed at Week 4 and Week 12.

The DAYLIGHT trial had an approved Version 1 of the protocol dated 10 June 2021.[54] It fully describes an estimand strategy, including the handling of intercurrent events. The primary estimand strategy is treatment policy, in accordance with the requirements of Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) in Germany[55], and a supplementary analysis was also presented using a hypothetical estimand. For the treatment policy analysis all data from the 24-week placebo-controlled treatment period were to be considered, regardless of participant discontinuation from treatment prior to Week 24. The study has a single primary endpoint, at Week 24.

The study population in DAYLIGHT was based on the HRT-unsuitable subgroup that had previously been derived for various pre-planned and post hoc analyses of data from the SKYLIGHT trials but was otherwise very similar to SKYLIGHT 1 and SKYLIGHT 2. The assessment and analysis of the mean change in moderate to severe VMS frequency, and of the mean change in VMS severity, was the same in all three studies.

The impact of these different estimand approaches is likely to be small. In the DAYLIGHT trial, following a treatment policy strategy, Astellas included data collected after treatment discontinuation from 27 of 226 fezolinetant participants (median of 7 days of data) and 41 of 226

placebo participants (median of 8 days of data). The results from the supplementary hypothetical estimand analysis at Week 24 align closely with the those from the primary treatment policy analysis (presented in DAYLIGHT CSR Tables 9.3.1.1 and 9.3.1.4).[56]

A22. For the SKYLIGHT trials, please define what constituted an intercurrent event and explain how these events were subsequently handled in the analyses.

Intercurrent events were not explicitly defined in the SKYLIGHT 1 and SKYLIGHT 2 trial protocols, and have a single mention in the Statistical Analysis Plans.[51, 52] The context for this approach is described in the response to A21 above.

For the primary analysis, a hypothetical estimand strategy was followed. A mixed model repeated measures analysis of covariance (MMRM) used all available on-treatment data, without requiring explicit imputation for missing data due to intercurrent events (i.e. for discontinued subjects). The MMRM approach under the hypothetical estimand uses a missing at random (MAR) assumption for missing data.

A23. For each of the SKYLIGHT trials, please state whether or not patients continued to have follow-up after they experienced an intercurrent event. If follow-up was continued, please present data on how many patients in each trial arm had imputed data used in the analyses when actual trial data were available for those participants.

For the SKYLIGHT 1 and SKYLIGHT 2 trials there was no protocol-mandated collection of outcome data following treatment discontinuation.[51, 52] These data were not required for the hypothetical estimand strategy, and a treatment policy estimand was not described in the protocol or in the statistical analysis plan.

No subjects had imputed data used in the analyses when actual trial data were available.

Database searches: Clinical SLR searches

A24. Extra searches for escitalopram are referred to in Section D.1.1.2 SLR Updates 1 and 2 (p. 24, p. 29, p. 34.) but the search strategies are missing from Appendix D. Please provide.

Escitalopram was not included in the original SLR and, as such, was included as part of the SLR updates.

In the main search strategy for SLR Update 1, presented in Table 5 of Appendix D, escitalopram was searched with a date limit in order to identify studies published since the original SLR was conducted in 2021. Following this, search terms for escitalopram were combined with disease area terms and study design terms, re-run without a date limit and deduplicated against the SLR Update 1 results to ensure that all studies investigating escitalopram were screened, regardless of publication date. The search strategies used in MEDLINE, Embase and Cochrane databases are presented in Table 23, Table 24 and Table 25, respectively.

In SLR Update 2, the main search strategy (i.e., including all interventions) was run without a date limit and de-duplicated against the EndNote library from SLR Update 1 to ensure that no articles were missed between the two SLR updates. As such, no extra searches for escitalopram were required for SLR Update 2.

Table 23. Escitalopram search terms for use in MEDLINE databases (searched via the OVID SP platform)

Term Group	#	Searches	Results (30.08.23)
Menopause	1.	exp climacteric/	65902
	2.	(menopaus$ or perimenopaus$ or peri-menopaus$ or postmenopaus$ or climacteri$).ti,ab,kf.	108257
	3.	1 or 2	122835
Vasomotor symptoms	4.	exp hot flashes/ or sweating/	10443
	5.	(fl?sh$ or sweat$ or sleep hyperhidrosis or nocturnal diaphoresis).ti,ab,kf.	90864
	6.	exp vasomotor system/	17013
	7.	(vasomotor or VMS).ti,ab,kf.	16605
	8.	or/4-7	121914
	9.	3 and 8	6878
Intervention: Escitalopram	10.	escitalopram/ or (escitalopram or cipralex or enlift or entact or esciprex or Lexapro or "lu 26054 0 lu 260540" or lu260640 or premalex or prilect or seroplex or sipralexa or zecidec or zocital).ti,ab,kf.	3277
Study design	11.	Randomized Controlled Trials as Topic/	163635
	12.	Randomized Controlled Trial/	598860
	13.	Random Allocation/	106955
	14.	Double-Blind Method/	176023
	15.	Single-Blind Method/	32886
	16.	Placebos/	35932
	17.	exp Clinical Trials as topic/	384081
	18.	Clinical Trial/	538684
	19.	Clinical Trial, Phase I/ or Clinical Trial, Phase II/ or Clinical Trial, Phase III/ or Clinical Trial, Phase IV/	81415
	20.	Controlled Clinical Trial/ or Adaptive Clinical Trial/	95448
	21.	randomized controlled trial.pt.	598860
	22.	clinical trial.pt.	538684
	23.	(clinical trial, phase i or clinical trial, phase ii or clinical trial, phase iii or clinical trial, phase iv).pt.	81415
	24.	(controlled clinical trial or multicenter study).pt.	428820
	25.	(clinical adj trial$).ti,ab,kf.	497486
	26.	((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3)).ti,ab,kf.	199067
	27.	placebo$.ti,ab,kf.	249750
	28.	(allocat$ adj2 random$).ti,ab,kf.	44292
	29.	(Randomi?ed adj2 trial$).ti,ab,kf.	435347
	30.	rct.ti,ab,kf.	34718

Term Group	#	Searches	Results (30.08.23)
	31.	or/11-30	2047623
	32.	exp animals/ not exp humans/	5150040
	33.	(comment or editorial or case reports or historical article).pt.	4160310
	34.	(case stud$ or case report$).ti.	386820
	35.	or/32-34	9304766
Combinations	36.	9 and 10 and 31	26
	37.	limit 36 to yr="1946-2020"	21
	38.	37 not 35	21

Database: Ovid MEDLINE(R) and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations and Daily 1946 to August 28, 2023.

Table 24. Escitalopram search terms for use in Embase databases (searched via the OVID SP platform)

Term Group	#	Searches	Results (30.08.23)
Menopause	1.	exp "menopause and climacterium"/	140045
	2.	exp menopause related disorder/	28975
	3.	(menopaus$ or perimenopaus$ or postmenopaus$ or climacteri$).ti,ab,kf.	157640
Vasomotor symptoms	4.	or/1-3	200325
	5.	exp hot flush/ or sweating/	41371
	6.	(fl?sh$ or sweat$ or sleep hyperhidrosis or nocturnal diaphoresis).ti,ab,kf.	124573
	7.	exp vasomotor disorder/ or exp vasomotor system/	3116
	8.	(vasomotor or VMS).ti,ab,kf.	19037
	9.	or/5-8	166265
Intervention: Escitalopram	10.	4 and 9	13006
Study design	11.	escitalopram/ or (escitalopram or cipralex or enlift or entact or esciprex or Lexapro or "lu 26054 0 lu 260540" or lu260640 or premalex or prilect or seroplex or sipralexa or zecidec or zocital).ti,ab,kf.	16217
	12.	"randomized controlled trial (topic)"/	260351
	13.	randomized controlled trial/	780717
	14.	randomization/	98292
	15.	double blind procedure/	210098
	16.	single blind procedure/	51552
	17.	crossover procedure/	75163
	18.	placebo/	401523
	19.	exp "clinical trial (topic)"/	441703
	20.	clinical trial/	1069468
	21.	phase 1 clinical trial/ or phase 2 clinical trial/ or phase 3 clinical trial/ or phase 4 clinical trial/	218797
	22.	controlled clinical trial/ or adaptive clinical trial/ or multicenter study/	778602

Term Group	#	Searches	Results (30.08.23)
	23.	(clinical adj trial$).ti,ab,kf.	719853
	24.	((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3)).ti,ab,kf.	281199
	25.	placebo$.ti,ab,kf.	367318
	26.	(allocat$ adj2 random$).ti,ab,kf.	54942
	27.	(Randomi?ed adj2 trial$).ti,ab,kf.	585940
	28.	rct.ti,ab,kf.	57145
	29.	or/12-28	2947519
	30.	("conference abstract" or "conference review").pt.	4884174
	31.	exp animal/ not exp human/	5132385
	32.	editorial.pt.	777874
	33.	editorial/ or case report/	3650764
	34.	(case stud$ or case report$).ti.	468653
	35.	or/30-34	13020015
Combinations	36.	10 and 11 and 29	104
	37.	limit 36 to yr="1974-2020"	94
	38.	37 not 35	77

Database: Embase 1974 to 2023 August 29.

Table 25. Escitalopram search terms for use in CDSR and CENTRAL databases (searched simultaneously via the Cochrane Library Wiley Online platform

Term Group	#	Searches	Results (30.08.23)
Menopause	1.	[mh "menopause"] or [mh "climacterium"]	9611
	2.	(menopaus* or perimenopause* or postmenopaus* or climacteri*):ab,ti,kw	30858
	3.	7-#2	31606
Vasomotor symptoms	4.	[mh "hot fl*sh"] or [mh "sweating"]	545
	5.	(fl?sh* OR sweat* OR night NEXT sweat* OR "sleep hyperhidrosis" OR "nocturnal diaphoresis"):ab,ti,kw	13064
	6.	(vasomotor OR VMS):ab,ti,kw	2443
	7.	{OR #4-#6}	14630
	8.	#3 AND #7	3102
Intervention: escitalopram	9.	(escitalopram or cipralex or enlift or entact or esciprex or Lexapro or "lu 26054" or "lu 260540" or lu260640 or premalex or prilect or seroplex or sipralexa or zecidec or zocital):ab,ti,kw	2027
Combinations	10.	#8 AND #9	33
	11.	#10 in Trials	33
	12.	#10 in Cochrane Reviews	0

Databases: Cochrane Database of Systematic Reviews (CDSR): Issue 8 of 12, August 2023. Cochrane Central Register of Controlled Trials (CENTRAL): Issue 8 of 12, August 2023.

A25. Were any RCT study design search filters used within the search strategies in Tables 3, 5, and 6 of Appendix D? If yes, please provide a reference to the filters used.

The search strategy used in the original SLR (presented in Table 5 and Table 6 of Appendix D) did not use a published RCT search filter; however, the strategy is not expected to be less sensitive than the Cochrane RCT filter.[57]

The search strategy used in SLR Updates 1 and 2 were adapted from the SIGN filter.[58]

A26. The search strategies in Tables 5 and 6 of Appendix D have a section to limit to adverse effects studies of fezolinetant. Please clarify if an adverse effects search filter was used and provide a reference to the search filter.

The section of the search strategy used for the clinical SLR Updates 1 and 2 which limits to adverse effect studies is not an adverse effect search filter. As a well-established, validated adverse effect filter is not available, a string of free text search terms was developed for use.

Database searches: Cost-effectiveness SLR searches

A27. The search terms perimenopause$ and climacteri$ are missing from the MEDLINE search line 2 in Table 50 on page 140, Section G.1.1.1, Appendix D. They seem to have been replaced by codes. Please check if the searches were run with this error.

Astellas thanks the EAG for flagging this mistake and can confirm that this was due to a copying error. The correct strategy used in the searches is presented in Table 26 with the amended search line 2 in bold text.

Table 26. Search terms used in MEDLINE databases (via the Ovid SP platform)

Term group	#	Search terms	Results	
			Original SLR (2023)	SLR Update (2024)[a]
Disease area: VMS associated with menopause	1	exp climacteric/	65608	66794
	2	**(menopaus$ or perimenopaus$ or peri-menopaus$ or postmenopaus$ or climacteri$).ti,ab,kf.**	107459	110985
	3	1 or 2	122017	125624
	4	exp hot flashes/ or sweating/	10407	10549
	5	(fl?sh$ or sweat$ or sleep hyperhidrosis or nocturnal diaphoresis).ti,ab,kf.	90012	93436
	6	exp vasomotor system/	17010	17023
	7	(vasomotor or VMS).ti,ab,kf.	16521	16906
	8	or/4-7	120991	124758
	9	3 and 8	6834	7047
Fezolinetant	10	(fezolinetant or ESN 364 or ESN364).ti,ab,kf.	18	50
SSRIs	11	exp “serotonin and noradrenaline reuptake inhibitors”/	5465	5619

Term group	#	Search terms	Results	
			Original SLR (2023)	SLR Update (2024)[a]
	12	exp selective serotonin reuptake inhibitors/	46056	46846
	13	(SSRI or selective serotonin reuptake inhibitor or serotonin reuptake inhibitor or serotonin uptake inhibitor or serotonin specific reuptake inhibitor).ti,ab,kf.	11056	11402
Paroxetine	14	paroxetine/ or (34aroom3434am* or arketis or aropax or aroxat or brisdelle or brl 29060 or brl29060 or daparox or deroxat or dexorat or divarius or dropax or dropaxin or euplix or eutimil or fg 7051 or fg7051 or frosinor or mesafem or motivan or nnc 207051 or nnc207051 or optipar or paluxetil or paluxon or paroc or parogen or paroxedura or paroxet or paroxia or paxan or paxil or paxil cr or paxtine or paxxet or pexeva or serestill or sereupin or 34aroom34 or setine or si 211103 or si211103 or solben or syntopar or tagonis).ti,ab,kf.	6844	6963
Citalopram	15	citalopram/ or (citalopram or aurex or ceform or 34aroom or cilopress or cinavol or ciprager or cipram or cipramil or cipraned or ciprotan or ciral or citabax or citacip or 34aroom34 or cital or citalec or citalich or citalon or citalonte or citalostad or citalox or citalvir or citapram or citaxin or citesint or citopam or citrol or citronil or 34aroom3434am or dalsan or elopram or exenadil or frimaind or 34aroom34 or galopran or humorap or kaidor or kitapram or linisan or lopracil or lopraxer or loxopram or lu 10 171 or lu 10171 or lu10 171 or lu10171 or lupram or malicon or nitalapram or oropram or percitale or pralotam or pramital or prefucet or pricital or prisdal or psiconor or recital or renevil or 34aroo or ropramin or selon or sepram or seralgan or seregra or serital or seropram or seror or sintopram or sotovon or talam or talosin or 34aroom or vesema or xadorek or zanipram or zd 211 or zd211 or zeclicid or zentius or zitolex or zyloramor).ti,ab,kf.	10397/	10947
Escitalopram	16	escitalopram/ or (escitalopram or cipralex or enlift or entact or esciprex or Lexapro or "lu 26054 0 lu 260540" or lu260640 or premalex or prilect or seroplex or sipralexa or zecidec or zocital).ti,ab,kf.	3233	3425

Term group	#	Search terms	Results	
			Original SLR (2023)	SLR Update (2024)[a]
Fluoxetine	17	fluoxetine/ or (fluox$ or actan or adofen or afeksin or alzac 20 or andep or andepin or ansilan or "at 001" or at001 or "atd 20" or auroken or auscap or bioxetin or captaton or compound 110140 or daforin or dagrilan or depren or deprex or deprex leciva or deprexetin or deprexin or deprizac or deproxin or diesan or digassim or elizac or exostrept or felicium or fldiss or flotinal or floxet or fluctin or fluctine or fludac or flufran or fluketin or flunil or flunirin or fluohexal or fluoksetin or fluoksetyna or fluoxtab or fluronin or flusac or flustad or flutin or fluxemed or fluxen or fluxet or fluxil or fluxomed or fluzac or fokeston or fontex or foxetin or foxtin or fropine or fuloren or gerozac or her 801 or her801 or ladose or lanclic or lilly 110140 or lilly110140 or lorien or lovan or luramon or ly 110140 or ly110140 or magrilan or margrilan or meropan or modipran or mutan or nopres or npl 2008 or npl2008 or nuzak or olena or oxactin or oxedep or plazeron or plinzene or pragmaten or prizma or proctin or prodep or prosac or prozac or prozac 20 or prozamel or prozamin or prozep or prozit or psipax or qualisac or rapiflux or reconcile or reneuron or rowexetina or salipax or sanzur or sarafem or sartuzin or selfemra or serelsa or seromex or seronil or sinzac or sofelin or stephadilat-s or xeredien or zactin or zepax or zinovat).ti,ab,kf.	26642	27479
SNRI	18	((("dual monoamine reuptake inhibitor" or "dual monoamine uptake inhibitor" or "dual reuptake inhibitor" or "dual uptake inhibitor" or "noradrenalin serotonin reuptake inhibitor" or "noradrenalin serotonin uptake inhibitor" or "norepinephrine serotonin reuptake inhibitor" or "norepinephrine serotonin uptake inhibitor" or "selective serotonin noradrenalin reuptake inhibitor" or serotonin) and noradrenaline reuptake inhibitor$) or "serotonin and noradrenaline uptake inhibitor" or "serotonin and norepinephrine reuptake inhibitor" or "serotonin and norepinephrine uptake inhibitor" or "serotonin noradrenalin uptake inhibitor" or "serotonin norepinephrine reuptake inhibitor" or "serotonin norepinephrine uptake inhibitor" or SNRI$).ti,ab,kf.	2806	2946

Term group	#	Search terms	Results	
			Original SLR (2023)	SLR Update (2024)[a]
Venlafaxine	19	Venlafaxine Hydrochloride/ or (venlafaxin$ or apclaven or dobupal or duofaxin or efectin or efexor or effexor or effexstad or elafax or faxiprol or genexin or hapixed or pracet or serosmine or sigven or sivion or sunveniz or trevilor or trewilor or vandral or vaxor or venix-xr or venla or venlabrain or venlalic or venlaneo or venlax or venlaxin or venlazid or venprimeven or vensuerteven or venxin or venzip or viepax or wy 45030 or wy 45651 or wy 45655 or wy45030 or wy45651 or wy45655 or xenalven or zarelis).ti,ab,kf.	5004	5153
Desvenlafaxine	20	(desvenlafaxine or dvs 233 or dvs233 or ellefore or khedezla or md 120 or md120 or "o desmethylvenlafaxine " or "o norvenlafaxine " or pristiq or wy 45233 or wy45233).ti,ab,kf.	715	759
Clonidine	21	exp clonidine/	13509	13558
	22	(clonidin$ or adesipress$ or arkamin or atensina or caprysin or catapres$ or catasan or chlofazolin or chlophazolin or chlophelin or clinidine or clofelin or clofeline or clomidine or clondine or clonicel or clonipresan or clonistada or clonnirit or clophelin or clopheline or daipres or dcai or dichlorophenylaminoimidazoline or dichlorophenylaminoimidazoline hydrochloride or dixarit or duraclon or haemiton or hemiton or huma-clonidine or hypodine or isoglaucon or jenloga or kapvay or m 5041t or melzin or normopresan or normopresin or paracefan or st 155 or sulmidine or taitecin or tenso timelets).ti,ab,kf.	15682	15829
Gabapentin	23	exp gabapentin/	4367	4494
	24	(gabapen$ or ci 945 or ci945 or dineurin or dm 1796 or dm 5689 or dm1796 or dm5689 or gabalept or gabaliquid geriasan or gabatin or gantin or go 3450 or go3450 or goe 3450 or goe3450 or gralise or kaptin or keneil or neurontin or neurotonin or nupentin or sefelsa or serada).ti,ab,kf.	8252	8676
Oxibutynin	25	(oxibutynin or anturol or ap 1034 or ap1034 or cs 801 or cs801 or cystonorm or cystrin or delifon or ditropan or diutropin or dr 3001 or dr3001 or dridase or driptane or esoxybutynin or fp 1097 or fp1097 or frenurin or gelnique or hob 294 or hob294 or iliaden or "jnp 0101" or jnp0101 or kentera or "kl 007" or kl007 or lenditro or lyrinel xl or mj 4309 1 or mj 43091 or mj43091 or mm 801 or mm801 or mutum cr or nefryl or neoxy tape or novitropan or oxibutinin or oxyban or oxybutinin or oxybutyin or oxytrol or oyrobin or pollakis or reteven or tropan or uricont or uroflax or urotrol or velariq or vesoxx or zatur ge).ti,ab,kf.	244	248
Sertraline	26	exp sertraline/	3414	3499

Term group	#	Search terms	Results	
			Original SLR (2023)	SLR Update (2024)[a]
	27	(sertralin$ or adjuvin or a?truline or aremis or besitran or cp 51974 or "cp 51974 01" or cp 51974 1 or cp 519741 or cp51974 or "cp51974 01" or cp51974 1 or cp5197401 or cp519741 or dominum or doxime or fatral or fridep or gladem or lesefer or lustral or nudep or seltra or serad or sercerin or serlain or serlift or sertraline hydrochloride or sertranex or sertranquil or sosser or tatig or tresleen or zolof? or zosert).ti,ab,kf.	5470	5703
Pregabalin	28	exp pregabalin/	2456	2543
	29	(bonqat or ci 1008 or ci1008 or lyrica or lyrica cr or pd 144723 or pd144723 or vronogabic or ynp 1807 or ynp1807).ti,ab,kf.	819	929
CBT	30	exp Behavior Therapy/	89006	92008
	31	(behavio$ therap$ or behavio$ treatment$ or cognitive therap$ or cognitive treatment$ or CBT).ti,ab,kf.	42247	44727
Interventions combined	32	or/10-31	209061	216635
Economic evaluations	33	Economics/ or exp "Fees and Charges"/ or exp Budgets/	69498	69664
	34	exp Models, Economic/ or exp Cost-Benefit Analysis/ or exp "Costs and Cost Analysis"/	270838	275700
	35	Economics, Nursing/ or exp Economics, Medical/ or Economics, Pharmaceutical/ or exp Economics, Dental/ or exp Economics, Hospital/	50485	50638
	36	Markov Chains/ or Monte Carlo Method/ or Decision Theory/	46210	46821
	37	(cost$ adj2 (effective$ or utilit$ or benefit$ or minimi$ or consequence$)).ti,ab,kf.	205780	220041
	38	((economic$ or pharmacoeconomic$ or cost$ or price$ or pricing$ or expenditure$ or financ$) adj2 (evaluat$ or model$ or analys?s or outcome$)).ti,ab,kf.	93113	98610
	39	(value adj2 (money or monetary)).ti,ab,kf.	3024	3184
	40	(economic model$ or markov or monte carlo).ti,ab,kf.	86885	91227
	41	(decision$ adj2 (tree or analys?s or model$)).ti,ab,kf.	33904	38254
	42	exp Value of Life/ or Quality-Adjusted Life Years/	21231	21831
	43	(quality adjusted life year$ or quality-adjusted life year$ or qaly$ or disability adjusted life year$ or disability-adjusted life year$ or daly$ or life year$ gained or life year$ equivalent$ or incremental cost effective$ or icer or qald$ or qale$ or qtime$).ti,ab,kf.	30402	32716
	44	or/33-43	628621	656219
HSUVs	45	Quality-Adjusted Life Years/ or Value of Life/	21231	21831

Term group	#	Search terms	Results	
			Original SLR (2023)	SLR Update (2024)[a]
	46	(quality adjusted life year$ or quality-adjusted life year$ or qaly$ or disability adjusted life year$ or disability-adjusted life year$ or daly$ or life year$ gained or life year$ equivalent$ or incremental cost effective$ or icer or qald$ or qale$ or qtime$).ti,ab,kf.	30402	32716
	47	(health utilit$ or health state$1 or illness state$1 or HSUV or HSUVs or utility assessment$ or preference based or utility based).ti,ab,kf.	12107	12833
	48	(utility adj3 (score$1 or scoring or valu$ or measur$ or evaluat$ or scale$1 or instrument$1 or weight or weights or weighting or information or data or unit or units or health$ or life or estimat$ or elicit$ or disease$ or mean or cost$ or expenditure$1 or gain or gains or loss or losses or lost or analysis or index$ or indices or overall or reported or calculat$ or range$ or increment$ or state or states or status)).ti,ab,kf.	43591	46281
	49	utility.ab. /freq=2	23452	25154
	50	(utilities or disutilit$).ti,ab,kf.	9756	10362
	51	((index adj3 wellbeing) or (quality adj3 wellbeing) or qwb).ti,ab,kf.	1180	1302
	52	(multiattribute$ or multi attribute$).ti,ab,kf.	1304	1437
	53	(euro qual or euro qual5d or euroqual or euroqual5d or euroqol or euroqol5d or euro qol or euro qol5d or eq-5d or eq5-d or eq5d or eq 5d or eq-sdq or eqsdq).ti,ab,kf.	16560	18079
	54	(short form$ or shortform$).ti,ab,kf.	43799	46520
	55	(sf36$ or sf 36$ or sf thirtysix or sf thirty six).ti,ab,kf.	26473	27588
	56	(sf6 or sf 6 or sf6d or sf 6d or short form 6D or shortform 6D or sf six D or sfsixD or shortform six D or short form six D or sf six or sfsix or sf8 or sf 8 or sf eight or sfeight or short form 6 or shortform 6 or shortform six or short form six).ti,ab,kf.	4114	4284
	57	(sf12 or sf 12 or sf twelve or sftwelve or short form 12 or shortform 12 or shortform twelve or short form twelve).ti,ab,kf.	7633	8063
	58	(sf16 or sf 16 or sf sixteen or sfsixteen or shortform 16 or short form 16 or shortform sixteen or short form sixteen).ti,ab,kf.	39	41
	59	(sf20 or sf 20 or sf twenty or sftwenty or short form 20 or shortform 20 or shortform twenty or short form twenty).ti,ab,kf.	456	468
	60	(15D or 15-D or 15 dimension or 15-dimension).ti,ab,kf.	6105	6301
	61	visual analog$ scale$.ti,ab,kf.	72990	77898
	62	(standard gamble$ or sg).ti,ab,kf.	14136	15149
	63	(time trade off$1 or time tradeoff$1 or tto or timetradeoff$1).ti,ab,kf.	2368	2489
	64	(health$1 year$1 equivalent$1 or hye or hyes).ti,ab,kf.	85	87
	65	(hui or hui1 or hui2 or hui3).ti,ab,kf.	1965	2083

Term group	#	Search terms	Results	
			Original SLR (2023)	SLR Update (2024)[a]
	66	or/45-65	257816	274384
	67	quality of life/ and ((quality of life or qol or hrqol) adj (score$1 or measure$1)).ti,ab,kf.	18417	19506
	68	quality of life/ and health-related quality of life.ti,ab,kf.	44561	47390
	69	quality of life/ and ec.fs.	10875	10883
	70	quality of life/ and (health adj3 status).ti,ab,kf.	11693	12305
	71	((qol or hrqol or quality of life) adj2 (increas$ or decreas$ or improv$ or declin$ or reduc$ or high$ or low$ or effect$ or worse or score$ or change$1 or impact$ or deteriorat$)).ti,ab,kf.	136817	147988
	72	(MENQoL or "women's health questionnaire" or WHQ or "Greene climacteric scale" or PROMIS or "Sleep Disturbance SF" or "Sleep Disturbance short form" or "Sleep Disturbance short-form" or "Sleep Disturbance shortform").ti,ab,kf.	4237	4871
	73	Patient report$.ti,ab,kf.	59938	67044
	74	or/67-73	232204	251778
Direct Healthcare Cost and Resource Use	75	Economics/ or exp "Fees and Charges"/ or exp Budgets/	69498	69664
	76	Economics, Nursing/ or exp Economics, Medical/ or Economics, Pharmaceutical/ or exp Economics, Dental/ or exp Economics, Hospital/	50485	50638
	77	Cost allocation/ or Cost control/ or Cost savings/ or Cost of illness/ or Cost sharing/ or "Deductibles and coinsurance"/ or Medical savings accounts/ or Health care costs/ or Direct service costs/ or Drug costs/ or Employer health costs/ or Hospital costs/ or Health expenditures/ or Capital expenditures/ or Financial management/	163405	165633
	78	(economic$ or pharmacoeconomic$).ti,ab,kf.	389165	415848
	79	((healthcare$ or health care or health-care or drug$ or medication$ or treatment$ or physician$ or nurse$ or nursing or hospital$ or illness$) adj2 cost$).ti,ab,kf.	93618	98938
	80	((unit adj cost$) or (low adj cost$) or (high adj cost$) or (cost adj estimate$) or (cost adj variable$)).ti,ab,kf.	117596	127147
	81	((resource$ or healthcare$ or health care or health-care or service$) adj3 (use$ or utilis$ or utiliz$ or consume$ or consuming or consumption$)).ti,ab,kf.	164800	176157
	82	(price$ or pricing$ or fiscal or funding or financial or finance).ti,ab,kf.	247984	266548
	83	or/75-82	1029882	1093668
	84	limit 83 to yr=2013-current	577533	641290
Exclusion terms	85	exp animals/ not exp humans/	5133020	5210272
	86	(comment or editorial or case reports or historical article).pt.	4135250	4239249
	87	(case stud$ or case report$).ti.	380552	406889
	88	or/85-87	9261625	9446510
Combination	89	9 and 32 and 44	13	14

Term group	#	Search terms	Results	
			Original SLR (2023)	SLR Update (2024)[a]
	90	9 and (66 or 74 or 84)	949	1007
	91	89 or 90	957	1015
	92	91 not 88	**940**	**997**

Database: Ovid MEDLINE(R) and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations and Daily 1946 to June 19, 2023 (Original SLR) and Ovid MEDLINE(R) and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations and Daily 1946 to April 8, 2024 (SLR update). **Footnotes:** [a]Searches were conducted without an amended date limit for the SLR update and the results were de-duplicated against the EndNote library from the original SLR (2023) to ensure that no articles were missed between the original and updated search dates.
Abbreviations: CBT: cognitive behavioural therapy; VMS: vasomotor symptoms; SLR; systematic literature review; SNRI: serotonin norepinephrine reuptake inhibitor; SSRI: selective serotonin reuptake inhibitors.

A28. The search strategies in tables 50 and 51 of Appendix G have a sections to limit to economic evaluations and health state utility values (HSUV). Please clarify if search filters were used for each, and provide references to the search filters.

The search filter used to limit to economic evaluations was adapted from the published SIGN, CADTH and NHS-EED filters.[58-60] The search filter to limit to studies reporting on HSUVs was adapted from the search filters published by ScHARR and YHEC.[61, 62]

Section B: Clarification on cost-effectiveness data

Structure of the economic model

B1. PRIORITY. The health states of the model are defined in terms of VMS frequency rather than severity of VMS. The CS justifies this choice largely because the primary outcome of the efficacy trials is frequency of VMS and the separation of frequency and severity outcomes in the trials limits the construction of a singular endpoint that captures both frequency and severity. However, as expressed by several clinical experts, capturing the severity of VMS plays an important role and VMS frequency alone may not provide an accurate proxy for the severity of VMS. Furthermore, VMS severity was calculated as a secondary outcome in the trials. The CS states that combining both frequency and severity into a single outcome was not performed because

statistical measures of dispersion around the mean for a combined outcome would not be possible to obtain.

a) **Please justify why the modelled health states only considers frequency and not severity of VMS. The company has access to the IPD from the trials for both frequency and severity outcomes; therefore, it should, in principle, be possible to obtain precision estimates for a combined outcome.**

The choice to define health states in terms of VMS frequency rather than severity in the model was not solely based on the primary outcome of the efficacy trials or limitations in the construction of a singular endpoint that captures both frequency and severity. Rather, this decision was heavily informed by extensive feedback from both UK clinical experts and patient representatives, as noted in Section B.3.2.2.1 of Document B.

Feedback from three UK clinical experts with menopause specialty and four patient representatives from three patient advocacy groups (Menopause UK; Menopause Alliance; and Menopause and Cancer) strongly indicated that VMS frequency is more clinically relevant to patients and more closely aligned with NHS clinical practice, when compared with VMS severity.[1, 63] Importantly, they highlighted that frequency serves as a more objective and consistent measure of treatment effect than VMS severity, which can be highly subjective and difficult to quantify accurately in clinical practice.[1, 63]

To further support the use of moderate to severe VMS frequency to define VMS frequency health states in the model, Astellas conducted two analyses that support the relevance of a frequency-based structure. First, a correlation analysis demonstrated that VMS frequency and severity are correlated (Table 27), indicating that the frequency-based health states in the model implicitly captures severity. As such, incorporation of VMS severity as an add on to the existing frequency-based structure would risk introducing double counting of treatment effects.

Table 27: Correlation coefficients between VMS frequency and severity for the fezolinetant 45 mg

Trial	Timepoint	N	Correlation coefficients
SKYLIGHT 1	Week 4	[redacted]	[redacted]
	Week 12	[redacted]	[redacted]
SKYLIGHT 2	Week 4	[redacted]	[redacted]
	Week 12	[redacted]	[redacted]
Pooled SKYLIGHT 1 and 2	Week 4	[redacted]	[redacted]
	Week 12	[redacted]	[redacted]

Abbreviations: VMS: vasomotor symptoms.

Secondly, Astellas conducted an analysis of the relationship between Patient Global Impression of Change in VMS (PGI-C VMS) and VMS frequency. The results showed that improvements in PGI-C VMS were associated with lower VMS frequency (Table 28), underscoring that the model structure aligns well with how patients perceive symptom improvement in the lower VMS

frequency health states. Together, these analyses support the suitability of using VMS frequency alone in the model to capture both clinical relevance and patient experience.

The current modelling approach is also consistent with the NICE clinical guidelines for menopause (NG23), which notes that frequency is often used as a proxy for severity in menopausal symptoms. The NICE NG23 (Appendix L) states: "*The frequency of hot flushes is often used as a proxy for the severity of menopausal symptoms. Therefore, a reduction in the frequency of flushing can be expected to bring a relief in symptoms and concomitant improvement in health-related quality of life.*" Moreover, the cost-effectiveness model used in NICE NG23 also relied on VMS frequency as the primary outcome, explicitly omitting severity due to variability in scoring methods across studies – feedback consistent with both UK clinical experts and patient representative consulted by Astellas for this appraisal.

While Astellas acknowledge that in theory it may be possible to combine frequency and severity into a single scoring scale that could be used to define health states, issues still arise in that there is no accepted combined scale that is quantitative, well known or accepted by clinicians and that can be calculated robustly when you consider the highly subjective nature of severity. As such, Astellas' position remains that modelling VMS health states with frequency alone is appropriate, given the uncertainty around the credibility of a combined scale. Moreover, Astellas' decision to focus on VMS frequency is ultimately grounded in clinical relevance, ease of interpretation, and the unanimous endorsement of both UK clinical experts and patient representatives, ensuring that the model reflects NHS clinical practice and ultimately captures the health effects of treatment in a way that is meaningful for the target population.

Table 28: Association between mean moderate and severe VMS frequency per 24h and PGI-C VMS for fezolinetant 45 mg and placebo (Full Analysis Set)

Mean VMS frequency per 24h	PGI-C VMS, n (%)						
	Much Better	Moderately Better	A little better	No change	A little worse	Moderately worse	Much worse
DAYLIGHT							
<2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥2 to <7	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥7 to 9	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥9	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pooled SKYLIGHT 1 and SKYLIGHT 2 (HRT-unsuitable subpopulation, up to Week 12) and DAYLIGHT							
<2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥2 to <7	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥7 to 9	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥9	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pooled SKYLIGHT 1 and SKYLIGHT 2 (HRT-unsuitable subpopulation, up to Week 52 for fezolinetant patients and up to Week 12 for placebo patients) and DAYLIGHT							
<2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥2 to <7	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥7 to 9	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥9	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Each cell represents the number of visits at which a particular PGI-C VMS response and a particular mean frequency of moderate to severe VMS occurs. Only data from the placebo and fezolinetant 45mg are included in this analysis. Each patient contributed all the visits for which they had both endpoints recorded.
Abbreviations: PGI-C: Patient Global Impression of Change; VMS: vasomotor symptoms.

b) Please justify why the model did not include the key secondary outcome of mean weighted severity of VMS experienced per 24 hours from the trials (Figure 8 of CS for DAYLIGHT).

While the key secondary outcome of mean weighted severity of VMS per 24 hours was not explicitly included in the model, it is important to note that VMS severity and frequency are correlated, as demonstrated in Table 27. This correlation indicates that changes in VMS frequency implicitly capture changes in VMS severity for patients and therefore treatment effects on VMS severity are captured in the current model structure. In addition to this, including VMS severity measures and an add to the existing current model structure would likely result in double counting treatment effects as implicit effects are already captured.

In order to include the key secondary outcome of mean weighted severity without the risk of double counting, this endpoint would need to be combined with frequency to generate a singular combined frequency and severity 'scoring scale' for the model. This scoring scale would lack credibility due to lack of validation and absence from common clinical practice. Moreover, the subjective nature of VMS severity introduces variability, which could negatively impact the robustness and reliability of a combined scale.

B2. PRIORITY. The transition probabilities used in the base-case analysis are said to be based on the population deemed unsuitable for HRT, but appear to include people classified as HRT-averse (as the model includes a separate scenario that excludes the HRT-averse population). Further, as it was discussed during the decision problem meeting, people who are 'HRT-averse' would be categorised within 'People for whom HRT is not considered suitable' because the people themselves have deemed themselves not suitable, and should be included as the target population in the decision problem.

***a)* Please clarify whether the base case population includes people classified as HRT-averse given the information in A3.**

Astellas can confirm that the base case population in the model included menopausal people classified as HRT-averse, despite the target population for the original submission being menopausal people who are deemed HRT-unsuitable for medical reasons (i.e. excluding HRT-averse). All transition probabilities used in the base case are therefore based on data from the full DAYLIGHT population and HRT-unsuitable subpopulation in the SKYLIGHT 1 and SKYLIGHT 2 trials (which all included HRT-averse patients). The full HRT-unsuitable population was included because it was not expected to significantly impact the overall results (given that categorisation based on HRT-unsuitability is due to safety considerations for HRT rather than for treatment modifying effects), and it was considered important to preserve a large sample size of the efficacy data. In the DAYLIGHT trial, 37% of patients were identified as HRT-averse and

excluding such a significant proportion of the trial cohort may have introduced uncertainty into the trial outcomes.

b) **Please clarify whether any other base case input parameters in the model exclude people classified as HRT-averse and, if so, please provide justification.**

Astellas can confirm that all base case inputs based on trial data (transition probabilities, discontinuation rates and utilities) included people classified as HRT-averse. Baseline age was based on the median age of VMS onset reported by the NHS (discussed further in response to B11) for all menopausal people and median duration of VMS symptoms (used to derive natural VMS cessation rate) was based on a study of menopausal people who reported frequent VMS.[64, 65] Therefore, neither of these inputs are based on populations which exclude people classified as HRT-averse.

c) **Please provide two analyses each for a separate population:**

i. ***An analysis for the population deemed unsuitable for HRT that <u>includes</u> people classified as HRT-averse, where all the model inputs are specific to the population of the decision problem. Where it is not possible to provide model inputs specific to this population, please justify the appropriateness of the assumptions used in the model.***

As confirmed in B2 b) the base case analysis was based on inputs for an HRT-unsuitable population, which included people classified as HRT-averse.

ii. ***An analysis for the population deemed unsuitable for HRT that <u>excludes</u> people classified as HRT-averse, where all the model inputs are specific to that population. Where it is not possible to provide model inputs specific to this population, please justify the appropriateness of the assumptions used in the model.***

A scenario was included in the submission model which explored a patient population that excludes people classified as HRT-averse scenario (scenario 17). In this scenario the following inputs were used:

Input/Setting	Base Case	Excluding HRT-averse Scenario
Fezolinetant Efficacy	Week 0–24: • Per cycle transition probabilities from DAYLIGHT (HRT-unsuitable population) Week 24–52: • Average transition probabilities	Week 0–24: • Per cycle transition probabilities from DAYLIGHT (HRT-unsuitable population excluding HRT-averse) Week 24–52: • Average transition probabilities

	from week 24 to 52 from SKYLIGHT 1, 2 (HRT-unsuitable population)	from week 24 to 52 from SKYLIGHT 1, 2 (HRT-unsuitable population excluding HRT-averse)
No Treatment Efficacy	Week 0–24: • Per cycle transition probabilities from DAYLIGHT (HRT-unsuitable population)	Week 0–24: • Per cycle transition probabilities from DAYLIGHT (HRT-unsuitable population excluding HRT-averse)
Fezolinetant and No Treatment Discontinuation	Week 0–24: • DAYLIGHT (HRT-unsuitable population) Week 24+ (fezolinetant only) • Pooled SKLIGHT 1 & 2 (HRT-unsuitable population)	Week 0–24: • DAYLIGHT (HRT-unsuitable excluding averse) Week 24+ (fezolinetant only) • Pooled SKLIGHT 1 & 2 (HRT-unsuitable population)
Utilities	Clinician adjusted estimates from DAYLIGHT (HRT-unsuitable population)	Clinician adjusted estimates from DAYLIGHT (HRT-unsuitable population excluding HRT-averse)*
Baseline distribution	Baseline distribution from DAYLIGHT (HRT-unsuitable population)	Baseline distribution from DAYLIGHT (HRT-unsuitable population)

Abbreviations: HRT: hormone replacement therapy; NA: not applicable.

*Note: In the submission model clinician adjustment for the 0 ≤ VMS Frequency < 2 health state had been mistakenly omitted for the excluding HRT-averse scenario. This has been corrected in the latest model and the clinician provided value for the 0 ≤ VMS Frequency < 2 health state has now been applied in the excluding HRT-averse utility values.

As such, the only trial inputs which were not updated to be specific to the HRT-unsuitable excluding HRT-averse population were:

- The discontinuation rate for fezolinetant from Week 24 onwards, which is based on the discontinuation rate in the HRT-unsuitable subpopulation of pooled SKYLIGHT 1 and 2 from Week 24 to 52 in the base case. The discontinuation rate from week 24 onwards from the HRT-unsuitable subpopulation excluding HRT-averse patients from pooled SKYLIGHT 1 and 2 has now been included in the model and can be selected for in the 'model specifications' tab. In the pooled SKYLIGHT 1 and 2 (HRT-unsuitable, excluding HRT-averse) population, [redacted] treatment discontinuations from [redacted] initial people at risk occurred in the period between Weeks 24 and 52. This represented a per 3 cycle discontinuation probability of [redacted]%.
- The baseline distribution, the baseline distribution from the DAYLIGHT HRT-unsuitable excluding HRT-averse population has been added to the model as an additional option for the excluding HRT-averse scenario. The data for this are presented in Table 29.

Table 29: Baseline distribution in DAYLIGHT HRT-unsuitable excluding HRT-averse population

	0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
Number of people	[redacted]	[redacted]	[redacted]	[redacted]
% of people	[redacted]	[redacted]	[redacted]	[redacted]

iii. Please provide a revised version of the model with sufficient flexibility to incorporate the information for the two analyses with

the population stated in B2.c.i and in B2.c.ii separately. Please signpost the changes made to the model.

The model has been updated such that analyses for the HRT-unsuitable population and HRT-unsuitable population excluding HRT-averse patients can be performed in full, as detailed in the response to B2)c)ii. Any additional data added into the model for the excluding HRT-averse scenario have been signposted using green highlighting.

d) For the clarification questions B3 to B10 and B14 to B15, please provide the corresponding information for the populations stated in B2.c.i and B2.c.ii separately, i.e., for the analyses in part c) above.

i. Please provide a revised version of the model with sufficient flexibility to incorporate the information from part d). Please signpost the changes made to the model.

Where possible, additional data for questions B3 to B10 and B14 and B15 has been added into the model for both the HRT-unsuitable and HRT-unsuitable excluding HRT-averse populations. Any additional data or scenarios that have been included in the model in response to clarification questions have been signposted using green highlighting.

Treatment effectiveness used in the model

B3. PRIORITY. Relative efficacy of fezolinetant compared to no treatment.

a) Please provide estimates of the relative treatment effect, with sampling uncertainty, for fezolinetant compared to placebo from the trials for each VMS frequency health state in each 4-weekly cycle, up to week 24 from DAYLIGHT and up to week 12 from SKYLIGHT 1 & 2 (HRT-unsuitable).

Astellas have not provided estimates of the relative treatment effects from DAYLIGHT or SKYLIGHT 1, 2 because they are not deemed the most appropriate source of efficacy in this model. Further justification for this is provided in response to B3)b)ii.

b) Please apply the estimates of relative treatment effect from part a) to the estimates of natural history in the absence of active treatment to provide an estimate of the transitions between VMS frequency health states over time for fezolinetant and use the estimates of natural

history to represent the no treatment comparator (rather than the transition probabilities from the placebo arm of the trials).

i. Please provide the corresponding estimates of cost-effectiveness of fezolinetant compared to no treatment.

Please see response in b)ii).

ii. Please provide a revised version of the model incorporating the changes made in part a) and part b), with sufficient flexibility to switch between the alternative methods used to estimate the transitions between health states over time. Please signpost the changes made to the model.

The model has not been updated to incorporate the functionality to apply estimates of relative treatment effect as Astellas believes that this would introduce unnecessary uncertainty into the modelling of efficacy for fezolinetant. Applying relative treatment effects to the structure expert elicitation (SEE) estimates for natural history would lead to a loss of granularity in the modelling of fezolinetant efficacy compared with using the observed trial data. At any given timepoint, the combination of SEE estimates and a singular relative treatment effect are significantly less granular than the data available from the clinical trials. In particular, the trial data are able to provide direct evidence of how many women benefit from moving to the lowest frequency health state, which was considered by both clinical experts and patient representatives as the most important factor to patients. Indeed, as shown in Table 28, improvements in PGI-C VMS were associated with lower VMS frequency, underscoring that the model structure aligns well with how patients perceive symptom improvement in the lower VMS frequency health states.

As noted in Section B.3.3.2.3 of Document B, despite SEE being an approved methodology by NICE, there was considerable uncertainty in the SEE estimates for the natural progression of VMS, reflective of the wide 95% credible intervals in Figure 28 of Document B. This was likely due to the inherent challenge associated with estimating the course of symptoms in postmenopausal women who are receiving no active treatment and are deemed unsuitable for HRT. Ultimately, the natural history data elicited from UK clinicians were estimates and not based on observed data and therefore applying relative treatment effects to these natural history estimates to estimate fezolinetant efficacy would introduce unnecessary additional uncertainty into the model.

B4. Average efficacy estimates from the trials over time.

a) Please provide estimates of the average treatment effect over time (with sampling uncertainty) for fezolinetant and placebo from the trials for each VMS frequency health state (e.g., derived from a multinomial

regression model that includes variables to adjust for treatment and 4-weekly visit period and interaction terms). Specifically, please provide:

i. Average health state transition probabilities, reported separately, for fezolinetant and placebo (week 0 to week 24) from DAYLIGHT.

Table 30 and Table 31 present the average health state transition probabilities for fezolinetant for the period of week 0 to week 24 in the DAYLIGHT HRT-unsuitable and HRT-unsuitable excluding HRT-averse populations respectively. These average transition probabilities are included in the model for the period of week 0 to week 12 and week 12 to week 24 for fezolinetant.

Table 32 and Table 33 present the average health state transition probabilities for placebo for the period of week 0 to week 24 in the DAYLIGHT HRT-unsuitable and HRT-unsuitable excluding HRT-averse populations respectively. These average transition probabilities are included in the model for the period of week 0 to week 12 and week 12 to week 24 for placebo.

Table 30: Average transition probabilities for fezolinetant week 0 to week 24 in DAYLIGHT HRT-unsuitable women

From	To			
	0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: HRT: hormone replacement therapy; VMS: vasomotor symptoms.

Table 31: Average transition probabilities for fezolinetant week 0 to week 24 in DAYLIGHT HRT-unsuitable excluding HRT-averse women

From	To			
	0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: HRT: hormone replacement therapy; VMS: vasomotor symptoms.

Table 32: Average transition probabilities for placebo week 0 to week 24 in DAYLIGHT HRT-unsuitable women

From	To			
	0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]

2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: HRT: hormone replacement therapy; VMS: vasomotor symptoms.

Table 33: Average transition probabilities for placebo week 0 to week 24 in DAYLIGHT HRT-unsuitable excluding HRT-averse women

From	To			
	0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: HRT: hormone replacement therapy; VMS: vasomotor symptoms.

ii. Average health state transition probabilities, reported separately, for fezolinetant and placebo (week 0 to week 12) from SKYLIGHT 1 & 2 (HRT-unsuitable).

Table 34 and Table 35 present the average health state transition probabilities for fezolinetant for the period of week 0 to week 12 in the pooled SKYLIGHT 1 & 2 HRT-unsuitable and HRT-unsuitable excluding HRT-averse populations respectively. These average transition probabilities are included in the model for the period of week 0 to week 12 for fezolinetant.

Table 36 and Table 37 present the average health state transition probabilities for placebo for the period of week 0 to week 12 in the pooled SKYLIGHT 1 & 2 HRT-unsuitable and HRT-unsuitable excluding HRT-averse populations respectively. These average transition probabilities are included in the model for the period of week 0 to week 12 for placebo.

Table 34: Average transition probabilities for fezolinetant week 0 to week 12 in pooled SKYLIGHT 1 & 2 HRT-unsuitable women

From	To			
	0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: HRT: hormone replacement therapy; VMS: vasomotor symptoms.

Table 35: Average transition probabilities for fezolinetant week 0 to week 12 in pooled SKYLIGHT 1 & 2 HRT-unsuitable excluding HRT-averse women

From	To

	0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: HRT: hormone replacement therapy; VMS: vasomotor symptoms.

Table 36: Average transition probabilities for placebo week 0 to week 12 in pooled SKYLIGHT 1 & 2 HRT-unsuitable women

From	To			
	0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: HRT: hormone replacement therapy; VMS: vasomotor symptoms.

Table 37: Average transition probabilities for placebo week 0 to week 12 in pooled SKYLIGHT 1 & 2 HRT-unsuitable excluding HRT-averse women

From	To			
	0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: HRT: hormone replacement therapy; VMS: vasomotor symptoms.

iii. Average health state transition probabilities for fezolinetant (week 0 to week 52) from SKYLIGHT 1 & 2 (HRT-unsuitable).

Table 38 and Table 39 present the average health state transition probabilities for fezolinetant for the period of week 0 to week 52 in the pooled SKYLIGHT 1 & 2 HRT-unsuitable and HRT-unsuitable excluding HRT-averse populations respectively. These average transition probabilities are included in the model for the period of week 0 to week 12, week 12 to week 24, week 24 to week 52 and week 52+ for fezolinetant.

Table 38: Average transition probabilities for fezolinetant week 0 to week 52 in pooled SKYLIGHT 1 & 2 HRT-unsuitable women

From	To

	0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: HRT: hormone replacement therapy; VMS: vasomotor symptoms.

Table 39: Average transition probabilities for fezolinetant week 0 to week 52 in pooled SKYLIGHT 1 & 2 HRT-unsuitable excluding HRT-averse women

From	To			
	0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: HRT: hormone replacement therapy; VMS: vasomotor symptoms.

iv. Average health state transition probabilities for fezolinetant (week 24 to week 52) from SKYLIGHT 1 & 2 (HRT-unsuitable).

These average health state transition probabilities are already provided in the model for both the pooled SKYLIGHT 1 & 2 HRT-unsuitable and SKYLIGHT 1 & 2 HRT-unsuitable excluding HRT-averse populations. Please find these repeated in Table 40 and Table 41.

Table 40: Average transition probabilities for fezolinetant week 24 to week 52 in pooled SKYLIGHT 1 & 2 HRT-unsuitable women

From	To			
	0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: HRT: hormone replacement therapy; VMS: vasomotor symptoms.

Table 41: Average transition probabilities for fezolinetant week 24 to week 52 in pooled SKYLIGHT 1 & 2 HRT-unsuitable excluding HRT-averse women

From	To			
	0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]

7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: HRT: hormone replacement therapy; VMS: vasomotor symptoms.

v. Average health state transition probabilities for fezolinetant (week 0 to week 52) from pooled DAYLIGHT and SKYLIGHT 1 & 2 (HRT-unsuitable).

Table 42 and Table 43Table present the average health state transition probabilities for fezolinetant for the period of week 0 to week 52 in the pooled DAYLIGHT and SKYLIGHT 1 & 2 HRT-unsuitable, and HRT-unsuitable excluding HRT-averse populations, respectively. These average transition probabilities are included in the model for the period of week 0 to week 12, week 12 to week 24, week 24 to week 52 and week 52+ for fezolinetant. It is noted that the pooled data set only includes data from DAYLIGHT up to week 24 due to the trial design.

Table 42: Average transition probabilities for fezolinetant week 0 to week 52 in pooled DAYLIGHT and SKYLIGHT 1 & 2 HRT-unsuitable women

From	To			
	0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: HRT: hormone replacement therapy; VMS: vasomotor symptoms.

Table 43: Average transition probabilities for fezolinetant week 0 to week 52 in pooled DAYLIGHT and SKYLIGHT 1 & 2 HRT-unsuitable excluding HRT-averse women

From	To			
	0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: HRT: hormone replacement therapy; VMS: vasomotor symptoms.

vi. Please provide a revised version of the model with sufficient flexibility to incorporate the average health state transition probabilities from part i) to v), and the option to switch between alternative sources used to estimate health state transition probabilities over time for both fezolinetant and placebo. Please signpost the changes made to the model.

Adaptations have been made to the CEM to include the average transition probabilities outlined in B4)a)i-v at the appropriate timepoints. Additions to the model have been highlighted in green.

Health-related quality of life utility values

B5. PRIORITY. Please provide details of the numbers of participants providing EQ-5D data at each time point (baseline and each 4-weekly cycle) for each treatment arm (fezolinetant and placebo arm) in the trials (reported separately for DAYLIGHT and SKYLIGHT 1, 2) to inform the health state utility values used in the model.

These data are presented in Table 44 and Table 45.

Table 44: EQ-5D Observations: DAYLIGHT

	DAYLIGHT		DAYLIGHT excluding HRT-averse	
Timepoint	**Fezolinetant**	**Placebo**	**Fezolinetant**	**Placebo**
Baseline	[redacted]	[redacted]	[redacted]	[redacted]
Week 4	[redacted]	[redacted]	[redacted]	[redacted]
Week 12	[redacted]	[redacted]	[redacted]	[redacted]
Week 16	[redacted]	[redacted]	[redacted]	[redacted]
Week 24	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: HRT: hormone replacement therapy; NA: not applicable

Table 45: Number EQ-5D Observations: Pooled SKYLIGHT 1&2

	Pooled SKYLIGHT 1&2 HRT-unsuitable		Pooled SKYLIGHT 1 & 2 HRT-unsuitable excluding HRT-averse	
Timepoint	**Fezolinetant**	**Placebo**	**Fezolinetant**	**Placebo**
Baseline	[redacted]	[redacted]	[redacted]	[redacted]
Week 4	[redacted]	[redacted]	[redacted]	[redacted]
Week 12	[redacted]	[redacted]	[redacted]	[redacted]
Week 24	[redacted]	NA	[redacted]	NA
Week 52	[redacted]	NA	[redacted]	NA

Abbreviations: HRT: hormone replacement therapy; NA: not applicable.

B6. Please clarify whether missing EQ-5D data was imputed and, if appropriate, please provide details on the methods used.

No imputation was performed for missing EQ-5D data. Patients who did not have an evaluation on a scheduled visit were excluded from the analysis for the timepoint of that visit.

B7. Please provide comprehensive details on the methods used to estimate health state utility values, including results of model fit.

EQ-5D-5L data were obtained from the phase 3 trials (DAYLIGHT, SKYLIGHT 1 & 2) and EQ-5D index values were estimated using the mapping function developed by the Decision Support Unit (Hernández Alava et al. 2017) and the "EEPRU dataset" (Hernández Alava *et al.*, 2020) as

recently recommended by NICE. In DAYLIGHT, EQ-5D-5L data were obtained from each available visit (Week 0, 4, 12, 16, 24) for patients who received fezolinetant 45 mg or placebo.

The generalized estimating equation (GEE) model was used to estimate health state utility values for the VMS frequency health states to account for repeated measures. In this model, utility values were the dependent variable and VMS frequency health state was the independent variable. GEE is a marginal or population-average model that generates population-averaged estimates of the parameters. Conceptually, the GEE model was equivalent to taking the average of the patients' utility scores as long as they belong to the same health state. An exchangeable working correlation structure was used, which assumed a constant correlation. Since the GEE model is robust to the misspecification of the working correlation structure, the parameter estimates are still unbiased even though the working correlation structure is wrong. The variance-covariance matrix was estimated by a robust sandwich variance estimator. The statistical fit of the GEE utilities was assessed using Quasi Information Criterion (QIC). The QIC results for all utilities included in the model are presented in Table 46. Utility for the cessation of VMS health state was estimated using descriptive data of the utility of any patients who reported a VMS frequency of 0 during any visit, including at baseline.

Table 46: Results of Utility Model Fits

Trial data	Population	Treatment Specific	QIC
DAYLIGHT	HRT-unsuitable	No	[redacted]
		Fezolinetant	[redacted]
		Placebo	
	HRT-unsuitable excluding HRT-averse	No	[redacted]
		Fezolinetant	[redacted]
		Placebo	
Pooled SKYLIGHT 1 & 2	HRT-unsuitable	No	[redacted]
		Fezolinetant	[redacted]
		Placebo	
	HRT-unsuitable excluding HRT-averse	No	[redacted]
		Fezolinetant	[redacted]
		Placebo	
Pooled DAYLIGHT and SKYLIGHT 1 & 2	HRT-unsuitable	No	[redacted]
	HRT-unsuitable excluding HRT-averse	No	[redacted]

Abbreviations: HRT: hormone replacement therapy; QIC: Quasi Information Criterion; SE: standard error; VMS: vasomotor symptoms.

B8. PRIORITY. Please provide the utility values (mean and standard error) for each VMS frequency health state and cessation of VMS separately for the fezolinetant and placebo arms of the trials (i.e., utility values reported separately by treatment arm) for DAYLIGHT and SKYLIGHT 1, 2.

These data are presented in Table 47 and Table 48. It is noted that the treatment effect on utilities is different between the DAYLIGHT and pooled SKYLIGHT 1 and 2 trials. Given this discrepancy further analyses were performed, in which the treatment-specific utilities from DAYLIGHT were adjusted for baseline utility and an interaction term between baseline utility and

treatment. The results of these analyses are presented in Table 49 and show that the discrepancy between the utility values in each treatment arm in DAYLIGHT is in part due to the differences in the baseline utility of patients in each treatment arm. When this is adjusted for the utilities for both treatment arms in the DAYLIGHT trial are similar.

It should be noted that treatment-specific utilities for pooled SKYLIGHT 1 and 2 are based on EQ-5D observations during the 12-week double-blind period, as after this placebo patients were crossed-over to the fezolinetant arm.

Table 47: Treatment specific health-state utilities DAYLIGHT

Health state	HRT-unsuitable				HRT-unsuitable excluding HRT-averse			
	Fezolinetant		Placebo		Fezolinetant		Placebo	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Cessation of VMS	[redacted]							
0 ≤ VMS Frequency < 2	[redacted]							
2 ≤ VMS Frequency < 7	[redacted]							
7 ≤ VMS Frequency < 9	[redacted]							
VMS Frequency ≥ 9	[redacted]							

Abbreviations: HRT: hormone replacement therapy; SE: standard error; VMS: vasomotor symptoms.

Table 48: Treatment specific health-state utilities pooled SKYLIGHT 1 & 2

Health state	HRT-unsuitable				HRT-unsuitable excluding HRT-averse			
	Fezolinetant		Placebo		Fezolinetant		Placebo	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Cessation of VMS	[redacted]							
0 ≤ VMS Frequency < 2	[redacted]							
2 ≤ VMS Frequency < 7	[redacted]							
7 ≤ VMS Frequency < 9	[redacted]							
VMS Frequency ≥ 9	[redacted]							

Abbreviations: HRT: hormone replacement therapy; SE: standard error; VMS: vasomotor symptoms.

Table 49. Treatment specific health-state utilities, adjusted for baseline utility, from DAYLIGHT

Health state	HRT-unsuitable			
	Fezolinetant		Placebo	
	Mean	SE	Mean	SE
Cessation of VMS	[redacted]	[redacted]	[redacted]	[redacted]
0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
QIC Statistical Fit	[redacted]			

Abbreviations: HRT: hormone replacement therapy; SE: standard error; VMS: vasomotor symptoms.

B9. If available, please provide information about other symptoms (over and above VMS) experienced by patients in each VMS frequency category by treatment arm in DAYLIGHT and SKYLIGHT 1, 2.

Astellas do not have any information regarding symptoms other than VMS that would be experienced by patients in each treatment arm of the DAYLIGHT and SKYLIGHT 1,2 trials.

B10. PRIORITY. Adjustment of health state utility values used in the model.

a) **Please justify why the health state utility values based on EQ-5D data from DAYLIGHT were adjusted with opinion from clinical experts. When providing justification, please refer to the NICE health technology evaluations manual 2022, which provides recommendations for circumstances when the EQ-5D is considered to not be an appropriate measure. Specifically the manual requires a case to be made on why the EQ-5D is inappropriate. The manual further details the need to provide qualitative empirical evidence on the lack of content validity of the EQ-5D, showing that key dimensions of health are missing. A flowchart (Fig 4.1, NICE manual) details the hierarchy of preferred HRQoL methods. Further detailed recommendations are provided in Rowen, et al. (2020).[66]**

The decision to adjust the utility values for the VMS frequency health states was driven by insights from three UK clinical experts and four patient representatives. The feedback was unanimous that the utility values derived using EQ-5D data from DAYLIGHT appeared overly optimistic compared with the real-world patient population they encounter. One clinical expert estimated a utility value of 0.810 for the 0 ≤ VMS Frequency < 2 health state and all clinical experts suggested that utilities for the 7 ≤ VMS Frequency < 9 and VMS Frequency ≥ 9 health states should be lower. Based on this feedback, a 5% reduction in utility was applied to these

health states relative to the values reported in the DAYLIGHT trial, with the adjusted values receiving full support from patient representatives.

These insights are further supported by evidence from the literature, which indicates that generic measures like EQ-5D may not fully capture the range of health impacts associated with menopause and its treatment.[67, 68] In a study involving 3,721 women aged 50–69, HRT was shown to significantly alleviate VMS, improve sexual functioning, and reduce sleep problems.[67] However, these benefits did not correspond to statistically significant differences in overall HRQoL as measured by EQ-5D/VAS scores, despite the reductions in hot flushes, night sweats, aching joints, insomnia, and vaginal dryness. This discrepancy underscores the limitations of EQ-5D in reflecting improvements in menopausal symptoms. Given this evidence, the adjusted utility values were considered appropriate for the model base case.

In line with the NICE health technology evaluations manual (2022), which advises the use of qualitative empirical evidence when a generic measure like EQ-5D may lack content validity, adjustments to the utility values were made. Clinical and patient expert feedback was leveraged to ensure that the utility values in the model more accurately represent the real-world impact of VMS on quality of life.

b) Please provide comprehensive information on the methods used to make the health state utility adjustments.

As outlined in the previous response, the adjustments to health state utility values were primarily informed by feedback from UK clinical experts and patient representatives. To reflect this feedback effectively in the model, we implemented a combination of expert-elicited values and pragmatic assumptions. Specifically, one clinical expert provided an elicited utility value of 0.810 for the 0 ≤ VMS Frequency < 2 health state. For the higher-frequency health states (7 ≤ VMS Frequency < 9 and VMS Frequency ≥ 9), a pragmatic assumption was made by applying a uniform 5% reduction to the utility values observed in the DAYLIGHT trial data. This approach was based on expert feedback suggesting that these values appeared overly optimistic in the context of real-world patient experiences.

Underlying evidence used in model

B11. PRIORITY. The baseline distribution and transition probabilities used in the model are based on data from the postmenopausal population of DAYLIGHT, with a mean age of 54.5 years and a baseline minimum average of 7 moderate to severe events of VMS per day during the last 10 days, while the mean age used in the model is 51 years based on the average age of menopause onset in the UK.

a) Please explain why the baseline distribution and transition probabilities would not be expected to differ for a population at menopause onset

compared to the population of DAYLIGHT who have ≥7 moderate to severe VMS per day and a higher mean age, not at menopause onset.

As noted in response to A6 b)i, there is clear evidence in the literature which indicates that the physiological mechanism underlying VMS is consistent across the menopause transition, from perimenopause to postmenopause, and is not exclusive to postmenopausal women. As such, there is no physiological rationale to suggest that the transition probabilities would differ between a population at menopause onset when compared with the DAYLIGHT trial population. It is worth noting that the subgroup analysis of age category 1 (<55 years versus ≥55 years) of DAYLIGHT (Figure 19 of Document B) shows consistent results across both subgroups for the primary endpoint of VMS frequency, which is suggestive of consistent treatment effects across older and younger subgroups of the trial.

For the baseline distribution, it is worth noting that the distribution observed in DAYLIGHT was influenced by inclusion criteria mandated by guidance from the FDA on clinical trials for VMS. As such, because only women with ≥7 moderate to severe VMS during the last ten days were enrolled, it is expected that women with higher VMS frequency are more likely to seek treatment when compared with a population at menopause onset and as such patients would be skewed toward more severe health states. However, without direct data for this population, it is challenging to determine the exact distributional differences, and modelling would be required to confirm the direction and magnitude of any differences.

That said, the baseline age of 51 years was chosen to align with the mean age of people seen in UK clinical practice based on UK clinical expert opinion and advice by the NHS. Importantly, scenario analyses using mean age data from DAYLIGHT and pooled SKYLIGHT and DAYLIGHT were provided in Table 76 of Document B that provide reassurance that this modelling assumption has minimal impact on the ICER.

b) **Please justify this considering the estimates of natural history of postmenopausal women in the UK from the structured expert elicitation (SEE) exercise, which suggests a much lower proportion of ≥7 moderate to severe VMS per day at each of the three time points of year 1, 3 and 6 in the absence of treatment (Table 61 of CS); noting that even with further clinical validation of the SEE estimates, the proportions in the higher VMS frequency categories of ≥7 VMS per day were much lower (60%) at year 1 compared to the estimate of 100% used in the model at baseline.**

It is necessary for the proportion of ≥7 VMS per day to be ~100% at baseline, as this was a trial design criterion specified by the FDA guidance on trial design for clinical studies on VMS.[23] This baseline requirement was integral to the trial design, shaping the evidence available to inform both the licence and the cost-effectiveness model.

However, Astellas acknowledge that this created a disconnect with the SEE data, which were not anchored on the trial baseline distribution when elicited. Consequently, as noted in Section B.3.3.2.3 of Document B, three UK clinical experts were consulted to validate the SEE estimates when anchored to the DAYLIGHT baseline. Experts highlighted that the substantial resolution in

symptoms at Year 1 was clinically implausible in the absence of active treatment and ultimately preferred the Year 6 estimates, as they were more reflective of the natural progression of VMS observed in their clinical practice. As such, in alignment with expert opinion, the base case model calculates the distributions at each four-week cycle from Year 0 up to Year 6, assuming a linear change from baseline to Year 6 in each VMS frequency health state. Accordingly, the model estimates that 86.41% of menopausal people with moderate to severe VMS would occupy the higher VMS frequency categories of ≥7 VMS per day at Year 1, which is consistent with feedback from UK clinical experts and supported by patient representatives who validated this modelling approach.

B12. In the CS, it states that no relevant studies on the natural history of VMS for the target population were identified in the clinical SLR (p154 of CS). Please clarify whether a SLR was performed to identify studies of natural history of VMS? If so, please provide details of the search terms used and the inclusion/exclusion criteria applied in the search strategy. If not, please perform a SLR to identify any relevant evidence on the natural history of VMS and report the findings from the published literature. These findings should be incorporated into the model, if appropriate.

While an SLR was not specifically conducted to identify studies on the natural history of VMS, Astellas are aware of two studies that report on the natural history of VMS for the target population of this appraisal. However, neither study provides data on the specific outcomes used in the model (i.e., moderate to severe VMS frequency per day).[69, 70] As such, Astellas undertook a structured expert elicitation exercise to generate robust estimates of the natural history of VMS that could be suitably incorporated in the model.

The variation in estimates obtained from the SEE reflects the considerable uncertainty surrounding the natural progression of VMS, which is highly individualised. This underscores the inherent challenges in eliciting natural history estimates to inform the cost-effectiveness modelling. Given this heterogeneity and the individualised nature of VMS, an SLR would not be expected to resolve this uncertainty.

B13. The model assesses the cost-effectiveness of fezolinetant in a postmenopausal population. Please justify why the median duration of VMS of 7.4 years from Avis et al. (2015)[64] is used in the model when this represents the median total VMS duration for all participants in the SWAN study, i.e., all women transitioning from pre-menopause or early perimenopause into late post-menopause, instead of the median duration of 3.4 years for postmenopausal women, reported in Avis et al. (2015).

The model uses a median VMS duration of 7.4 years from Avis et al. (2015), which represents the total VMS duration across all participants in the SWAN study.[64] This choice aligns with the licensed indication for fezolinetant, which includes both perimenopausal and postmenopausal women. Since treatment is intended for women across the menopausal transition, it is important

to model VMS duration in a way that reflects the experience of both perimenopausal and postmenopausal individuals, rather than limiting to postmenopausal duration alone.

In the SWAN study, the majority of participants (87%, n=1257) were either peri- or postmenopausal, supporting the relevance of the 7.4-year duration for our model population. Although the use of a median duration of 3.4 years, specific to postmenopausal women, would result in a more favourable ICER, this alternative scenario is not presented in line with the response to Question A6.

B14. PRIORITY. Treatment discontinuation rates used in model.

a) Please provide the treatment discontinuation rates separately for i) week 0-12 and ii) week 12-24 from DAYLIGHT, for the fezolinetant and placebo arms of the trial.

These data are presented in Table 50.

Table 50: Treatment discontinuation rates in DAYLIGHT

	HRT-unsuitable				HRT-unsuitable excluding HRT-averse			
	Fezolinetant		Placebo		Fezolinetant		Placebo	
	Week 0-12	Week 12-24	Week 0-12	Week 12-24	Week 0-12	Week 12-24	Week 0-12	Week 12-24
Discontinuation events / number of patients at risk	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Per 4-week cycle probability of discontinuation	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: HRT: hormone replacement therapy.

b) Please provide the treatment discontinuation rates separately for i) week 0-12 and ii) week 12-24 from DAYLIGHT, by the fezolinetant and placebo arms of the trial.

Question 14)b was included in error and therefore has not been addressed, as discussed with NICE and the EAG.

Scenario analysis based on SKYLIGHT 1 & 2 (HRT-unsuitable)

B15. PRIORITY. Please provide cost-effectiveness results for a scenario based on pooled estimates from SKYLIGHT 1 & 2 (HRT-unsuitable) separately from the pooled estimates of SKYLIGHT with DAYLIGHT. Specifically,

a) Please provide the transition probabilities (TPs) for pooled SKYLIGHT 1 & 2 (HRT-unsuitable) in each 4-weekly cycle, up to week 12 for placebo and up to week 52 for fezolinetant.

Table 51 and Table 52 present the per cycle health state transition probabilities for placebo for the period of week 0 to week 12 in the pooled SKYLIGHT 1 & 2 HRT-unsuitable and HRT-unsuitable excluding HRT-averse populations respectively. These per cycle transition probabilities are included in the model for the period of week 0 to week 12 for placebo.

Table 53 and Table 54 present the per cycle health state transition probabilities for fezolinetant for the period of week 0 to week 52 in the pooled SKYLIGHT 1 & 2 HRT-unsuitable and HRT-unsuitable excluding HRT-averse populations respectively. These per cycle transition probabilities are included in the model for the period of week 0 to week 52 for fezolinetant.

Table 51. Transition probabilities for placebo pooled SKYLIGHT 1 & 2 HRT-unsuitable

Week	From	To			
		0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
0–4	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
4–8	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
8–12	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: VMS: vasomotor symptoms.

Table 52. Transition probabilities for placebo pooled SKYLIGHT 1 & 2 HRT-unsuitable excluding HRT-averse

Week	From	To			
		0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
0–4	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]

	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
4–8	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
8–12	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: VMS: vasomotor symptoms.

Table 53. Transition probabilities for fezolinetant pooled SKYLIGHT 1 & 2 HRT-unsuitable

Week	From	To			
		0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
0–4	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
4–8	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
8–12	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
12–16	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
16–20	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
20–24	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
24–28	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]

	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
28–32	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
32–36	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
36–40	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
40–44	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
44–48	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
48–52	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: VMS: vasomotor symptoms.

Table 54. Transition probabilities for fezolinetant pooled SKYLIGHT 1 & 2 HRT-unsuitable excluding HRT-averse

Week	From	To			
		0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
0–4	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
4–8	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
8–12	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]

	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
12–16	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
16–20	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
20–24	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
24–28	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
28–32	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
32–36	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
36–40	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
40–44	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
44–48	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]
48–52	0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
	2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
	7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
	VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: VMS: vasomotor symptoms.

b) Please provide the baseline distribution for pooled SKYLIGHT 1 & 2 (HRT-unsuitable).

These data are presented in Table 55.

Table 55: Baseline distribution of people pooled SKYLIGHT 1 & 2

Source		0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
Pooled SKYLIGHT1 & 2 HRT-unsuitable	Number of people	[redacted]	[redacted]	[redacted]	[redacted]
	% of people	[redacted]	[redacted]	[redacted]	[redacted]
Pooled SKYLIGHT1 & 2 HRT-unsuitable excluding HRT-averse	Number of people	[redacted]	[redacted]	[redacted]	[redacted]
	% of people	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: HRT: hormone replacement therapy; VMS: vasomotor symptoms.

c) Please provide the treatment discontinuation rates for pooled SKYLIGHT 1 & 2 (HRT-unsuitable) for i) week 0-12; ii) week 12-24; and iii) week 0-24.

These data are presented in Table 56.

Table 56: Treatment discontinuation rates in pooled SKYLIGHT 1 & 2

	HRT-unsuitable				HRT-unsuitable excluding HRT-averse			
	Fezolinetant		Placebo		Fezolinetant		Placebo	
	Week 0-12	Week 12-24	Week 0-24	Week 0-12	Week 0-12	Week 12-24	Week 0-24	Week 0-12
Discontinuation events / number of patients at risk	17/287	11/270	28/287	28/297	[redacted]	[redacted]	[redacted]	[redacted]
Per 4-week cycle probability of discontinuation	2.01%	1.38%	1.70%	3.25%	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: HRT: hormone replacement therapy; NA: not applicable.

d) Please provide the utility values for pooled SKYLIGHT 1 & 2 (HRT-unsuitable).

These data are presented in Table 57.

Table 57: Health-state utilities pooled SKYLIGHT 1 & 2

Health state	HRT-unsuitable		HRT-unsuitable excluding HRT-averse	
	Mean	SE	Mean	Mean
Cessation of VMS	[redacted]	[redacted]	[redacted]	[redacted]
0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: SE: standard error; VMS: vasomotor symptoms.

e) Please provide a revised version of the model with sufficient flexibility to incorporate the information from part a) to d). Please signpost the changes made to the model.

The model has been updated to include a pooled SKYLIGHT 1 & 2 scenario based on the data provided in questions B15)a–d.

Section C: Textual clarification and additional points

None.

References

1. Astellas Data on File. UK Individual Expert Validation Meetings Report. 2024.
2. National Institute for Health and Care Excellence (NICE). Fezolinetant for treating vasomotor symptoms associated with the menopause (ID5071). Response to stakeholder organisation comments on the draft remit and draft scope. Available at: https://www.nice.org.uk/guidance/indevelopment/gid-ta11058 [Accessed: 8 November 2024].
3. Astellas Data on File. Exploratory Network Meta-Analysis.
4. National Institute for Health and Care Excellence (NICE). Menopause: diagnosis and management [NG23]. Available at: https://www.nice.org.uk/guidance/NG23 [Accessed: 8 November 2024].
5. Clinical Knowledge Summaries (CKS). Menopause Scenario: Managing women with menopause, perimenopause, or premature ovarian insufficiency – Non-hormonal Treatments. Available at: https://cks.nice.org.uk/topics/menopause/management/management-of-menopause-perimenopause-or-premature-ovarian-insufficiency/#non-hormonal-treatments [Accessed: 8 November 2024].
6. Astellas. UK HRT-unsuitable Consensus Statement.
7. National Institute for Health and Care Excellence (NICE). CKS Hormone Replacement Therapy (HRT). Available at: https://cks.nice.org.uk/topics/menopause/prescribing-information/hormone-replacement-therapy-hrt/. [Accessed: October 2024].
8. Electronic Medicines Compendium (eMC). Zumenon 1 mg Film-coated Tablets. Summary of Product Characteristics (SmPC).
9. Dacks PA, Krajewski SJ, Rance NE. Activation of neurokinin 3 receptors in the median preoptic nucleus decreases core temperature in the rat. Endocrinology 2011;152:4894-905.
10. Dacks PA, Rance NE. Effects of estradiol on the thermoneutral zone and core temperature in ovariectomized rats. Endocrinology 2010;151:1187-93.
11. Jayasena CN, Comninos AN, Stefanopoulou E, et al. Neurokinin B administration induces hot flushes in women. Sci Rep 2015;5:8466.
12. Krajewski SJ, Burke MC, Anderson MJ, et al. Forebrain projections of arcuate neurokinin B neurons demonstrated by anterograde tract-tracing and monosodium glutamate lesions in the rat. Neuroscience 2010;166:680-97.
13. Krajewski-Hall SJ, Blackmore EM, McMinn JR, et al. Estradiol alters body temperature regulation in the female mouse. Temperature (Austin) 2018;5:56-69.
14. Padilla SL, Johnson CW, Barker FD, et al. A Neural Circuit Underlying the Generation of Hot Flushes. Cell Rep 2018;24:271-277.
15. Mittelman-Smith MA, Krajewski-Hall SJ, McMullen NT, et al. Neurokinin 3 Receptor-Expressing Neurons in the Median Preoptic Nucleus Modulate Heat-Dissipation Effectors in the Female Rat. Endocrinology 2015;156:2552-62.
16. Mittelman-Smith MA, Williams H, Krajewski-Hall SJ, et al. Role for kisspeptin/neurokinin B/dynorphin (KNDy) neurons in cutaneous vasodilatation and the estrogen modulation of body temperature. Proc Natl Acad Sci U S A 2012;109:19846-51.
17. Rance NE, Dacks PA, Mittelman-Smith MA, et al. Modulation of body temperature and LH secretion by hypothalamic KNDy (kisspeptin, neurokinin

B and dynorphin) neurons: a novel hypothesis on the mechanism of hot flushes. Front Neuroendocrinol 2013;34:211-27.
18. Rance NE, McMullen NT, Smialek JE, et al. Postmenopausal hypertrophy of neurons expressing the estrogen receptor gene in the human hypothalamus. J Clin Endocrinol Metab 1990;71:79-85.
19. Rance NE, Young WS, 3rd. Hypertrophy and increased gene expression of neurons containing neurokinin-B and substance-P messenger ribonucleic acids in the hypothalami of postmenopausal women. Endocrinology 1991;128:2239-47.
20. Krajewski-Hall SJ, Miranda Dos Santos F, McMullen NT, et al. Glutamatergic Neurokinin 3 Receptor Neurons in the Median Preoptic Nucleus Modulate Heat-Defense Pathways in Female Mice. Endocrinology 2019;160:803-816.
21. Lambrinoudaki I, Armeni E, Goulis D, et al. Menopause, wellbeing and health: A care pathway from the European Menopause and Andropause Society. Maturitas 2022;163:1-14.
22. Baber RJ, Panay N, Fenton A. 2016 IMS Recommendations on women's midlife health and menopause hormone therapy. Climacteric 2016;19:109-50.
23. Food and Drug Administration (FDA). Guidance for Industry. Estrogen and Estrogen/Progestin Drug Products to Treat Vasomotor Symptoms and Vulvar and Vaginal Atrophy Symptoms — Recommendations for Clinical Evaluation. Available at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/estrogen-and-estrogenprogestin-drug-products-treat-vasomotor-symptoms-and-vulvar-and-vaginal-atrophy [Accesssed: 8 November 2024], 2003.
24. European Medicines Agency (EMA). GUIDELINE ON CLINICAL INVESTIGATION OF MEDICINAL PRODUCTS FOR HORMONE REPLACEMENT THERAPY OF OESTROGEN DEFICIENCY SYMPTOMS IN POSTMENOPAUSAL WOMEN. Available at: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-hormone-replacement-therapy-oestrogen-deficiency_en.pdf [Accessed: 8 November 2024].
25. European Medicines Agency (EMA). Vezoa Public Assessment Report. Available at: https://www.ema.europa.eu/en/documents/assessment-report/veoza-epar-public-assessment-report_en.pdf. [Accessed: 8 November 2024].
26. Medicines and Healthcare Products Regulatory Agency (MHRA). Veoza 45 mg film-coated tablets. Summary of Product Characteristics. Available at: https://mhraproducts4853.blob.core.windows.net/docs/5a51dd236244e9ab5bb6e3353ca74dd766e673a6 [Accessed: 8 November 2024].
27. Morga A, Ajmera M, Gao E, et al. Systematic review and network meta-analysis comparing the efficacy of fezolinetant with hormone and nonhormone therapies for treatment of vasomotor symptoms due to menopause. Menopause 2023:10.1097.
28. Astellas Data on File. Astellas UK 2024 Vasomotor Symptoms Associated (VMS) with Menopause Advisory Board Transcript report
29. Fraser GL, Obermayer-Pietsch B, Laven J, et al. Randomized Controlled Trial of Neurokinin 3 Receptor Antagonist Fezolinetant for Treatment of Polycystic Ovary Syndrome. J Clin Endocrinol Metab 2021;106:e3519-e3532.

30. Depypere H, Timmerman D, Donders G, et al. Treatment of Menopausal Vasomotor Symptoms With Fezolinetant, a Neurokinin 3 Receptor Antagonist: A Phase 2a Trial. J Clin Endocrinol Metab 2019;104:5893-5905.
31. Fraser GL, Lederman S, Waldbaum A, et al. A phase 2b, randomized, placebo-controlled, double-blind, dose-ranging study of the neurokinin 3 receptor antagonist fezolinetant for vasomotor symptoms associated with menopause. Menopause 2020;27:382-392.
32. Takamatsu K, Miki T, Miyazaki K, et al. Phase II study of fezolinetant for treatment of vasomotor symptoms associated with menopause in Japan. Climacteric 2024;27:389-397.
33. Lederman S, Ottery, F. D., Cano, A., Santoro, N., Shapiro, M., Stute, P., Thurston, R. C., English, M., Franklin, C., Lee, M. and et al. Fezolinetant for treatment of moderate-to-severe vasomotor symptoms associated with menopause (SKYLIGHT 1): a phase 3 randomised controlled study. Lancet 2023;401:1091–1102.
34. Johnson KA, Martin N, Nappi RE, et al. Efficacy and safety of fezolinetant in moderate to severe vasomotor symptoms associated with menopause: a phase 3 RCT. The Journal of Clinical Endocrinology & Metabolism 2023.
35. Neal-Perry G, Cano A, Lederman S, et al. Safety of Fezolinetant for Vasomotor Symptoms Associated With Menopause: A Randomized Controlled Trial. Obstet Gynecol 2023;141:737-747.
36. Ruan X, Bai W, Ren M, et al. Efficacy and safety of fezolinetant for moderate to severe vasomotor symptoms associated with menopause among women in East Asia: a phase 3 randomized study (MOONLIGHT I). J Int Med Res 2024;52:3000605241247684.
37. Yu Q, Ming F, Ma J, et al. Long-term safety of fezolinetant in Chinese women with vasomotor symptoms associated with menopause: the phase 3 open-label MOONLIGHT 3 clinical trial. J Int Med Res 2024;52:3000605241246624.
38. National Institute for Health and Care Excellence (NICE). Early and locally advanced breast cancer: diagnosis and management [NG101]. Available at: https://www.nice.org.uk/guidance/NG101 [Accessed: 8 November 2024].
39. Nappi RE SE, Todorova L, Rea C, Gemmen E, Schultz NM. Prevalence and quality-of-life burden of vasomotor symptoms associated with menopause: A European cross-sectional survey. Maturitas 2023;167.
40. Astellas. UK Clinican Estimates of HRT-Averse Population.
41. Morga A, Zimmermann L, Valluri U, et al. Validation and Application of Thresholds to Define Meaningful Change in Vasomotor Symptoms Frequency: Analysis of Pooled SKYLIGHT 1 and 2 Data. Advances in Therapy 2024;41:2845-2858.
42. Shapiro C.M. M SK, Hirschberg A L, Blogg M, Morga A, Martins K, Bouchard C,. Analysis of Responders in the Phase 3b DAYLIGHT Study: Treatment of Moderate-to-Severe Vasomotor Symptoms due to Menopause in Women Considered Unsuitable for Hormone Therapy. The Menopause Society (TMS) 2024 Annual Meeting.
43. Schultz NM, Morga A, Siddiqui E, et al. Psychometric evaluation of the PROMIS SD-SF-8b instrument in individuals experiencing vasomotor symptoms due to menopause. Health Qual Life Outcomes 2023;21:126.
44. Schultz NM, Morga A, Siddiqui E, et al. Psychometric Evaluation of the MENQOL Instrument in Women Experiencing Vasomotor Symptoms Associated with Menopause. Advances in Therapy 2024;41:2233-2252.

45. Douxfils J, Beaudart C, Dogné JM. Risk of neoplasm with the neurokinin 3 receptor antagonist fezolinetant. Lancet 2023;402:1623-1625.
46. Food and Drug Administration (FDA). Clinical review 216578Orig1s000. Available at: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2023/216578Orig1s000MedR.pdf. [Accessed: 8 November 2024].
47. Therapeutic Goods Administration (TGA). Australian Public Assessment Report for Veoza. Available at: https://www.tga.gov.au/sites/default/files/2024-05/auspar-veoza-240503.pdf. [Accessed: 8 November 2024].
48. Yoshida S, Shimo T, Takabatake K, et al. Expression of Neurokinin B Receptor in the Gingival Squamous Cell Carcinoma Bone Microenvironment. Diagnostics (Basel, Switzerland). Volume 11, 2021:1044.
49. Steinhoff MS, von Mentzer B, Geppetti P, et al. Tachykinins and their receptors: contributions to physiological control and the mechanisms of disease. Physiol Rev 2014;94:265-301.
50. Coveñas R, Muñoz M. Cancer progression and substance P. Histol Histopathol 2014;29:881-90.
51. Astellas Data on File. SKYLIGHT 1 Study Protocol.
52. Astellas Data on File. SKYLIGHT 2 Study Protocol.
53. European Medicines Agency (EMA). ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials . Available at: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-and-sensitivity-analysis-clinical-trials-guideline-statistical-principles-clinical-trials-step-5_en.pdf. [Accessed: 8 November 2024].
54. Astellas Data on File. DAYLIGHT Study Protocol.
55. Morga A, Latimer NR, Scott M, et al. Is Intention to Treat Still the Gold Standard or Should Health Technology Assessment Agencies Embrace a Broader Estimands Framework?: Insights and Perspectives From the National Institute for Health and Care Excellence and Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen on the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use E9 (R1) Addendum. Value Health 2023;26:234-242.
56. Astellas Data on File. DAYLIGHT Clinical Study Report.
57. Higgins JPT TJ, Chandler J, Cumpston M, Li T, Page MJ, Welch VA Cochrane Handbook for Systematic Reviews of Interventions version 6.5 (updated August 2024), 2023.
58. SIGN Filters. Available at: https://www.sign.ac.uk/what-we-do/methodology/search-filters/ [Accessed: 8 November 2024].
59. Strings attached: CADTH database search filters [Internet]. Ottawa: CADTH; 2016. Available from: https://www.cadth.ca/resources/finding-evidence/strings-attached-cadths-database-search-filters [Accessed 8 November 2024].
60. University of York Centre for Reviews and Dissemination. Search Strategies. NHS-EED. Available from: http://www.crd.york.ac.uk/crdweb/searchstrategies.asp#nhseedmedline [Accessed 8 November 2024].
61. Arber, M, Garzcia, S, Veale, T, Edwards, M, Shaw, A and Glanville, J. Performance of search filters to identify health state utility studies. Available

from: https://www.ispor.org/docs/default-source/ispor-19th-annual-european-congress/prm183--strong-u-arber-m-u-strong-garcia-s-veale-t-edwards-m-shaw-a-glanville-j-br-york-health-economics-consortium-york-uk.pdf?sfvrsn=8fedab18_0 [Accessed 8 November 2024].
62. Papaioannou D, Brazier J, Paisley S. Systematic searching and selection of health state utility values from the literature. Value in health 2013;16:686-695.
63. Astellas Data on File. PAG Engagement Meetings: Fezolinetant (Veoza™) HTA appraisals.
64. Avis NE, Crawford SL, Greendale G, et al. Duration of menopausal vasomotor symptoms over the menopause transition. JAMA Intern Med 2015;175:531-9.
65. NHS inform. Menopause. Available at: https://www.nhsinform.scot/healthy-living/womens-health/later-years-around-50-years-and-over/menopause-and-post-menopause-health/menopause. [Accessed: 8 November 2024].
66. Rowen D, Brazier J, Wong R, et al. Measuring and valuing health-related quality of life when sufficient EQ-5D data is not available. Report by the Decision Support Unit. Sheffield: Decision Support Unit, ScHARR, University of Sheffield, 2020:48.
67. Welton AJ, Vickers MR, Kim J, et al. Health related quality of life after combined hormone replacement therapy: randomised controlled trial. BMJ 2008;337:a1190.
68. Brazier JE, Roberts J, Platts M, et al. Estimating a preference-based index for a menopause specific health quality of life questionnaire. Health and Quality of Life Outcomes 2005;3:13.
69. Matthews KA, Chang Y, Brooks MM, et al. Identifying women who share patterns of reproductive hormones, vasomotor symptoms, and sleep maintenance problems across the menopause transition: group-based multi-trajectory modeling in the Study of Women's Health Across the Nation. Menopause 2021;28:126-134.
70. Mishra GD, Dobson AJ. Using longitudinal profiles to characterize women's symptoms through midlife: results from a large prospective study. Menopause 2012;19:549-555.

Appendices

Simulation analyses for inclusion of exploratory NMA results in the cost-effectiveness model

For paroxetine 7.5 mg (SSRI), the transition probabilities between VMS frequency health states at 4-week intervals from baseline to Week 52 were derived based on exploratory NMA results and baseline VMS frequency data from the DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 trials.

The exploratory NMA results from the fixed effects model were used and included 50,000 iterations of VMS frequency change from baseline at Week 12 for fezolinetant 45mg, SSRI, and placebo. Fezolinetant 45mg was considered as the benchmark when calculating the relative treatment effects among treatments to ensure the fezolinetant 45mg efficacy inputs are consistent with the observed data in the DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 trials. The data used for the exploratory NMA are summarised below:

NMA population:

- NMA results based on 12-week VMS frequency outcome (fixed effects, 50,000 iterations) from the DAYLIGHT trial only.
- Fezolinetant data used for simulations:
    - Week 0 to 24: DAYLIGHT.
    - Week 24 to 52: Pooled SKYLIGHT 1 and SKYLIGHT 2 (HRT-unsuitable subpopulation).

The detailed steps taken to derive the transition probability inputs for comparators are summarised below:

Step 1. Calculate treatment effects of comparator versus fezolinetant 45 mg

- For each of the 50,000 iterations, the treatment effect of SSRI versus fezolinetant 45mg at Week 12 were calculated by dividing the change in VMS frequency from baseline of comparators by the change in VMS frequency from baseline of fezolinetant 45mg:

$$ratio_{comparator\ versus\ fezo\ 45\ mg} = \frac{VMS\ frequency\ change\ from\ baseline_{\ comparator}}{VMS\ frequency\ change\ from\ baseline_{fezo\ 45\ mg}}$$

- Since the NMA results were only reported at week 12 due to data availability, the treatment effects at each 4-weekly interval up to 52 weeks were assumed to be the same as those at week 12.

Step 2. Calculate VMS frequency change from baseline at each 4-week interval up to 52 weeks, for fezolinetant 45 mg

- VMS frequency change from baseline at each 4-week interval of the fezolinetant 45mg treatment arm was calculated for each individual using the relevant trial data from the NMA population:

$$Observed\ VMS\ frequency\ change\ from\ baseline\ of\ fezo\ 45\ mg\ arm_{timepoint} = Observed\ VMS\ frequency\ of\ fezo\ 45\ mg\ arm_{timepoint} - Observed\ VMS\ frequency\ of\ fezo\ 45\ mg\ arm_{baseline}$$

$$timepoint = 4\ weekly\ intervals\ up\ to\ week\ 52$$

Step 3. Calculate VMS frequency change from baseline at each 4-week interval, for SSRI

- For each individual and across each of the 50,000 iterations, the change from baseline in VMS frequency for SSRI was calculated by multiplying the ratio of relative change of comparators and fezolinetant 45mg obtained in Step 1 by the observed VMS frequency change from baseline of fezolinetant 45mg at each timepoint obtained in Step 2:

$$simulated\ VMS\ frequency\ change\ from\ baseline_{comparator,timepoint} = ratio_{comparator\ vs.fezo\ 45mg} \times observed\ VMS\ frequency\ change\ from\ baseline\ of\ fezo\ 45mg\ arm_{timepoint}$$

$$timepoint = 4\ weekly\ intervals\ up\ to\ week\ 52$$

Step 4. Calculate VMS frequency at each 4-week interval, for SSRI

- For each individual and across each of the 50,000 iterations, the absolute VMS frequency at each 4-week interval were calculated by adding the simulated change from baseline in VMS frequency for SSRI obtained in Step 3 and the observed VMS frequency at baseline for fezolinetant 45mg from the relevant trial for the NMA population:

$$Simulated\ VMS\ frequency_{comparator,timepoint} = Simulated\ VMS\ frequency\ change\ from\ baseline_{comparator,timepoint} + Observed\ VMS\ frequency\ of\ fezo\ 45\ mg\ arm_{baseline}$$

$$timepoint = 4\ weekly\ intervals\ up\ to\ week\ 52$$

- The simulated VMS frequencies were then categorized into four health states, i.e., 0 ≤ VMS Frequency < 2, 2 ≤ VMS Frequency < 7, 7 < VMS Frequency < 9, and VMS Frequency ≥ 9.

Step 5. Calculate transition probabilities between VMS frequency states from baseline to 4-week intervals, up to 52 weeks, for SSRI

- VMS frequency matrices were created for the SSRI comparator based on the results in Step 4, i.e., the number of individuals in each health state at each timepoint. Three VMS frequency matrices were created from baseline to 4-week intervals, up to 52 weeks. Each frequency matrix is a four-by-four matrix with the four health states as the rows and columns. For example, in the frequency matrix for baseline to Week 4, cell 3,2 represents the number of individuals who belong to health state 7 ≤ VMS Frequency < 9 at baseline and transitioned to health state 2 ≤ VMS Frequency < 7 at Week 4.
- The transition probability matrices were derived based on the frequency matrices. Each entry in the transition probability matrices was computed by dividing each cell in the frequency matrix by the row sum. For example, cell 3,2 in the transition probability matrix is computed as cell 3,2 in the frequency matrix divided by the sum of row 3 in the same frequency matrix. If the row sum equals to zero, the diagonal entry in the transition probability matrix is assumed to be 1 with the rest of entries assumed to be 0 (i.e., remain in the current state and assume no transition to other health states).

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Fezolinetant for treating vasomotor symptoms associated with the menopause
Professional 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 the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

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 13 pages.

About you

1. Your name	[redacted]
2. Name of organisation	British Menopause Society (BMS)
3. Job title or position	[redacted]
4. Are you (please select Yes or No):	An employee or representative of a healthcare professional organisation that represents clinicians? **Yes** A specialist in the treatment of people with this condition? **Yes** A specialist in the clinical evidence base for this condition or technology? **Yes**
5a. Brief description of the organisation (including who funds it).	The BMS is the specialist authority for menopause and post reproductive health in the UK. The BMS educates, informs and guides healthcare professionals, working in both primary and secondary care, on menopause and all aspects of post reproductive health. Income is generated primarily through annual membership subscriptions and the provision of an annual education programme for healthcare professionals.
5b. Has the organisation received any funding from the manufacturer(s) of the technology and/or comparator products in the last 12 months? [Relevant manufacturers are listed in the appraisal matrix.] **If so, please state the name of manufacturer, amount, and purpose of funding.**	Astellas Pharma – £2,250 Exhibition table at BMS 33rd annual scientific conference, 27 & 28 June 2024
5c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No

The aim of treatment for this condition

6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	The aim of the treatment is to reduce vasomotor symptoms (VMS) that can cause considerable distress for peri- and post-menopausal women and impair quality of life. VMS are the commonest symptoms of the menopause in the UK and severe symptoms occur in 20-25% of women. All women who live long enough become menopausal and thus, a large number of women are affected. VMS can interrupt sleep and cause considerable distress and embarrassment for women. Successful treatment allows women to be able to continue to be active through mid-life and into older age.
7. 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.)	Usually the patient defines the response subjectively. It is rare that VMS are assessed objectively. A response that is statistically significantly greater than placebo (which is often 30-50%) is usually considered meaningful. The clinically meaningful response is pre-defined in the clinical studies, which assess the difference between pre and post treatment rather than looking at a specific percentage change or a decrease in number. A decreased impact on the patient herself is also important. However, if the VMS are decreased by as much as 60-70% then this is meaningful to most patients.
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Currently the mainstay of treatment is Hormone Replacement Therapy (HRT). It is extremely effective in treating menopausal symptoms including VMS and may decrease the number by 80% or more. Further, HRT treats other consequences of estrogen deficiency and is believed to have long-term benefits on the cardiovascular system and bone health. However, HRT is associated with some risks (e.g. increased incidence of breast cancer) and is contraindicated in women with a history of hormone dependant cancer. In addition, for some women it can be associated with unacceptable side effects, and some choose not to take it. Its long-term use and initiation in women over 65 is controversial. Non-hormonal alternatives are needed to give women more options. Some are available but are associated with significant side effects, which means that continuation is low (see NICE Guideline NG23 discussion of SSRIs). Available data suggests that Fezolinetant does not appear to be associated with significant side effects.

What is the expected place of the technology in current practice?

9. How is the condition currently treated in the NHS?	HRT SSRIs/ SNRIs Gabapentin and Pregabalin Oxybutinin Clonidine CBT
9a. Are any clinical guidelines used in the treatment of the condition, and if so, which?	NICE Guideline NG23 *Menopause: Diagnosis and management* (2015), update due for publication in November 2024. British Menopause Society publications – BMS joint guidelines, BMS Consensus Statement, BMS Tools for Clinicians. Other Menopause Society publications, e.g. International Menopause Society, North American Menopause Society, European Menopause and Andropause Society, Indian Menopause Society The Endocrine Society (US)
9b. 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.)	Yes, the pathway of care is well defined. Yes, there are differences in opinion across the NHS as to the importance of VMS (and other menopausal symptoms), and whether medical intervention is needed. Much work is being undertaken by the BMS, and other societies, to provide education for healthcare professionals so that if women choose to attend a healthcare professional, initially with their primary care team, they will be offered consistent advice. A large part of the variation in pathways is related to the fact that all women are affected differently with huge variation in type, severity, duration and impact of symptoms. Some women will self-manage with good effect, some will require support form their primary care team, and others will require referral to a specialist service. Specialist services, both NHS and private, exist across the UK.

9c. What impact would the technology have on the current pathway of care?	It would provide an effective, non-hormonal alternative to hormones that is likely more effective than SSRIs, CBT etc. The use of the latter is decreased, due to lack of availability of the service generally and self-administration is not appropriate for many women. Fezolinitent could also be used in those with hormone-dependent cancers e.g. breast cancer. This group is important as the treatment offered to decrease recurrence (tamoxifen in particular) is often associated with severe VMS that lead to discontinuation in a significant proportion of women. Although SSRIs are effective, they are associated with significant side effects and are not recommended for use in the NICE Menopause Guideline.
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Currently Fezolinetant is available privately but not on the NHS which leads to discrepancies in healthcare. HRT is generally quite cheap and readily available.
10a. How does healthcare resource use differ between the technology and current care?	
10b. In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	It is hoped that Fezolinetant could be initiated in primary care, as long as clear guidance is available on prescribing as well as the benefits and risks of its use.
10c. What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	None beyond extending and improving general menopause training.
11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes. It will mean that women who cannot take hormones or who do not wish to, can receive effective treatment for VMS. This will be invaluable for women with hormone sensitive cancer in whom HRT is contraindicated.

11a. Do you expect the technology to increase length of life more than current care?	Fezolinetant may increase length of life if women with breast cancer are able to continue their adjuvant endocrine therapy (tamoxifen or aromatase inhibitors) for the recommended length of time, rather than stopping them early due to worsening VMS, as unfortunately many women do.
11b. Do you expect the technology to increase health-related quality of life more than current care?	Yes as the current evidence supports the safety of the drug and there appears to be no reason why it should not be taken on the long-term if necessary. Most women only take hormones for a relatively short time and may get symptom recurrence when they stop
12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Women with hormone dependant cancer such as breast cancer who cannot take hormones. The efficacy would not be greater than the general population but the suitability would be much greater.

The use of the technology

13. 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.)	There is no reason why it should be more difficult to use and no concomitant medications would be needed. There would be no additional clinical requirements.

14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	There may be informal rules regarding a trial without therapy after a particular length of time. There is currently nothing to suggest that it should only be taken for a short time or that it will be required for the long term to prevent recurrence. There is no additional testing required at present.
15. 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?	Possibly but too early to say.
16. 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?	Treatment for a wider group of women.
16a. Is the technology a 'step-change' in the management of the condition?	Yes, using a medication, which works through a different pathway from previous options in this field.
16b. Does the use of the technology address any particular unmet need of the patient population?	Yes for those who cannot or prefer not to use hormones and have significant VMS.

17. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	Only minor side effects have been reported which should not impact greatly on QoL.

Sources of evidence

18. Do the clinical trials on the technology reflect current UK clinical practice?	Yes although women will be treated if they need to be even if they don't have the large numbers required for inclusion in the clinical trials. 2 or 3 night sweats can have more impact than a large number of day-time flushes but it is less easy to demonstrate change in a small number of flushes than with larger numbers.
18a. If not, how could the results be extrapolated to the UK setting?	
18b. What, in your view, are the most important outcomes, and were they measured in the trials?	Number of VMS and number of night sweats Impact of these on quality of life and ability to function on a daily basis
18c. If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	

18d. Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	Not as far as I know.
19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No
20. Are you aware of any new evidence for the comparator treatment(s) since the publication of NICE technology appraisal guidance [TAXXX]? [delete if there is no NICE guidance for the comparator(s) and renumber subsequent sections]	There is a NICE Guideline on the use of HRT currently being updated although the section on VMS is largely unchanged as there is little new evidence.
21. How do data on real-world experience compare with the trial data?	I do not think many doctors in the UK have seen enough patients to answer this although the patients report a good response and few side effects.

Equality

22a. Are there any potential equality issues that should be taken into account when considering this treatment?	Currently yes, as it is only available privately and most of the population cannot afford it.
22b. Consider whether these issues are different from issues with current care and why.	There are some equality issues with all Menopause care as there is still some variation in practice and in priority in providing services across the UK. In addition, there is variation across cultural and ethnic backgrounds around access to services

Topic-specific questions

23 [To be added by technical team at scope sign off. Note that topic-specific questions will be added only if the treatment pathway or likely use of the technology remains uncertain after scoping consultation, for example if there were differences in opinion; this is not expected to be required for every appraisal.] **if there are none delete highlighted rows and renumber below**	

Key messages

24. In up to 5 bullet points, please summarise the key messages of your submission.	• Fezolinetant is a promising new treatment for vasomotor symptoms with several placebo-controlled RCTs, currently all pharma funded. • The efficacy is statistically significantly greater than placebo. • It is non-hormonal and has none of the risks associated with HRT although it does not treat other hypoestrogenic symptoms as HRT does. • It can be used in those with estrogen-dependent cancer e.g. breast • There do not appear to be significant side effects and it is well tolerated by patients.

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

Fezolinetant for treating moderate to severe vasomotor symptoms associated with the menopause [ID5071]

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 note that the population we are seeking input on is specifically people with moderate to severe vasomotor symptoms associated with the menopause when HRT is not considered suitable.

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.

NICE National Institute for Health and Care Excellence

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 **Wednesday 19 February**. 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.

NICE National Institute for Health and Care Excellence

Part 1: Treating moderate to severe vasomotor symptoms associated with the menopause and current treatment options

Table 1 About you, aim of treatment, place and use of technology, sources of evidence and equality

1. Your name	Professor Waljit Dhillo
2. Name of organisation	Imperial College London
3. Job title or position	Professor of Endocrinology & Metabolism
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 moderate to severe vasomotor symptoms associated with the menopause? ☐ A specialist in the clinical evidence base for moderate to severe vasomotor symptoms associated with the menopause 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.	nil

8. What is the main aim of treatment for moderate to severe vasomotor symptoms associated with the menopause when HRT is not considered suitable? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability)	Improve symptoms and quality of life
9. What do you consider a clinically significant treatment response? (In the case of VMS, how would an improvement in vasomotor symptoms be measured [i.e. what scale] and what would be considered a significant response [i.e. number on that scale]?)	Frequency & severity measures on MENQOL
10. In your view, is there an unmet need for people and healthcare professionals in moderate to severe vasomotor symptoms associated with the menopause when HRT is not considered suitable? **Additionally, how would moderate to severe vasomotor symptoms be defined/determined in NHS practice to assess who is eligible for treatment? Would both severity of an individual episode and the frequency be taken into account?**	Yes Yes - frequency & severity
11. How is moderate to severe vasomotor symptoms associated with the menopause currently treated in the NHS when HRT is not considered suitable? • 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?	Poor other treatments No Offer a credible alternative to HRT

• Are non-hormonal pharmacological treatments used in clinical practice, such as anti-depressants [selective serotonin reuptake inhibitors (SSRIs) and serotonin and norepinephrine reuptake inhibitors (SNRIs)], clonidine, or anti-convulsants such as gabapentin and pregabalin? If so, to what extent. •	Yes but limited efficacy
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) • What investment is needed to introduce the technology? (for example, for facilities, equipment, or training)	the issue will be the cost of the drug secondary care or primary care with GP with a specialist interest as need liver function test monitoring nil
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?	No Yes
14. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Yes in patients with moderate to severe vasomotor symptoms associated with the menopause when HRT is not considered suitable

15. Will the technology be easier or more difficult to use for people 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 acceptability to people, or ease of use or additional tests or monitoring needed)	Oral tablet so no issues. need liver function test monitoring
16. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Yes if liver function test deteriorates
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	no
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? • 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 targeted population?	Yes novel treatment Yes in patients with moderate to severe vasomotor symptoms associated with the menopause when HRT is not considered suitable the technology provides an effective alternative

19. How do any side effects or adverse effects of the technology affect the management of the condition and a person's quality of life? Would you consider additional monitoring of specific adverse effects to be required?	need liver function test monitoring whilst on technology
20. Do the clinical trials on the technology reflect current UK clinical practice? **Related to this, does the population in the trial reflect the population that would have this treatment in the NHS? (note that the trials only included people who were postmenopausal, had at least 7 moderate to severe events per day, and excluded those with high blood pressure)** • 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?	Yes Frequency & severity of flushes need liver function test monitoring whilst on technology
21. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	no
22. How do data on real-world experience compare with the trial data?	favourably
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	no

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. Find more general information about the Equality Act and equalities issues here.	

NICE National Institute for Health and Care Excellence

Part 2: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

Effective technology for patients with moderate to severe vasomotor symptoms associated with the menopause when HRT is not considered suitable

need liver function test monitoring whilst on technology

Click or tap here to enter text.

Click or tap here to enter text.

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.

CONFIDENTIAL UNTIL PUBLISHED

External Assessment Group Report
Fezolinetant for treating vasomotor symptoms associated with the menopause [ID5071]

Produced by	Centre for Reviews and Dissemination (CRD) and Centre for Health Economics (CHE) Technology Assessment Group, University of York, Heslington, York, YO10 5DD
Authors	Mark Corbett, Research Fellow, CRD Natalia Kunst, Senior Research Fellow, CHE Mark Perry, Research Fellow, CRD Anqian Zhou, Research Fellow, CHE, Melissa Harden, Senior Information Specialist, CRD Connor Evans, Research Assistant, CRD Sofia Dias, Professor in Health Technology Assessment, CRD Claire Rothery, Professor of Health Economics, CHE
Correspondence to	Prof Claire Rothery, CHE, University of York
Date completed	05/12/2024 (updated 07/01/2025)

Source of funding

This report was commissioned by the NIHR Evidence Synthesis Programme as project number NIHR169592.

Declared competing interests of the authors

None.

Acknowledgements

We would like to thank Dr Sarah Glynne, MB BS BSc MSc MRCP MRCGP, The Portland Hospital, London, for her advice on clinical issues during this appraisal.

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:
Corbett M, Kunst N, Perry M, Zhou A, Harden M, Evans C, Dias S, Rothery C. Fezolinetant for treating vasomotor symptoms associated with the menopause [ID5071]: A Single Technology Appraisal. CRD and CHE Technology Assessment Group, University of York, 2024.

Contributions of authors
Please refer to the International Committee of Medical Journal Editors (ICMJE) Uniform Requirements for Manuscripts Submitted to Biomedical Journals see http://www.icmje.org/

Mark Corbett wrote the critique of the clinical effectiveness evidence and contributed to the critique of the decision problem and safety evidence. Natalia Kunst performed the critical review of the economic analyses, contributed to drafting Sections 1, 4, 5 and 6 of the report, co-led the economic analyses, and takes joint responsibility for the report. Mark Perry wrote the critique of the decision problem and safety evidence and contributed to the critique of the clinical effectiveness evidence. Anqian Zhou performed the critical review of the economic analyses, conducted the EAG additional analyses and contributed to drafting Sections 4, 5 and 6 of the report. Melissa Harden wrote the critique of the search strategies. Connor Evans contributed to the critique of the clinical effectiveness and safety evidence. Sofia Dias provided advice on statistical methods and report content, commented on drafts of the report and reviewed the whole report. Claire Rothery performed the critical review of the economic analyses, contributed to drafting Sections 1, 4, 5, 6 and 7 of the report, co-led the economic analyses, and takes joint responsibility for the report.

Note on the text
All commercial-in-confidence (CON) data have been redacted (black highlight [redacted]).

Table of Contents

Table of Tables

Table of Figures

List of abbreviations

AE	Adverse event
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
BMS	British Menopause Society
BNF	British National Formulary
CBT	Cognitive behavioural therapy
CE	Common effect
CHE	Centre for Health Economics
CI	Confidence interval
CNS	Central nervous system
CRD	Centre for Reviews and Dissemination
CrI	Credible interval
CS	Company submission
CSR	Clinical study report
DSA	Deterministic sensitivity analysis
DIC	Deviance information criterion
EAG	External Assessment Group
EMA	European Medicines Agency
EQ-5D	EuroQoL Group's 5 dimension questionnaire
EQ-5D VAS	EuroQoL Group's 5 dimension questionnaire visual analogue scale
ER	Extended release
FDA	The US Food and Drug Administration
FE	Fixed effect
FSH	Follicle-stimulating hormone
GEE	Generalized estimating equations
HCRU	Healthcare resource use
HR	Hazard ratio
HRQoL	Health-related quality of life
HRT	Hormone replacement therapy
HTA	Health Technology Assessment
ICER	Incremental cost-effectiveness ratio
INHB	Incremental net health benefit
LOCF	Last-observation-carried-forward
LS	Least squares
LYG	Life year gain
MAR	Missing at random
MCID	Minimum clinically important difference
MENQOL	The Menopause-specific quality of life questionnaire
MHRA	Medicines and Healthcare products Regulatory Agency
MI	Multiple imputation
MMRM	Mixed models repeated measures analyses
NICE	National Institute for Health and Care Excellence
NK3R	Neurokinin 3-receptor
NMA	Network meta-analysis
OMA	Office for Market Access
OR	Odds ratio

PfC	Points for clarification
PHQ-4	Patient health questionnaire-4
PROM	Patient reported outcome measure
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
PSSRU	the Personal Social Services Research Unit
QALY	Quality-adjusted life year
QIC	Quasi information criterion
QoL	Quality of Life
RCT	Randomised controlled trial
RE	Random effects
SAE	Serious adverse event
SD	Standard deviation
SE	Standard error
SEE	Structured expert elicitation
SLR	Systematic literature review
SNRI	Serotonin and norepinephrine reuptake inhibitors
SSRI	Selective serotonin reuptake inhibitor
STEER	Structured expert elicitation resources
SWAN	The Study of Women's Health Across the Nation
TBL	Total bilirubin
TEAE	Treatment emergent adverse event
UK	United Kingdom
ULN	Upper limit of normal
VMS	Vasomotor symptoms
WPAI-VMS	Work productivity and activity impairment questionnaire – vasomotor symptom domain

1 EXECUTIVE SUMMARY

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

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

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

1.1 Overview of the EAG's key issues

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

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

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

Table 1 List of key issues

ID	Summary of issue	Report sections
1	Exclusion of relevant comparator treatments from the decision problem	2.2.5, 2.3, 4.2.4
2	All trial continuous efficacy estimates are at high risk of bias due to missing outcome data	3.2.1.1
3	Applicability of the trial populations to NHS practice	3.2.1.2, 3.2.2.2
4	Uncertainty about fezolinetant's safety profile: risk of neoplasms and serious liver injury	3.2.3
5	Use of frequency of moderate to severe VMS to define the health states in the model	4.2.2
6	Arbitrary cut-off thresholds used to define moderate to severe VMS frequency health states in the model	4.2.2
7	The baseline distribution with a moderate to severe VMS frequency limit of greater than seven per day	4.2.3
8	Treatment effects used in the model	4.2.6
9	Uncertainty in the natural history of VMS in the absence of treatment	4.2.6.2
10	The modelled health state utility values are highly uncertain	4.2.8

The key differences between the company's preferred assumptions and the EAG's restricted base case assumptions are: (i) an increase in the placebo effect from DAYLIGHT from week 12 (company's base case) to week 24 (EAG restricted base case) for no active treatment; although the EAG considers it more appropriate to incorporate relative treatment effects in the model rather than absolute effects from the separate arms of the trials; (ii) median duration of VMS of 3.4 years for the modelled population of postmenopausal people rather than the company's base case duration of 7.4 years that includes perimenopausal people and results in 40% of people still experiencing VMS at the end of the modelled time horizon of 10 years; and (iii) health state utility values based on the pooled DAYLIGHT and SKYLIGHT 1 & 2 (HRT-unsuitable) data (UK) rather than DAYLIGHT only (company's base case) which showed higher utility values in the placebo arm than the fezolinetant arm. The selection of changes made to the EAG's restricted base case comes with the caveat that the EAG is unable to address the key structural uncertainties in the company's base case analysis.

1.2 Overview of key model outcomes

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

Overall, the technology is modelled to affect QALYs by:

- Increasing the proportion of postmenopausal people moving to lower moderate to severe VMS frequency health states over time, which are associated with improved health-related quality of life compared to higher VMS frequency health states.
- Use of health-related quality of life (HRQoL) utility values from the DAYLIGHT trial adjusted with the opinion of one clinical expert to ensure that higher frequencies of moderate to severe VMS are associated with lower HRQoL utility values.

Overall, the technology is modelled to affect costs by:

- Lower health care resource use associated with occupancy of improved health states, where health states with higher moderate to severe VMS frequency are associated with higher costs.

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

- The baseline VMS frequency distribution.
- Adjustment for placebo effect from the trials.
- Use of Year 1, 3 and 6 SEE estimates for natural history rather than Year 6 estimates only.

- Utility values based on data from the pooled SKYLIGHT 1 and 2 trials that are not adjusted for clinical opinion.

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

Issue 1 Exclusion of relevant comparator treatments from the decision problem

Report section	2.2.5, 2.3, 4.2.4
Description of issue and why the EAG has identified it as important	The company stated that non-hormonal pharmacological treatments (such as anti-depressants, clonidine or gabapentin) were not appropriate comparator treatments. Although only clonidine is licensed for treating VMS, several non-hormonal pharmacological treatments are used in the NHS and many have been evaluated in RCTs. The company considered 'no active treatment' as the sole relevant comparator, but it is likely that menopausal people presenting with moderate to severe VMS and deemed HRT-unsuitable would be offered non-hormonal pharmacological treatments as opposed to 'nothing' in NHS clinical practice because these treatments have demonstrated some efficacy compared to placebo in clinical trials.
What alternative approach has the EAG suggested?	In addition to no active treatment, other non-hormonal pharmacological treatments, including anti-depressants (such as selective serotonin reuptake inhibitors [SSRIs] and serotonin-norepinephrine reuptake inhibitors [SNRIs]), clonidine, and anti-convulsants (such as gabapentin and pregabalin), and non-pharmacological treatments should be considered as relevant comparators. The company should provide relative efficacy estimates and cost-effectiveness evaluation of fezolinetant compared to these treatment options and sufficient flexibility in the model to incorporate evidence from the NMA. The EAG requested that the company undertake a NMA comparing fezolinetant with non-hormonal pharmacological treatments used in the NHS and adapt the model to incorporate relative effects. The company undertook an 'exploratory' NMA, but included only one comparator (paroxetine). The model was not updated to allow for incorporation of relative effects.
What is the expected effect on the cost-effectiveness estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	The EAG identified a published NMA - which was sponsored by the company - which compared a broader range of non-hormonal pharmacological treatments with fezolinetant. Its results indicated that there is little evidence to suggest that fezolinetant has a clinically-meaningful benefit over SSRIs, SNRIs and gabapentin for reducing the frequency of moderate to severe vasomotor symptoms, and that there is no evidence to suggest any increased benefit (from taking fezolinetant) for reducing the severity of vasomotor symptoms. Sufficient flexibility in the model is needed to incorporate evidence from the NMA.

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

Issue 2 All trial continuous efficacy estimates are at high risk of bias due to missing outcome data

Report section	3.2.1.1
Description of issue and why the EAG has identified it as important	The EAG believes the continuous trial outcomes of all three fezolinetant efficacy trials are at high risk of bias, with the bias favouring fezolinetant. The EAG believes this bias is a consequence of using over-optimistic methods when dealing with missing trial data; these were assumed to be missing at random (MAR, in which the missing data were considered similar to the treatment group mean). The EAG considers that for most patients with missing data, the value of the missing outcome data is likely to be related to the reason it is missing (rather than being missing randomly). The EAG has concerns about the company's reporting clarity on reasons for treatment discontinuation e.g. not explicitly reporting how many patients discontinued due to lack or loss of efficacy. The EAG also notes that data from the SKYLIGHT trial active treatment extension periods suggest that patients with missing data up to week 12 (of the placebo-controlled period) did not continue to participate in the extension period. This bias is important because the use of the MAR assumption in the company's analyses does not take account of the fact that the treatment effect is likely to attenuate across the cohort of patients with missing data.
What alternative approach has the EAG suggested?	The EAG would have liked to have seen the impact on results of analyses which used more conservative assumptions about missing data. Although the company performed an analysis which used a more conservative assumption for *some* of the missing data (i.e. discontinuations due to having a treatment-emergent adverse event) this only related to a small number of the participants with missing data (so its impact on results was always likely to be minimal).
What is the expected effect on the cost-effectiveness estimates?	An increase in the ICER for fezolinetant relative to no treatment
What additional evidence or analyses might help to resolve this key issue?	See above

Issue 3 Applicability of the trial populations to NHS practice

Report section	3.2.1.2, 3.2.2.2
Description of issue and why the EAG has identified it as important	The fezolinetant RCTs enrolled a much narrower population than would be seen in NHS practice. Perimenopausal people were not recruited, nor were participants with: chronic diseases, elevated blood pressure, or patients with smaller numbers of moderate to severe VMS events per day (e.g. < 7/day) at baseline. Given that a large number of these patients were excluded, it is important to understand whether these subgroups will benefit from fezolinetant to the same extent as the trial cohorts. The EAG is particularly concerned about the lower baseline VMS frequency subgroup exclusions because pooled SKYLIGHT trial efficacy estimates for improvements in VMS frequency appear to be driven by the subgroup of patients with around 10 or more VMS events per day.
What alternative approach has the EAG suggested?	None, given the evidence currently available.
What is the expected effect on the cost-effectiveness estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	Considering the baseline VMS frequency subgroup effect seen in the pooled SKYLIGHT trial analyses, the EAG is concerned that the company did not investigate whether a similar effect was seen in the DAYLIGHT trial, given that this could be an important effect modifier.

Issue 4 Uncertainty about fezolinetant's safety profile: risk of neoplasms and serious liver injury

Report section	3.2.3
Description of issue and why the EAG has identified it as important	Although the company asserts that fezolinetant has a similar safety profile to placebo, the EAG identified independent analyses of fezolinetant trial data which suggest a significantly higher incidence of neoplasms in participants taking fezolinetant. The EAG is also aware that the FDA issued a warning in September 2024 about a rare occurrence of serious liver injury with the use of fezolinetant for VMS. The EAG acknowledges that there is uncertainty about any association between fezolinetant and neoplasms. However, the EAG considers it is important to raise awareness on these issues since: i) they relate to analyses not covered by the CS ii) there is potential for liver function tests to be required for patients taking fezolinetant in the future (depending on the outcome of impending discussions with the MHRA) iii) these issues may need to be taken into account by clinicians and patients when making treatment choices.
What alternative approach has the EAG suggested?	N/A
What is the expected effect on the cost-effectiveness estimates?	Expected to have minimal impact on cost-effectiveness results.
What additional evidence or analyses might help to resolve this key issue?	The EAG considers these safety concerns may be valid until further data suggest otherwise.

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

Issue 5 Use of frequency of moderate to severe VMS to define the health states in the model

Report section	4.2.2
Description of issue and why the EAG has identified it as important	In the model, VMS frequency was used to define the structure of the model. A frequency of seven moderate to severe VMS per day represents the baseline cut-off to start receiving fezolinetant in the trials. However, in the model it is assumed that postmenopausal people who experience a decrease in the frequency of moderate to severe VMS after treatment keep receiving the treatment, even if they experience only 0 to < 2 moderate to severe VMS per day. The EAG is concerned that the model structure is not adequately capturing the impact of fezolinetant on the severity of moderate to severe VMS, which was a key secondary endpoint of DAYLIGHT and a co-primary endpoint of SKYLIGHT 1 and SKYLIGHT 2
What alternative approach has the EAG suggested?	An alternative model structure and definition of health states that more accurately reflects the decision problem and incorporates severity of moderate to severe VMS as an important trial outcome.
What is the expected effect on the cost-effectiveness estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	Exploration of alternative model structures or health states that also takes into account severity of moderate to severe VMS.

Issue 6 Arbitrary cut-off thresholds used to define moderate to severe VMS frequency health states in the model

Report section	4.2.2
Description of issue and why the EAG has identified it as important	The company assumed a frequency of seven moderate to severe VMS per day, which was used in the trial eligibility criteria, as a fixed cut-off threshold to define model health states. The EAG notes that the utility values for frequency between 7 to < 8 and 8 to < 9 are higher than the utility values between 6 to < 7. The EAG notes that among the explored models, several other models appear to be associated with more favorable p-values, but the information provided about the GEE models is too scarce for the EAG to make a more detailed assessment. In general, the EAG is also concerned about the use of utility values to define VMS frequency health states because the EQ-5D utility values are likely to reflect changes in quality of life due to other menopause symptoms over and above moderate to severe VMS and the EAG notes that the utility values in the placebo arm were higher than the fezolinetant arm in DAYLIGHT. Consequently, the EAG expresses several concerns with this approach and considers the chosen cut-off thresholds arbitrary.
What alternative approach has the EAG suggested?	An alternative approach to define health states that more accurately reflects the decision problem and incorporates severity of moderate to severe VMS.
What is the expected effect on the cost-effectiveness estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	Exploration of alternative model structures or health states that adequately reflect the severity of VMS as defined in NHS practice.

Issue 7 The baseline distribution with a moderate to severe VMS frequency limit of greater than seven per day

Report section	4.2.3
Description of issue and why the EAG has identified it as important	The baseline distribution based on the DAYLIGHT trial is assumed to be a minimum average of seven moderate to severe events of VMS per day in the absence of treatment, which the EAG does not consider be an appropriate baseline to represent postmenopausal people likely to receive fezolinetant in NHS primary care.
What alternative approach has the EAG suggested?	The EAG believes that the baseline distribution should be closer to the elicited values at one year for the natural history of postmenopausal women in the UK because a significant reduction in VMS frequency would not be expected after one year in the absence of treatment. The EAG performed scenario analyses, Scenarios 1a and 1b, where alternative assumptions for the baseline distribution are assumed, while keeping the year 6 SEE estimates for the natural history as per the company's base case.
What is the expected effect on the cost-effectiveness estimates?	EAG Scenarios 1a and 1b had the largerst impact on the cost-effectiveness results, increasing the ICER for fezolinetant vs. no treatment. This is driven by a substantial decrease in costs for the no treatment arm but also greater increase in QALYs associated with no treatment compared to fezolinetant.
What additional evidence or analyses might help to resolve this key issue?	An understanding of the eligibility criteria to receive fezolinetant in NHS primary care.

Issue 8 Treatment effects used in the model

Report section	Section 4.2.6
Description of issue and why the EAG has identified it as important	The model did not include relative treatment effects for fezolinetant that adjust for the placebo effect observed in the trials. Instead, the company used the absolute changes from baseline in the frequency of moderate to severe VMS observed in the fezolinetant arm of the trials, which are not adjusted for the placebo effect from week 12 onwards in the model. The implication of this assumption is that the placebo effect is lost arbitrarily from week 12 in the model for no active treatment and reverts to the natural history of VMS, while fezolinetant continues to receive the absolute changes (including a placebo effect) in 4-weekly cycles observed in the separate fezolinetant arm of DAYLIGHT up to week 24 and the pooled absolute changes observed in the fezolinetant arm of SKYLIGHT 1 and 2 from week 24 to week 52, without any adjustment for placebo effect in the trials.
What alternative approach has the EAG suggested?	The EAG considers it more appropriate to derive estimates of relative treatment effect for fezolinetant vs. placebo as per EAG request at clarifications (question B3). However, given that the company did not address the EAG's concern and chose their base case approach to apply the absolute changes in 4-weekly cycles, the placebo effect should be extrapolated for a longer time horizon. The EAG performed Scenario 2 where alternative assumptions for placebo effect are assumed.
What is the expected effect on the cost-effectiveness estimates?	Extending the placebo effect over a longer time horizon (Scenarios 2a, 2c, 2d) worsens the cost-effectiveness of fezolinetant compared to no treatment.
What additional evidence or analyses might help to resolve this key issue?	Deriving estimates of relative treatment effect for fezolinetant vs. placebo as per EAG clarification question B3.

Issue 9 Uncertainty in the natural history of VMS in the absence of treatment

Report section	Section 4.2.6.2
Description of issue and why the EAG has identified it as important	The EAG is concerned about the credibility of the results of the structured expert elicitation (SEE) exercise for informing the natural history of moderate to severe VMS over time, when a baseline VMS frequency distribution with a minimum average of seven moderate to severe VMS per day is used in the

	model. The EAG believes that the baseline distribution should be closer to the elicited values at one year for the natural history of postmenopausal women in the UK and the rapid reduction in the frequency of moderate to severe VMS between baseline and one year in the absence of treatment leads to the cost-effectiveness of fezolinetant relative to no active treatment being highly sensitive to the estimates of natural history used in the model. The approach used by the company to overcome this concern involved discarding the SEE estimates at years 1 and 3 and only use the estimates elicited at year 6 in the base case analysis, with a linear change from the baseline distribution of DAYLIGHT up to year 6 applied. The EAG does not consider this approach satisfactory because it highlights a lack of credibility of the SEE estimates.
What alternative approach has the EAG suggested?	The EAG assessed the impact of this assumption with Scenario 3, which is split into four separate scenarios, 3a, 3b, 3c and 3d, where alternative assumptions used to inform the natural history of moderate to severe VMS are used.
What is the expected effect on the cost-effectiveness estimates?	The cost-effectiveness results for the scenario analyses demonstrate that the alternative options have a material impact on the ICER for fezolinetant relative to no active treatment, which highlights how sensitive the cost-effectiveness results are to the natural history assumed in the model.
What additional evidence or analyses might help to resolve this key issue?	Provide more robust estimates of natural history in the absence of treatment as it represents a key driver of cost-effectiveness of fezolinetant relative to no treatment.

Issue 10 The modelled health state utility values are highly uncertain

Report section	Section 4.2.8
Description of issue and why the EAG has identified it as important	The company provided scarce information on how the final GEE model to estimate health-related quality of life was selected making it difficult for the EAG to assess the appropriateness of the final model. Furthermore, the EAG is concerned that the utility values from the DAYLIGHT trial for all health states of the model were higher in the placebo arm compared to the fezolinetant arm. In contrast, the utility values from the SKYLIGHT 1 and 2 trials were higher in the fezolinetant arm than in the placebo arm but these values were relatively close to the age-adjusted utility values for women in the general population. The EAG also considers the adjustment of the utility values with the opinion of one clinical expert to be highly uncertain and not an appropriate approach that aligns with the NICE reference case. Consequently, given these concerns, the modelled health state utility values are highly uncertain.
What alternative approach has the EAG suggested?	The EAG assessed the impact of the use of different utility values through Scenario 5, which is split into three scenarios, 5a, 5b and 5c, where alternative assumptions for the utility values were used.
What is the expected effect on the cost-effectiveness estimates?	The cost-effectiveness results for these scenario analyses demonstrate that the alternative options have a negative impact on the ICER for fezolinetant relative to no active treatment.
What additional evidence or analyses might help to resolve this key issue?	Provide more robust information on how utility values were estimated. Assess alternative models where moderate to severe VMS is not based on a fixed frequency threshold used to reflect severity of VMS.

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

No other key issues identified.

1.7 Summary of EAG's preferred assumptions and resulting ICER

The EAG presents a restricted base case. It should be noted that the EAG is unable to address the key structural uncertainties in the company's base case analysis because the company have not addressed the EAG's concerns at points for clarification, namely:

- The need to incorporate relative treatment effects in the model to adjust for the placebo effect observed in the trials.
- The need to incorporate a baseline distribution of moderate to severe VMS frequency that reflects that seen in the NHS for menopausal people who are likely to receive treatment with fezolinetant.
- Structural concerns regarding the health states used in the model, where arbitrary cut-off thresholds are used to define moderate to severe VMS frequency health states in the model and the concern that frequency of moderate to severe VMS is not commonly used to define severity of VMS in NHS clinical practice.
- No active treatment is unlikely to be the sole relevant comparator to fezolinetant in the NHS.
- Uncertainty about the natural history in the absence of treatment, which is a key driver of the cost-effectiveness of fezolinetant relative to no treatment.
- Absence of data for perimenopausal people, with cost-effectiveness of fezolinetant only assessed in postmenopausal people.

Therefore, a number of important uncertainties remain unresolved to determine an EAG preferred base case. Table 2 and Table 3 summarise the EAG's restricted base case and additional scenarios on the EAG's restricted base case, respectively.

Table 2 Cumulative cost-effectiveness results for the EAG's restricted base case (deterministic analysis)

Scenario #	Name	Option	Inc. Costs	Inc. QALYs	ICER, /QALY	% change of company's base-case ICER
	Company's base case results	Fezolinetant	-	-	-	-
		No Treatment	£1,199.49	0.116	£10,364.17	-
2a	Increase the placebo effect from week 12 to week 24	Fezolinetant	-	-	-	-
		No Treatment	£1,241.87	0.107	£11,621.78	12%
2a+4a	Median duration of VMS of 3.4 years	Fezolinetant	-	-	-	-
		No Treatment	£915.25	0.087	£10,496.77	1%
2a+4a+5b (EAG restricted base case)	Utility values based on pooled DAYLIGHT and SKYLIGHT 1 & 2 (HRT-unsuitable) data (UK)	Fezolinetant	-	-	-	-
		No Treatment	£915.25	0.056	£16,470.01	59%

Footnote: Probabilistic analysis not presented due to the structural uncertainties in the EAG's restricted base case.

Table 3 Cost-effectiveness results for alternative assumptions on the EAG's restricted base case (deterministic analysis)

Scenario #	Name	Option	Inc. Costs	Inc. QALYs	ICER, /QALY
	EAG restricted base case (2a+4a+5b)	Fezolinetant	-	-	-
		No Treatment	£915.25	0.056	£16,470.01
1b	Use Year 1 clinical validation estimates for baseline distribution	Fezolinetant	-	-	-
		No Treatment	£1,111.22	0.035	£31,304.74
1b+3d	Use Year 3 and 6 clinical validation estimates for natural history	Fezolinetant	-	-	-
		No Treatment	£1,119.46	0.037	£30,433.64

Footnote: Probabilistic analysis not presented due to the structural uncertainties in the EAG's restricted base case.

2 INTRODUCTION AND BACKGROUND

2.1 Introduction

This report presents a critique of the company's submission to NICE on the clinical and cost-effectiveness of fezolinetant (Veoza®) for treating moderate to severe vasomotor symptoms associated with menopause.

Fezolinetant was authorised by the European Medicines Agency (EMA) on 7th December 2023 and licensed by the Medicines and Healthcare products Regulatory Agency (MHRA) on 14th December 2023 for the treatment of moderate to severe vasomotor symptoms associated with menopause.

2.2 Background

2.2.1 Vasomotor symptoms

The company describes vasomotor symptoms (VMS) in the company submission (CS), section B.1.3.1, which is briefly summarised below.

VMS involves hot flushes (also known as hot flashes) and night sweats. The hot flushes are described as 'transient periods of intense heat', occurring mostly around the head, neck, chest and upper back, and sometimes accompanied by sweating and flushing of the skin. This may be followed by chills, palpitations and anxiety. Night sweats result from nocturnal hot flushes and may adversely affect sleep. Hot flushes and night sweats may occur multiple times per day. The CS reports that a mean of 17 hot flushes and 11 night sweats are experienced weekly, citing Hunter, 2012[1] and Thurston, 2008.[2] However, the reported 17 hot flushes and 11 night sweats per week only relate to the data from Hunter, 2012.[1] The data in Thurston, 2008 [2] show a slightly lower frequency at 13.5 hot flushes per week and 6.2 night sweats per week.

VMS usually begins in the perimenopause, which is defined on the NHS website[3] as the period when symptoms of menopause occur but periods have not yet stopped. Perimenopause ends at the menopause – the point in time when a period has not occurred for 12 months. VMS may continue into the postmenopause period (any time after 12 months of amenorrhoea) alongside other symptoms associated with menopause such as sleep disturbance, depression, mood changes and urogenital symptoms. The mean duration of VMS is reported as between four years [4] and 7.4 years[5] with a mean of 4.5 years during postmenopause,[5] and there may be a tendency for symptoms to worsen over time. VMS onset and progression may largely be due to changes in the levels of oestrogen in the perimenopausal and postmenopausal stages, or to other mechanisms, relating to neurotransmitter signalling and vascular processes.

The CS reports that about 80% of postmenopausal people are believed to be affected by VMS, and that in the UK, the annual prevalence of moderate to severe VMS is approximately 21%. The prevalence figure of 21% is based on a retrospective database study[6] of women aged 40–65 years registered at GP clinics. The study authors assumed that any patients on the database would have moderate/severe VMS because a patient would probably not visit their GP for mild symptoms. This prevalence figure may therefore be different to that which would apply when defining moderate/severe using the EMA grading (see section 2.2.4.1). The company also cite moderate/severe VMS prevalence figures of 25-38%, though only the 25%[7] figure was found in the references cited. Predisposing factors may include smoking, alcohol consumption, increased body mass index, lower educational level and anxiety. Ethnicity may also affect prevalence.

2.2.2 Burden of disease

The company provide an account of the burden of disease (CS, section B1.3.2), which has been summarised below.

The CS cites evidence that supports an association between VMS and reductions in quality of life,[8] difficulty with activities of daily living,[9 10] poorer sleep and daytime functioning,[10-12] depression,[13]absenteeism,[10]decreased workplace productivity,[14 10] relationship problems[15], decreased social activities,[15] and increased use of healthcare resources.[16] However, other evidence cited by the company does not demonstrate what the company suggests.

For example, the company states how VMS causes reduced quality of life or depressed mood, but the cited studies[1, 2] were designed to demonstrate how mood and other factors affect how bothersome VMS are, and were not designed to show the *effects* of VMS on such factors. Likewise, the review by Woods (2005)[17] is referenced as evidence that VMS causes an impact on relationships, involving less intimacy, a loss of closeness, arguments, and tension. However, whilst an effect on intimacy is supported, the other effects do not appear to be demonstrated in the article.

In terms of social and work-related effects, a reference to back up a statement that VMS has a negative socioeconomic impact is a webpage that focussed on general menopause symptoms, and does not specifically refer to the effects of VMS on socioeconomic variables.[18] In the same way, Hardy, 2018[19] is cited in order to support the claim that VMS leads to reduced work performance, greater workplace stress and intention to leave the labour force. However, the survey carried out by the authors only shows a link between VMS at work and intention to stop working, whilst specifically stating that VMS does not affect work outcomes.

Although healthcare resource costs are stated to be 'considerable', based on consideration of VMS prevalence and reduction in quality of life, there are no empirical data provided that demonstrate the actual costs.

Clinical advice to the EAG suggests that patients rarely visit their primary provider solely for VMS, and that other symptoms of the perimenopause or postmenopause are more likely to prompt a consultation. According to the updated NG23,[20] these symptoms include mood changes, musculoskeletal symptoms, urogenital symptoms, and sexual disorders. Although some of these may occur secondary to VMS, they can also occur independently of VMS.

2.2.3 Fezolinetant

Fezolinetant is an oral neurokinin 3-receptor (NK3R) antagonist. Its function is to block binding of neurokinin B to the KNDy neuron, and thereby help to moderate stimulation of the thermoregulatory centre of the brain. Figure 1 in the CS (section B.1.2) describes the mechanism.

Fezolinetant 45mg daily, by mouth, is licenced and approved for use in people with moderate to severe VMS. No investigations or additional tests are required prior to use. Absolute contraindications are hypersensitivity to the active substance or its excipients, concomitant use of moderate/strong CYP1A2 inhibitors, or known/suspected pregnancy.

2.2.4 Clinical pathway of care

The NICE clinical guideline on the menopause (NG23) provides recommendations for the management of VMS in the UK. The NICE guideline on early and locally advanced breast cancer (NG101) provides additional recommendations for people with breast cancer.

2.2.4.1 Diagnosis and classification

Diagnosis of VMS associated with the menopause is made clinically in primary care. This diagnosis may be made during the perimenopause (before the menopause in people aged 45 years or older if there are changes in the menstrual cycle alongside the VMS) or the postmenopause (after 12 months of amenorrhea, without use of hormonal contraception).[20] Serum follicle-stimulating hormone (FSH) measurements may be necessary to confirm postmenopause if there are atypical symptoms, if aged 40-45 with menopause symptoms, or aged < 40 years with suspected premature ovarian insufficiency.

The EMA[21] defines three grades of VMS: mild VMS involves a sensation of heat without sweating, moderate VMS involves a sensation of heat with sweating but continued ability to engage in activity, and serious VMS involves a sensation of heat with sweating, where activity has to stop.

VMS may also be classified by frequency, or by a combination of frequency and severity. The company suggests the following approach for calculating VMS severity by combining frequency and severity (p52, CS):

$$\frac{([\text{number of mild VMS/day} \times 1] + [\text{number of moderate VMS/day} \times 2] + [\text{number of severe VMS/day} \times 3])}{\text{Total number of daily (or weekly) mild/moderate/severe VMS}}$$

However, this classification implies an ordering and differential impact across mild, moderate and severe VMS which may not be valid in practice. For example, it is implied that moderate VMS has double the impact of mild, and that severe VMS has 50% more impact than moderate. These differentials have not been validated and may be misleading. Neither this approach, nor a similar method, was used in the NICE guideline (NG23).[22]

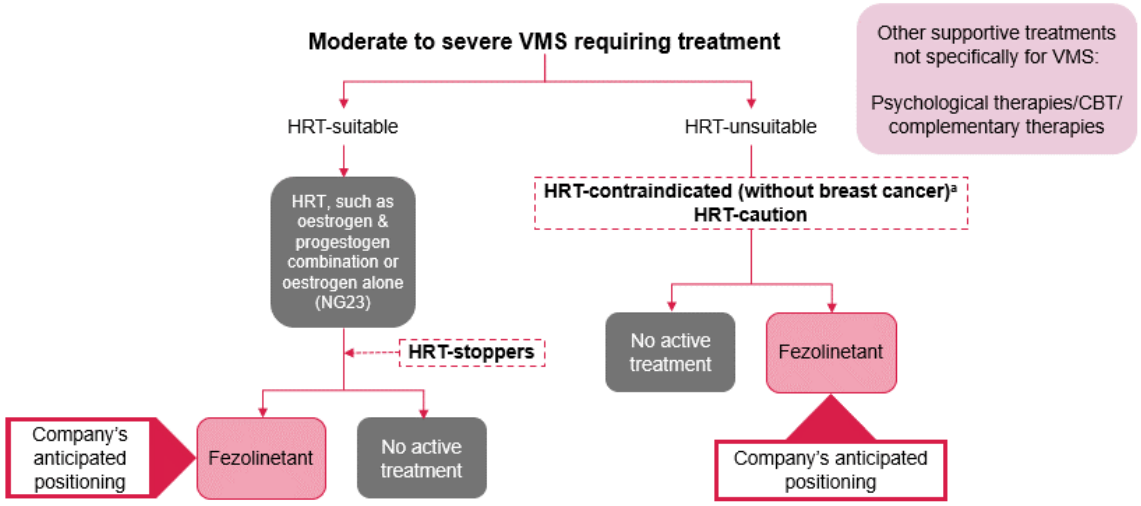
2.2.4.2 Current standard of care treatment

The company agrees with the NG23 recommendation that hormone replacement therapy (HRT) should be the only first-line treatment for VMS. In the absence of fezolinetant, the company believes that people for whom HRT is unsuitable, or people who have stopped HRT for any reason, would receive no active treatment.

2.2.5 Intended positioning of fezolinetant

Figure 1 summarises the company's anticipated positioning of fezolinetant in the treatment pathway. For people with moderate to severe VMS deemed suitable for HRT, HRT would be the first-line treatment. The company foresees that fezolinetant would be the only active treatment provided to those that have stopped HRT for any reason. For people with moderate to severe VMS deemed unsuitable for HRT (because of HRT contraindications, HRT cautions, or because people are HRT-averse) the company suggests that fezolinetant would be the only active treatment provided. However, other non-hormonal pharmacological treatments (such as anti-depressants, clonidine or gabapentinoids) are considered as comparators in the NG23 guideline, the NICE scope and, according to EAG clinician advice, are still part of NHS clinical practice.

Figure 1. Anticipated treatment pathway and proposed positioning of fezolinetant. Reproduced from CS figure 4.



[a] Fezolinetant is not licensed for use in women with breast cancer, or those being treated for breast cancer. Note that the definition of 'HRT unsuitable' omits 'HRT-averse', although people who are HRT-averse should be included in this category. This omission in the figure was due to a misunderstanding by the company during scoping, and the company has since stated that people who are HRT averse are included in the HRT-unsuitable category. Abbreviations: CBT: cognitive behaviour therapy; HRT: hormone replacement therapy; VMS: vasomotor symptoms.

The company stated that non-hormonal pharmacological treatments in the NICE scope (anti-depressants, clonidine or gabapentin/pregabalin) were not appropriate alternative treatments for those people unsuitable for HRT or who have stopped HRT for any reason, because NG23[22] recommended that anti-depressants and clonidine should not be used as first-line treatments for VMS. The EAG do not agree that this validates omission of these treatments as comparators. In NICE guideline NG23, the recommendation for antidepressants and clonidine not to be prescribed at first-line was based on the clear superiority of HRT as a treatment for VMS. This is identical to the company's own decision to place fezolinetant after HRT in patients for whom HRT is suitable. The NICE recommendation does not state that the non-hormonal pharmacological treatments should not be prescribed when HRT is unsuitable. The EAG considers that non-hormonal pharmacological treatments appear to be at a similar position in the treatment pathway as fezolinetant.

In addition, the company stated that all the non-hormonal pharmacological treatments except clonidine are unlicensed for the treatment of VMS, and their use is 'sporadic'. The EAG do not agree that this is a sufficient reason to exclude these treatments. Although antidepressants and gabapentinoids are not licensed for VMS, clinical advice to the EAG confirms that the pharmacological non-hormonal treatments in the NICE scope (clonidine, selective serotonin reuptake inhibitors (SSRIs)/ serotonin and norepinephrine reuptake inhibitors (SNRIs) and gabapentinoids) are commonly prescribed in the UK to about 1 in 5 people with VMS,[23] and will therefore be a more appropriate comparator than 'no active treatment' for estimations of efficacy and cost-effectiveness.

The company also notes that there is a lack of robust evidence supporting these treatments. However, the EAG disagrees with this assertion. All the non-hormonal pharmacological treatments have been

tested in randomised controlled trials (RCTs).[24] Reviews of such RCTs[25-28] have generally demonstrated some efficacy over placebo.

To further support their case for the omission of the non-hormonal pharmacological comparators, the company refers to the British Menopause Society (BMS) comments[29] on the draft NICE scope as having "*stressed the unmet need for a routinely reimbursed (i.e., through the National Health Service [NHS]) non-hormonal treatment option, that is both more efficacious and tolerable than the non-licensed non-hormonal therapies that may currently be prescribed to menopausal people for the treatment of VMS"*. The EAG do not think this provides sufficient rationale for omission. The BMS comments[29] state that non-hormonal treatments are not well-known to health care professionals but also that the NICE draft scope is appropriate, and reiterates the NICE scope comparator list (even adding oxybutynin as an additional comparator).

The efficacy and safety of the non-hormonal pharmacological treatments is regarded by the company and a group of clinical experts as 'less favourable' than that of fezolinetant. The EAG does not agree with this statement. In the original CS, the company assumes better safety and efficacy for fezolinetant compared to the non-hormonal pharmacological NICE comparators based on clinical expert advice alone. To fully evaluate fezolinetant's efficacy compared to other non-hormonal therapies, direct or indirect comparisons based on a full systematic literature review (SLR) and (network) meta-analysis (NMA) need to be provided.

In response to EAG requests for further clarification on why the non-hormonal pharmacological comparators were omitted from the decision problem, the company reiterated the expert opinion and provided evidence from an 'exploratory' Bayesian NMA[30] conducted by the company in January 2024 that considered 15 RCTs. These comprised the 14 trials identified by the company SLR (see section 3.1.5) with the addition of DAYLIGHT (which had not been in the company's original SLR). However, of the 15 trials considered, only five studies were included in the NMA. These were the three fezolinetant trials[31-33] and two paroxetine trials.[34] Of the remaining 10 studies identified by the company SLR, but not included in the NMA, eight were excluded for reasons that the EAG considered inappropriate (see section 3.1.5 for further details). The NMA suggested that fezolinetant was more effective than paroxetine in terms of the outcome of change in frequency of moderate/severe VMS.

However, having lower efficacy is not a suitable basis for excluding interventions as comparators, and the company has not provided convincing evidence that the other comparators are less effective and less safe than fezolinetant. The EAG therefore thinks that the non-hormonal pharmacological treatments listed in the NICE scope should be comparators in this submission and does not agree that

'no active treatment' in either the HRT-suitable or HRT-unsuitable groups is the only appropriate comparator to fezolinetant.

The company also does not regard non-pharmacological treatments such as cognitive behavioural therapy (CBT) as a central element of management, citing pre-publication reports from the updated NG23 guideline[35] that the committee view CBT as an added option rather than a routine treatment. The EAG agrees that CBT is probably not a relevant comparator, based on clinical advice to the EAG that CBT for VMS is rarely prescribed on the NHS, and is difficult to access, expensive, and time-consuming for the patient. It should be noted that the published 2024 update of the NG23 guideline[20] states that CBT should be considered as an option for VMS, "*in addition to HRT, or for people for whom HRT is contraindicated or for those who prefer not to take HRT*", although it also stresses the difficulty in accessing CBT services, and reports that the recommendations are based on low to very low quality evidence.

2.2.6 Equality considerations

The CS describes equality considerations in section B.1.4. It is noted that VMS may vary in prevalence and severity between ethnicities, with greater prevalence in black and Hispanic people, and a longer duration in people of African and Caribbean descent. People in semi-skilled or unskilled manual work may have greater difficulty with symptoms, and it is reported that 'working class' people may have more severe symptoms. There may also be disparities in access to specialist menopause services across locations.

No further equality considerations have been identified by the EAG.

2.2.7 Other relevant appraisals in progress

Elinzanetant is another treatment for VMS that is due to be evaluated by NICE in ID6359 (the appraisal is currently 'awaiting development').

2.3 Critique of company's definition of decision problem

A summary and critique of the decision problem addressed in the CS is provided in Table 1. The following sections provide a more detailed critique of the two aspects of the decision problem that differed from the NICE scope.

Population

The decision problem population is narrower than the NICE scope population. The NICE scope population includes all people with moderate or severe VMS, which equate to the licensed population, but the company restricts the decision problem population to people for whom HRT is unsuitable.

The EAG agrees with the company's placing of fezolinetant as secondary to HRT, as clinical advice to the EAG reiterates the company's view that HRT is the gold standard treatment for VMS. Therefore, focussing the decision problem on HRT stoppers, and people where HRT is unsuitable (due to contraindications, cautions or being HRT-averse) is reasonable.

The company define people unsuitable for HRT as those with HRT-contraindications, those with HRT cautions, those that have stopped HRT for any reason, or those who are HRT-averse (see final paragraph of this section) based on the UK HRT-Unsuitable Consensus Statement,[36] which elicited expert opinion on the populations that may be eligible for fezolinetant, largely based on their presumed unsuitability for HRT. The HRT contraindications described in the fezolinetant DAYLIGHT trial comprise undiagnosed vaginal bleeding, a history of breast cancer or oestrogen dependent tumours, liver disease, pregnancy, arterial thromboembolic disease, venous thrombophilic disorder, hypersensitivity to oestrogen and progesterone therapy or any of the excipients or porphyria. HRT cautions in the DAYLIGHT trial include diabetes mellitus, hyperlipidaemia, current smoking, migraine, obesity, systemic lupus erythematosus, epilepsy and family history of breast cancer in the first degree relative or a mutation of breast cancer gene (BRCA1 and BRCA2).

Clinical advice to the EAG suggests that whilst these contraindications or cautions would relate to orally administered, synthetic HRT formulations, most do not apply to transdermal 17β-oestradiol (which may have equivalent efficacy to oral HRT[24]). This is supported by two systematic reviews, which suggest that the risks of venous thromboembolism and possibly deep vein thrombosis are not elevated with transdermal HRT.[37, 38] Clinical advice to the EAG suggests that of these contraindications or cautions, only a history of oestrogen-receptor positive tumours would be a contraindication to transdermal 17β-oestradiol and body-identical progesterone. Therefore, the HRT-unsuitable population may be smaller than the company supposes, potentially comprising only those who are HRT averse, HRT stoppers and people with a history of oestrogen-receptor positive tumours. In response to clarification on the cautions and contraindications to HRT (question A2) the company stated that, in the absence of clear evidence, provision of HRT should be based on shared decision making between doctor and patient, weighing up the benefits and harms. Therefore, the EAG suggests that although the eligible population for fezolinetant (or its comparators such as non-hormonal pharmacological treatments) is possibly smaller than that suggested by the company, the extent of the eligible population that are 'HRT-unsuitable' will need to be made on a case-by-case basis, taking account of patient preferences after discussion with their medical provider.

People who are HRT-averse were originally not included in the decision problem, but this was due to a misunderstanding by the company at scoping that has since been rectified. Therefore, people who are HRT-averse are now included in the 'HRT-unsuitable' category. In response to clarification question A3, the company confirmed that the decision problem population is *"Menopausal people*

with moderate to severe vasomotor-predominant symptoms for whom HRT is deemed unsuitable i.e., HRT-contraindicated, HRT-caution, HRT-stoppers and HRT-averse." Clinical advice to the EAG stressed that the numbers of patients that are HRT averse may be reduced by provision of information on the full range of HRT formulations available, including the transdermal preparations.

In the NICE scope, the term 'associated with the menopause' is used to further define the population experiencing VMS, whilst in the decision problem the population is described as 'menopausal people'. In response to clarification question A6, the company confirmed that their submission covered both perimenopausal and postmenopausal people with moderate/severe VMS. The company also confirmed that both these groups are covered by the marketing authorisation. However, the company also stated that the fezolinetant RCT evidence did not cover the perimenopausal population. This may affect the applicability of the trials' results to the NHS population, and is a potential source of uncertainty.

Comparators

The NICE scope comparators include HRT where this is suitable, and no pharmacological treatment, non-hormonal pharmacological treatment, or non-pharmacological treatments where HRT is unsuitable.

The company's narrower population of people for whom HRT is unsuitable, means that omitting HRT as a comparator is appropriate. However, the company's decision problem also omits non-hormonal pharmacological treatment, or non-pharmacological treatments as possible comparators. The omission of non-pharmacological treatments appears appropriate given clinical advice to the EAG that such treatments (i.e., CBT) are rarely prescribed in the UK.

However, the EAG believes that the omission of non-hormonal pharmacological treatments is not justifiable (see section 2.2.5 for further details).

Table 4 Summary of decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	EAG comment
Population	People with moderate to severe VMS associated with the menopause.	Menopausal people with moderate to severe vasomotor-predominant symptoms for whom HRT is deemed unsuitable: HRT-contraindicated HRT-caution HRT-stoppers HRT-averse	UK clinical experts strongly indicated that HRT remains the treatment of choice in the first-line setting for menopausal people with moderate to severe VMS because the benefits of HRT extend far beyond alleviating VMS alone. Consequently, UK clinical experts expect that, in NHS primary care, fezolinetant will mainly be used in menopausal people with moderate to severe VMS for whom HRT is deemed unsuitable, due to the current unmet need for licensed, effective, and tolerable treatment options in clinical practice. As such, Astellas seeks to optimise reimbursement in this subpopulation. Menopausal people who are HRT-averse were originally excluded from the targeted population, as it is expected that HRT remains the relevant comparator for these people. However, following misunderstanding at scoping which was resolved at clarification, the decision problem population was updated to include HRT-averse.	The EAG agrees that HRT is the first-line treatment of choice and that it is appropriate for the decision problem to be narrowed to people who are HRT unsuitable. Clinical advice to the EAG indicates that the HRT contraindications and cautions provided by the company do not define in an absolute sense who should not have HRT – this would be the decision of the doctor and patient. The EAG believes that the population eligible for fezolinetant (or its comparators) may therefore be smaller than the company supposes (section 2.3).
Intervention	Fezolinetant	Fezolinetant	N/A – in line with the NICE final scope.	No comments
Comparator(s)	People for whom HRT is considered suitable: Hormonal pharmaceutical treatments (such as oestrogen and progestogen combination, or oestrogen alone) People for whom HRT is not considered suitable: No pharmacological treatment Non-hormonal pharmacological treatments, for example: Anti-depressants, such as SSRIs and SNRIs Clonidine Anti-convulsants, such as gabapentin and pregabalin	Menopausal people for whom HRT is not deemed suitable: No pharmacological treatment	Fezolinetant is a first-in-class, non-hormonal, pharmacological therapy, developed specifically to treat moderate to severe VMS. Current NICE clinical guidelines on the management of menopause (NG23) states, *"Do not routinely offer SSRI, SNRIs or clonidine as first-line treatment for VMS alone".* This statement reflects the lack of robust phase 3 trial evidence for these treatments in moderate to severe VMS, which are not licensed for VMS in the UK (except for clonidine). UK clinical experts consulted by Astellas, including those consulted as part of NICE's early engagement process through the Office for Market Access (OMA), highlighted the limited efficacy and unpleasant side effects of current non-hormonal treatments, such as SSRIs and SNRIs. These treatments are also not specifically targeted for VMS and are therefore less	The omission of HRT as a comparator is appropriate given the HRT-unsuitable population. The omission of non-pharmacological treatments also appears appropriate given that such treatments (i.e., CBT) are rarely prescribed in the UK, despite CBT being recommended in the updated NG23. However, the omission of non-hormonal pharmacological treatments is not appropriate. Antidepressants and gabapentiniods are commonly prescribed in the NHS for VMS, and will therefore be a more appropriate comparator than 'no active treatment'.

	Non-pharmacological treatments such as CBT		favourable compared with fezolinetant. Additionally, many patients who cannot take HRT, may also not tolerate the use of SSRIs or gabapentin. It is therefore expected that these non-hormonal therapies would be used in a later line compared with fezolinetant in the UK, if it is reimbursed. Additionally, UK clinical experts consulted as part of the NICE OMA meeting and the Draft Scoping Workshop indicated that psychological therapies and CBT do not represent relevant comparators in UK clinical practice, given that these treatments are typically used as add-on therapies. Given the above considerations, Astellas consider that no pharmacological treatment represents the most relevant comparator for this appraisal; a position which has been further validated, and agreed upon, with three UK clinical experts.	All the non-hormonal pharmacological treatments have been tested in RCTs which have generally demonstrated a benefit over placebo.
Outcomes	The outcome measures to be considered include: • Frequency of VMS • Severity of VMS • Sleep disturbance • Psychological symptoms (anxiety, low mood) • Adverse effects of treatment • Health-related quality of life	The outcome measures to be considered include: • Frequency of VMS • Severity of VMS • Sleep disturbance • Psychological symptoms (anxiety, low mood) • Adverse effects of treatment • Health-related quality of life	N/A – in line with the NICE final scope.	No comments
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.	The economic analysis has been conducted in line with the NICE reference case	N/A – in line with the NICE final scope.	The EAG notes that the modelled population is postmenopausal people but the population covered in the decision problem is both perimenopause and postmenopausal people. The time horizon is not sufficiently long to reflect all the long-term effects of fezolinentant. Also, the approach to estimate the utility values, with clinician adjustment of EQ-5D utility values does not follow the requrements outlined in the NICE manual. More details are presented in the Table 15.

	Costs will be considered from an NHS and Personal Social Services perspective.			
Subgroups	None.			
Special considerations including issues related to equity or equality	No special considerations raised.	The following equality issues should be considered relevant to this submission: • Differences in VMS prevalence and severity between ethnicities • Differences in VMS severity between socioeconomic levels • Location-dependent disparities in access to specialist menopause clinics	Equality considerations are further discussed in Section B.1.4 (CS)	No comment.

3 CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

3.1.1 Searches

The company searches to identify clinical evidence on fezolinetant and comparators for the treatment of VMS in postmenopausal women were detailed in Appendix D. Search strategies for the original review (covering the period 1990 – June 2021) were included, along with two further update searches (covering the period 2021 – April 2024).

The company searches in general matched the eligibility criteria for the review set out in tables 11 and 12, p38-40 of Appendix D. The database searches included search terms for the population: menopausal, perimenopausal and postmenopausal women with VMS. The population terms were combined with terms for the intervention fezolinetant or terms for non-hormonal therapies. The original search strategies (covering the period 1990 – June 2021) included an additional set of search terms for hormonal therapies.

The EAG found some weaknesses and errors with the searches which may have caused relevant studies to be missed. These are detailed in Table 5.

Table 5 EAG appraisal of evidence identification

Topic	EAG response	EAG Notes
Is the report of the search clear and comprehensive?	PARTLY	Extra searches for escitalopram are referred to in Section D.1.1.2 SLRs 1 and 2 however the search strategies were missing. These were provided in the company response to clarifications.
Were appropriate sources searched?	PARTLY	The International Health Technology Assessment database (INAHTA) was not searched in the original or update searches. Epistemonikos or KSR Evidence (databases containing non-Cochrane systematic reviews) were not searched in either the original or update searches. The Cochrane Database of Systematic Reviews was searched.
Was the timespan of the searches appropriate?	YES	Database searches covered the period 1990 to April 2024. Manual conference proceedings searches covered the period 2019 to April 2024. Clinical Trial registry searches covered the period inception to 15th April 2024.
Were appropriate parts of the PICOS included in the search strategies?	YES	[P] VSM in menopause AND ([I] fezolinetant OR [C] comparators) AND [S] RCTs n.b. Comparators in the original search were hormonal treatments or non-hormonal treatments. Comparators in the update searches were non-hormonal treatments only. The update searches also included a section to identify adverse effects: [I] fezolinetant AND [O] adverse effects.
Were appropriate search terms used?	PARTLY	**Missing search terms:** Trial registry and conference abstract searches – postmenopause, post-menopause, perimenopause and peri-menopause were missing. Search terms and subject headings for HRT included in original database search strategies were fairly limited. Examples of missing terms include: HRT, hormone replacement therapy, oestradiol, progesterone.

		Brand names for fezolinetant (veoza and veozah) missing from all search strategies. A limited range of search terms for adverse effects were used in the update searches. **Search term errors:** Update searches CENTRAL – MeSH headings incorrect at line 1 [mh "climacterium"] and at line 4 [mh "hot fl*sh"]. Both terms retrieved 0 hits when tested by the EAG. Update searches CENTRAL – incorrect truncation symbol used at line 31 (behavio?l NEXT therapy or behavio?l NEXT treatment). This part of the search line 31 retrieved 0 hits when tested by the EAG. Original searches - truncation symbol * is missing for climacter at line 5
Were any search restrictions applied appropriate?	NO	**Original searches (1990-2021):** Language bias possible as retrieval was restricted to English language only. Case reports and letters were excluded from the search results. Adverse effects and safety information can be reported in these publication types. As several databases were searched all at once the restrictions to remove various publication types and animal only studies may not have worked as expected in all databases and may have caused relevant studies to be missed. An inappropriate restriction to RCT studies was applied in the original searches. This should not have been applied to the results from CDSR, CENTRAL, DARE, HTA database, NHS EED. These databases are all pre-filtered to include particular study designs therefore the restriction to RCTs may have caused studies to be missed from these resources. n.b. The restrictions applied to the update searches were generally appropriate.
Were any search filters used validated and referenced?	PARTLY	No search filters were referenced in the submission. The company clarified in their response to EAG questions that RCT search filters by SIGN were used in the update searches. The SIGN RCT filters have not been externally validated. They also clarified that a validated search filter for adverse effects was not used in the update searches. A validated adverse effects search filter for use in MEDLINE and Embase has been available since 2022 and would have provided a more comprehensive method of searching for adverse effects data on fezolinetant.[39]

EAG response = YES/NO/PARTLY/UNCLEAR/NOT APPLICABLE

3.1.2 Inclusion criteria

An initial SLR included both hormonal and non-hormonal pharmaceutical therapies. Two subsequent versions (update 1 and 2) focussed mainly on non-hormonal therapies for moderate/severe VMS. As the two updated SLRs are more in line with the decision problem, these have been selected as relevant to this report and are henceforth referred to as the company SLR. Table 6 summarises the inclusion criteria for the company SLR.

Table 6. Eligibility criteria for SLR Update 1 (2023) and SLR Update 2 (2024) [Table 12 in CS Appendix]

Domain	Inclusion Criteria	Exclusion Criteria
Patient population	Post-menopausal individuals with moderate-to-severe VMS associated with menopause, defined as a minimum average of 7 to 8 moderate to severe VMS per day, or 50 to 60 per week	Any other conditions, including patients with breast cancer or a history of breast cancer
Interventions	• Fezolinetant • SSRIs: – Paroxetine (Paxil) – Citalopram (Celexa) – Fluoxetine (Prozac) – Sertraline – Escitalopram • SNRIs: – Venlafaxine (Effexor XR) – Desvenlafaxine ER (Pristiq) • Clonidine (Catapres) • Gabapentin (Neurontin) • Pregabalin • Oxybutynin (Oxytrol) • CBT	Studies not investigating a relevant intervention
Comparators	• Any relevant intervention, as listed above • Standard of care • Placebo	-
Outcomes	• Changes in the frequency of VMS • Changes in the severity of VMS • Number of responders • Psychological symptoms (anxiety, low mood) • Adverse effects of treatment	No relevant outcomes
Study design	• RCTs (phase III or IV) • Conference abstracts published in or since June 2021	Any other study designs
	Relevant SLRs and (N)MAs were identified at the title/abstract review stage and hand-searched for relevant studies. They were excluded at the full text review stage unless they presented primary research.	
Other considerations	• Abstract or full-text in the English language • Any location • Human subjects • Published in 2021 or later[a] (unless reporting on CBT, sertraline, pregabalin or escitalopram)	• Publications without an abstract or full-text in the English language • Not in human subjects • Published before 2021 (unless reporting on CBT, sertraline, pregabalin or escitalopram)

Abbreviations: CBT, cognitive behavioural therapy; (N)MA, (network) meta-analysis; RCT, randomised clinical trial; SLR, systematic literature review; SNRIs, serotonin-norepinephrine reuptake inhibitors; SSRIs, selective serotonin reuptake inhibitors; VMS, vasomotor symptoms. **Footnotes:** [a]Searches were conducted without an amended date limit for the second SLR update (2024) and the results were de-duplicated against the EndNote library from the first SLR update (2023) to ensure that no articles were missed between the two SLR updates.

Although closer to the decision problem than the original SLR, the company SLR and decision problem have several differences.

Population

There are three differences in terms of population.

1. The decision problem is restricted to people who are either HRT stoppers or HRT-unsuitable, but no restriction on HRT use is applied in the SLR. The effects of including people eligible for HRT are uncertain; in response to clarifications, the company stated that HRT-suitability is not an effect modifier. There does not appear to be any evidence suggesting that HRT-suitable and HRT-unsuitable people would respond differently to fezolinetant, nor does there appear to be a plausible mechanism to explain any effect modification.

2. The decision problem specifies moderate/severe VMS, but this is not precisely defined. In contrast, the SLR defines moderate/severe severity '*as a minimum average of 7 to 8 moderate to severe VMS per day, or 50 to 60 per week*'.

3. The decision problem refers to 'post-menopausal people', which was clarified by the company to mean both perimenopausal and postmenopausal people. The SLR restricts the population to 'post-menopausal individuals'.

Whilst no restriction by HRT-use tends to widen the SLR scope relative to the decision problem, the latter two criteria are expected to narrow it. Any overall narrowing to the company SLR scope may mean that some trials relevant to the decision problem may have been missed.

Intervention/comparison

Both decision problem and company SLR agreed that the intervention was fezolinetant, but disagreed on the comparators. Whilst the decision problem restricted comparators to 'no pharmacological treatment', the company SLR included a range of pharmacological non-hormonal treatments (anti-depressants, clonidine, gabapentinoids and oxybutynin), and CBT. As the company has decided to include only 'no pharmacological treatment' as a comparator in the decision problem, the only parts of the company SLR with relevance to the decision problem are the fezolinetant versus placebo trials.

As stated previously, the EAG believe that the decision problem should be expanded to include pharmacological non-hormonal treatments. Should the company decide to amend their decision problem, the enhanced scope of the company SLR can provide information for indirect comparisons between fezolinetant and relevant pharmacological non-hormonal treatments.

Outcomes

Outcomes also differ between decision problem and company SLR. Both cover frequency and severity of VMS, psychological symptoms and adverse events. However, the SLR does not contain

the important decision problem outcomes of sleep disturbance or health-related quality of life (HRQoL).

3.1.3 Study selection and data extraction

On the basis of the title and abstract, two independent reviewers used the eligibility criteria (see section 3.1.2) to select full-text copies of publications for further inspection. Two independent reviewers then decided on inclusion or exclusion of the full articles based on the same eligibility criteria. In cases where the full article gave insufficient information to achieve all inclusion criteria, the article was excluded. A third reviewer was used to achieve consensus in case of disagreements in both the first and second sifts. Data were extracted from the included studies by one reviewer onto a pre-specified extraction sheet, and data extraction accuracy was validated by another reviewer.

The EAG has no concerns about the methodology used for selection and extraction of data.

3.1.4 Quality assessment

For each study, quality was assessed using the University of York Centre for Reviews and Dissemination (CRD) quality assessment tool for RCTs. The risk of bias adjudications for each included study are provided in Table 25 of the CS appendix (part D) but no details were provided to justify the judgements.

3.1.5 Evidence synthesis

The overall company SLR included 35 publications, covering 14 studies. Details of these studies are provided in Table 19, CS Appendix D. In brief, these studies had similar populations in terms of disease severity, with most recruiting people with at least seven moderate to severe hot flushes per day, but no information is provided on other characteristics. Placebo comparisons were made with desvenlafaxine (5 studies), gabapentin (1 study), gabapentin extended release (3 studies), paroxetine (2 studies), clonidine (1 study) and fezolinetant (2 studies). It should be noted that the company SLR did not include the DAYLIGHT trial as no full published paper was available at the time of the SLR being conducted.

In the original CS, the company did not provide any outcome data from the 14 studies, and synthesis of evidence was not carried out due to the company's decision to exclude all comparators except 'no pharmacological treatment' from the decision problem. As explained in section 2.2.5, the EAG disagrees with this decision.

During clarification, the EAG requested that the company provide objective evidence of the efficacy and safety of the non-hormonal pharmacological comparators compared to fezolinetant. In response, the company provided an exploratory Bayesian NMA.[30] This is critiqued in Section 3.4

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

The CS stated that there were 19 completed studies (11 phase 1 studies, four phase 2 studies and four phase 3 studies) within the fezolinetant clinical development programme; however, direct references were not provided for all these trials. Given that the EAG had identified several phase 3 fezolinetant trials that were not referenced in the CS, and that the company restricted their SLR to include only phase 3 and 4 trials, the EAG requested at clarification a full list of all the fezolinetant trials under the clinical development programme to date.

The company provided a full list of the completed phase 2 and 3 trials, noting that there have now been five and six completed phase 2 and phase 3 fezolinetant trials, respectively. Besides the DAYLIGHT and SKYLIGHT 1, 2 & 4 trials presented in the CS, Table 5 of the clarification response reported a further seven fezolinetant trials. The EAG notes that six of these trials are outside of the scope of the appraisal, based on the dose of fezolinetant being investigated: 30mg for MOONLIGHT 1 and MOONLIGHT 3; 15mg or 30mg for STARLIGHT; 15mg, 30mg, 60mg and 90 mg for Fraser et al 2020;[40] 60mg or 180 mg for Fraser et al 2021;[41] and 90mg for Depypere et al.[42] The seventh trial, study ESN364-UF-02, had an ineligible population, being of women with uterine fibroids.

The CS included four eligible RCTs of fezolinetant 45mg: DAYLIGHT was placebo-controlled to 24-weeks; SKYLIGHT 1 and SKYLIGHT 2 were concurrent and identically designed trials which were placebo-controlled to 12-weeks; and SKYLIGHT 4 was a safety-only trial which was placebo-controlled to 52-weeks.

3.2.1 Critical appraisal of the fezolinetant RCTs

3.2.1.1 Risk of bias

The company judged the SKYLIGHT 1&2 and the DAYLIGHT trials to be at low risk of bias (CS, Table 16). The EAG identified two issues with missing data, described below, which it believes place the results of all three trials at high risk of bias (due to missing data) with respect to continuous outcomes (this includes the VMS frequency and VMS severity outcomes).

Issue 1: Use of different estimand approaches

'Estimands' provide a structured description of the treatment effects a trial intends to quantify. Differences in estimand approaches can affect how pragmatic or explanatory a trial is. An important difference in approaches relates to how intercurrent events are handled, i.e. events which affect a participant's assigned treatment, such as discontinuing trial treatment, or taking a trial-prohibited treatment. 'Hypothetical policy estimand' approaches consider how outcome data might have been if

intercurrent events had not occurred. 'Treatment policy estimand' approaches use outcome data regardless of any intercurrent events.

The SKYLIGHT 1 and 2 trials both used a hypothetical policy estimand approach, in which patients who discontinued treatment were not followed-up any further. The primary approach used in the DAYLIGHT trial was the treatment policy estimand, in which patients who discontinued treatment were continued to be followed-up, with their post-discontinuation data used in the analyses. The lack of follow up for all patients who discontinued treatment in the two SKYLIGHT trials is concerning: it unnecessarily increases the amount of missing outcome data and also ensures that no data can be recorded on subsequent therapies in patients who discontinue their trial treatment.

Issue 2: Approaches to handling missing data

For continuous outcomes, mixed models repeated measures analyses (MMRM) were used, which assumed that any missing outcome data were 'missing at random' (MAR). When MAR assumptions are used in a MMRM model, missing data are considered similar to the mean for their treatment group, with the analysis incorporating all available on-treatment data to calculate mean treatment effect estimates. The company's primary analyses assumed that data were missing at random, although the CS stated that there was no explicit imputation of missing data (for the primary analysis).

These methods can bias results if MAR is not a reasonable assumption. The EAG does not consider that a MAR approach is reasonable in any of the trials. This is because of concerns about the company's reporting clarity on reasons for treatment discontinuation e.g. not explicitly reporting how many patients discontinued due to lack of efficacy. Also, the numerator data from Figure 15 of the CS suggest that patients with missing data up to week 12 (of the placebo-controlled period) did not return to participate in the extended period. Although it is plausible that a small number of patients may have data which are missing at random (e.g. the participant emigrated) and so do not return to the study, missing data due to lack or loss of efficacy, or due to adverse events appear more likely scenarios for most participants and these reasons are not compatible with a MAR assumption.

At the final trial endpoints, data could be missing due to treatment discontinuation or due to other reasons. Data on discontinuations and all missing data (at the primary endpoints) are reported in Table 7.

Table 7 Proportions of patients who discontinued trial treatment or who had missing data at primary endpoints in the SKYLIGHT 1 & 2 and DAYLIGHT trials

Trial	Proportion of patients who discontinued trial intervention		Proportion of patients with missing data at primary endpoints	
	Fez 45 mg	Placebo	Fez 45 mg	Placebo
DAYLIGHT	14%, 5% due to TEAE	23%, 6% due to TEAE	22%	27%

SKYLIGHT 1	8%, 3% due to TEAE	13%, 5% due to TEAE	16%	21%
SKYLIGHT 2	7%, 1% due to TEAE	10%, 1% due to TEAE	13%	16%

Sources for missing data calculations: Table 10 Daylight CSR, Table 11 Skylight 1 CSR, Table 11 Skylight 2 CSR, Co-primary endpoints for SKYLIGHT 1 & 2 were frequency and severity of moderate to severe VMS at week 12. Primary endpoint for DAYLIGHT was frequency of moderate to severe VMS at week 24. TEAE Treatment-emergent adverse event

Impact on results of the different estimand and missing data approaches

Results for VMS outcomes using different MAR and estimand approaches across trials are presented in Table 8. In addition to the MAR analyses, the company performed an analysis which used a different assumption for *some* of the missing data, referring to it as a "discontinuation-reason based multiple imputation sensitivity analyses"; this assumed that treatment benefits are diminished after treatment discontinuation. However, this "jump to reference" algorithm (where placebo is the reference group) only related to participants who discontinued due to a treatment-emergent adverse event (TEAE); all other participants who discontinued were still assumed to have data missing at random.

Table 7 shows that across all three trials the proportion of participants who discontinued due to an TEAE was small, so this sensitivity analysis does not greatly affect the results. In both DAYLIGHT (Table 13, CS Appendices) and (pooled) SKYLIGHT 1 & 2 (Table 15, CS Appendices) the most frequent reason for discontinuation was the vague "*Withdrawal by patient*". The EAG does not consider this as a suitable reason for a MAR assumption, especially without exploring alternatives in sensitivity analyses. The EAG notes that discontinuation rates due to lack or loss of efficacy were not specifically reported in the clinical study reports (CSRs); it is unclear from the company's CSRs whether such patients are covered by the "*withdrawal by patient*" discontinuation reason, or whether no patients discontinued due to lack/loss of efficacy (the latter seems unlikely).

Table 8 Trial results for VMS outcomes using different estimand approaches and missing data assumptions for discontinuations

Trial, estimand and missing data assumptions	**Difference in least squares means: fezolinetant 45mg versus placebo**	
	Frequency of moderate to severe VMS (SE)	**Severity of VMS (SE)**
DAYLIGHT, Week 24		
Treatment policy, MAR	-1.93 (0.36)	-0.39 (0.09)
Treatment policy, Discontinuation-reason based MI & MAR	-1.88 (0.35)	Not analysed
Hypothetical policy, MAR	-2.14 (0.36)	-0.43 (0.09)
SKYLIGHT 1, Week 12		
Hypothetical policy, MAR	-2.55 (0.43)	-0.20 (0.08)

Hypothetical policy, Discontinuation-reason based MI & MAR	-2.49 (0.44)	-0.20 (0.08)
SKYLIGHT 2, Week 12		
Hypothetical policy, MAR	-2.53 (0.55)	-0.29 (0.08)
Hypothetical policy, Discontinuation-reason based MI & MAR	-2.48 (0.55)	-0.28 (0.08)

MAR missing at random, MI Multiple imputation

The EAG considers use of the MAR assumption for missing data at the final trial endpoints (regardless of whether or not this was associated with discontinuation) to not be a reasonable assumption. The EAG notes that the number of patients with available data for the analyses does not increase after week 12 (Figure 15, CS) i.e. there is little evidence to suggest that a significant proportion of patients with missing data at 12 weeks went on to continue with the study, so a MAR assumption is not reasonable.

For the DAYLIGHT and SKYLIGHT trials, use of the MAR assumption using a MMRM will tend to maintain the treatment effect (the difference between fezolinetant and placebo) in the significant number of patients with missing trial data. The EAG considers it much more likely that the treatment effect will attenuate, based on the more plausible assumption that only a small number of patients are likely to have data which are genuinely MAR.

The EAG therefore considers that all the company's analyses of continuous outcomes are at high risk of bias due to missing outcome data, with the bias favouring fezolinetant.

However, the company did use a conservative missing data assumption method (non-responder imputation) for the binary responder VMS outcomes in both SKYLIGHT trials and DAYLIGHT. These are reported in Section 3.2.2.2.

3.2.1.2 Applicability of the trial populations to NHS practice

The EAG noted several aspects of the company's trial populations which differ to the population anticipated to use fezolinetant on the NHS. Firstly, the company's RCTs enrolled only postmenopausal women (typically around 5-7 years postmenopausal), so clinical data for the perimenopausal population is not available. At clarification (question A6) the company stated that the licensed indication for fezolinetant covers both perimenopausal and postmenopausal people with moderate to severe VMS, adding that although the EMA's Committee for Medicinal Products for Human Use initially considered whether the indication should be limited to postmenopausal women only, it was persuaded by the company's "substantial evidence from the published literature demonstrates that the physiological mechanism underlying VMS remains consistent across the menopause transition, from perimenopause to postmenopause, and is not exclusive to postmenopausal

women". The EAG's clinical adviser thought that this assumption was reasonable, noting that HRT is only licensed in postmenopausal women, but is frequently prescribed for perimenopausal women because oestrogen deficiency underlies symptoms in both perimenopause and menopause.

Similarly, by enrolling selected cohorts who had a minimum average of seven moderate to severe VMS events per day, or 50 to 60 moderate to severe events per week (during the 10 days prior to randomisation), the company's trials have focussed only on a subgroup of the patients who will be eligible to receive fezolinetant in NHS practice. This means that there is some uncertainty about how well fezolinetant works in patients who have fewer, but severe, VMS events, or many more but less severe VMS events. The trial evidence suggests a significantly smaller effect in patients with lower VMS frequencies at baseline (see Section 3.2.2.2).

The second population issue related to the EMA noting that in SKYLIGHT 1 and SKYLIGHT 2 "a remarkable number of participants with a hysterectomy (i.e. 32.2%, 32.0%) and with ovariectomy (i.e. 21.5%, 21.8%) were enrolled". The EAG's clinical adviser thought that these are higher proportions than would be seen in clinical practice and that these procedures would cause early menopause, which is associated with more severe VMS; it is therefore plausible that these subgroups might respond differently to treatment. At clarification the EAG requested detailed subgroup results from the company. These were presented for the two SKYLIGHT trials and indicated that a history of hysterectomy or ovariectomy did not significantly modify treatment response in terms of frequency of moderate to severe VMS.

Finally, at clarification (question A17) the EAG asked the company to document how many patients were excluded for medical or blood pressure reasons, or for using prohibited interventions; further details on these reasons are reported in Table 8 of the CS. A slightly modified version of the company's tabulated response (numbers randomised have been added) is presented in Table 9; the company noted that some patients with these criteria were randomised in error. The EAG nevertheless notes that a substantial number of patients were excluded – especially from the two SKYLIGHT trials – who will be eligible to receive fezolinetant on the NHS; it remains unclear whether these excluded subgroups will benefit from fezolinetant to the same extent as was seen in the RCTs.

Table 9: Number of patients excluded prior to randomisation for select reasons in DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2

Reason for exclusion	Number of patients excluded from trial		
	DAYLIGHT	SKYLIGHT 1	SKYLIGHT 2
Number randomised (all arms)	453	527	500
Medical condition or chronic disease	[redacted]	[redacted]	[redacted]
Blood pressure measurements	[redacted]	[redacted]	[redacted]
Prohibited treatment for VMS	[redacted]	[redacted]	[redacted]

3.2.2 Clinical efficacy results of the fezolinetant RCTs

3.2.2.1 Baseline characteristics

Baseline characteristics of the trial cohorts were reported in Tables 10 (DAYLIGHT) and 11 (SKYLIGHT 1 & 2) of the CS. The mean age in all three trials was between 54 and 55 years. In response to the EAG's first clarification question, the company stated that HRT unsuitability was not considered to be a treatment effect modifier, but race or ethnicity was identified as a potential effect modifier. These characteristics were well-balanced across the three fezolinetant trials. The only imbalance noted by the EAG was a shorter mean time since onset of amenorrhoea in the DAYLIGHT placebo group (59 months) compared to the fezolinetant group (72 months). The EAG's clinical adviser thought this difference noteworthy, as it could have affected the results; VMS usually improves over time so it is possible that participants in the fezolinetant group were more likely to improve since they were further from menopause.

3.2.2.2 VMS outcomes

VMS frequency

Results for the continuous VMS outcomes across trials, using different analysis methods, have been presented in Table 8 and noted as being at high risk of bias; it is also not easy to interpret the clinical significance of those results.

Table 10 presents results for the company's responder analyses of change from baseline in frequency of moderate to severe VMS, using cut-offs of ⩾ 50%, ⩾ 75% and 100% reductions. On p149 of the CS the company stated that UK clinical expert opinion indicated that a 75% reduction from baseline in moderate to severe VMS frequency is a clinically meaningful improvement. The results for the responder analyses are judged to be at low risk of bias since non-responder imputation were used for missing data in all three trials. They indicate a consistent benefit at week 12 of fezolinetant 45mg when compared with placebo with around [redacted] more fezolinetant patients achieving a 75% reduction from baseline in frequency of moderate to severe VMS, and around [redacted] more fezolinetant patients achieving a 100% reduction.

Table 10 Responder Analyses of Change from Baseline in Frequency of Moderate to Severe VMS

Trial	Responders with ≥ 50% Reduction		Responders with ≥ 75% Reduction		Responders with 100% Reduction	
	Fez 45mg	Placebo	Fez 45mg	Placebo	Fez 45mg	Placebo
DAYLIGHT Week 24	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OR (95% CI)	[redacted]		[redacted]		[redacted]	

DAYLIGHT Week 12	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OR (95% CI)	[redacted]		[redacted]		[redacted]	
SKYLIGHT 1 Week 12	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OR (95% CI)	3.16 (2.04 to 4.94)		NR		3.26 (1.33 to 9.19)	
SKYLIGHT 2 Week 12	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OR (95% CI)	[redacted]		[redacted]		[redacted]	

Sources: DAYLIGHT CSR Table 26; SKYLIGHT 1 CSR Table 18; SKYLIGHT 2 CSR Table 18 and Table 9.3.3.6.1. Odds ratios > 1 favour fezolinetant. NR Not reported, OR Odds ratio. * from Morga 2023 NMA supplementary file 1.

As part of a supplemental prespecified analysis for the pooled SKYLIGHT trials dataset, the company also estimated the clinically meaningful within-subject change threshold in VMS frequency. A threshold of 6.2 was found to characterise the level of reduction from baseline to week 12 in the frequency of moderate to severe VMS that constituted meaningful improvements from the participants' perspective, based on the "moderately better" anchor level. In the pooled placebo group 118/293 (40%) achieved a response whereas in the fezolinetant 45mg group 209/303 (69.0%) achieved a response. However, based on these denominators, this analysis did not use non-responder imputation for participants with missing data, so the EAG considers the difference of around [redacted] between groups using the 75% threshold (Table 10) to be the most valid relative estimate of differences in clinically meaningful improvement.

The EAG notes important differences between the efficacy outcomes and datasets of the DAYLIGHT and SKYLIGHT 1 and 2 trials and the efficacy data used in the company's economic model. The four VMS frequency health states in the model do not relate to VMS frequency improvements/cut-offs used in the trial (unlike the VMS responder outcomes, which were not used in the model). The clinical validity of the data used for the health state transitions is therefore questionable. Also, the efficacy data used in the modelling were observed data which did not take account of any biases arising from missing trial data. Data from the different trials were also pooled naively, rather than using meta-analysis.

Subgroup analyses

Figure 20 of the CS reported pooled SKYLIGHT 1&2 subgroup analyses of difference in mean change in frequency of moderate to severe VMS from baseline to Week 12. These suggested an important difference between subgroups based on baseline VMS frequency (split by ≥ median vs < median, the median being around 9.7 across the SKYLIGHT 1&2 placebo and fezolinetant 45mg groups). The analysis suggested a significantly smaller effect in patients with < median VMS

frequencies at baseline (mean difference -1.36) vs ⩾ median (mean difference -3.66); the EAG requested the statistical significance (p-values) for all the tests for interaction, with the VMS frequency analysis reported as being [redacted] (company's response to clarification question A20).

It is concerning that despite this result no subgroup analysis of baseline VMS frequency was pre-specified or reported for DAYLIGHT. The SKYLIGHT 1&2 analysis nevertheless adds weight to the EAG's concerns about the populations recruited to the DAYLIGHT and SKYLIGHT trials. Given that trial efficacy estimates for improvements in VMS frequency appear to be driven by the subgroup of patients with around 10 or more VMS events per day, and that trial patients must have had a minimum average of 7 to 8 moderate to severe VMS event per day to be recruited to the trials, it is highly plausible that fezolinetant will not be as effective in the NHS population, where there is no requirement for a minimum number of VMS events per day. Given the VMS severity results outlined below this will be of particular concern for patients who have few, but severe, VMS events.

VMS severity

As noted in Section 2.2.4.1, the EAG has concerns regarding the validity of the outcome measure used to assess VMS severity in the three fezolinetant trials. Although statistically significant differences in VMS severity were seen between fezolinetant and placebo in all three trials (CS Tables 18 & 27), the clinical significance of these results is unknown. The EAG asked the company to report clinically important thresholds for all outcomes used in the DAYLIGHT and SKYLIGHT trials (clarification question A13) but no data were reported for VMS severity in the company's response.

3.2.2.3 Sleep disturbance

WPAI-VMS

The Work Productivity and Activity Impairment – VMS questionnaire (WPAI-VMS) evaluates VMS-related work productivity and activity in the previous seven days. It is a six-item patient-reported outcome measure (PROM) comprising the four domains of absenteeism, presenteeism, overall work productivity loss and activity impairment (CS, p53). Higher scores denote greater impairment.

In DAYLIGHT, the fezolinetant group demonstrated significantly greater improvements in WPAI-VMS (overall work productivity loss) than the placebo group at 24 weeks [least squares (LS) difference in means: -6.77 (-13.09, -0.45)] (Table 23, CS). Similar effects were seen for the activity impairment domain [LS difference in means: -8.38(-13.22, -3.55)] (Table 24, CS). Levels of absenteeism and presenteeism appeared to be similar between groups (Figure 10, CS). Similarly, in SKYLIGHT 1 and 2, fezolinetant showed significantly better improvements than placebo for WPAI-VMS overall productivity loss and activity impairment (Tables 35 and 36, CS), but levels of absenteeism and presenteeism again appeared to be similar between groups (Figure 18, CS). The company did not provide known minimum clinically important differences (MCIDs) for this outcome,

and no MCIDs relevant to a VMS population were found by the EAG through limited searching. Thus, it is unclear if the statistically significant effects observed were also clinically important.

3.2.2.4 Psychological symptoms

PHQ-4 total scores

The Patient Health Questionnaire for Anxiety and Depression (PHQ-4) questionnaire measures anxiety and depression during the previous two weeks, with higher scores denoting a worse outcome (CS, p53).

In DAYLIGHT there were no significant differences between groups [LS difference in means: -0.44 (-0.97, 0.08) (Table 25, CS).

3.2.2.5 Health-related quality of life

MENQOL

The menopause-specific quality of life questionnaire (MENQOL) is a PROM that assesses the impact on quality of life of four domains (vasomotor, psychosocial, physical and sexual) of symptoms associated with menopause over the previous week. Higher scores represent a worse impairment (CS, p 52).

Total score

In the DAYLIGHT trial there was a greater improvement in MENQOL total score at week 24 for the fezolinetant 45mg group compared to the placebo group [LS mean difference -0.44 (95% CI: -0.695, -00.185)] (DAYLIGHT CSR, Table 27). There were similarly greater improvements for fezolinetant at 12 weeks in SKYLIGHT1 [-0.49(-0.745, -0.235)][32] and in SKYLIGHT 2 [-0.47(-0.76, -0.19)] (SKYLIGHT 2 CSR, Table 9.3.4.9.1).

Vasomotor domain

In the DAYLIGHT trial there was a greater improvement in MENQOL vasomotor domain score at week 24 for the fezolinetant 45mg group compared to the placebo group [LS mean difference -0.96 (95% CI: -1.39, -0.54)] (Table 21, CS). In the SKYLIGHT 1 and 2 trials there were also greater improvements for fezolinetant at 12 weeks [SKYLIGHT1: -0.99 (-1.42, -0.57), SKYLIGHT 2: -0.76 (-1.20, -0.32)] (Table 33, CS).

Although the differences between groups were statistically significant in both the total score and the VMS domain, it is important to assess the clinical importance of these differences in improvement. As part of the clarification process, the company was asked to provide information on the MCIDs for all outcomes (question A13), including the MENQOL. The company suggested an MCID value of 2.0 for the vasomotor domain, based on a within-group anchor-based study of data from SKYLIGHT 1 and

2.[43] The differences between groups in either the total score and VMS domain did not reach this (within-group) threshold, so clinical significance cannot be assumed..

EQ-5D-5L VAS

The EQ-5D-5L visual analogue scale (VAS) provides a generic measure of quality of life on a scale of 0 to 100, with 100 indicating the best imaginable health and 0 representing the worst imaginable health.

In the DAYLIGHT trial, there were no significant differences between groups in improvements of EQ-5D-5L VAS score over 24 weeks [difference in LS means +1.6(-1.5, 4.6)] (Table 22, CS). In SKYLIGHT 1 and SKYLIGHT 2, there were also no significant differences in improvements between groups at 12 weeks (Table 34, CS).

It should be noted that there was a statistically significant between-group difference in the MENQOL vasomotor domain improvements but not in the EQ-5D VAS improvements. There was also a larger between-group difference in the MENQOL vasomotor domain improvements than the MENQOL total score improvements. One explanation for these observations relates to the fact that VMS make up only a proportion of symptoms associated with menopause, and so even if VMS are alleviated, the remaining symptoms associated with menopause may still have a substantial impact on quality of life. Therefore, an improvement in VMS may have a measurable impact on specific VMS scales, but not greatly affect a more general quality of life scale like the EQ-5D VAS. If we assume that a more general measure of quality of life is the most valid indicator of an individual's overall well-being, then this suggests that treating VMS alone may not have the level of impact on an individual suggested by the MENQOL vasomotor domain results.

3.2.3 Adverse event and safety data

In the CS (p32), the company asserts that fezolinetant has a similar TEAE rate to placebo. However, the EAG identified independent findings that question this.

Potential increase in risk of neoplasms

The EMA reported a higher number of neoplasms (benign, malignant and unspecified [including cysts and polyps]) in participants who took fezolinetant 45mg (7 participants) compared with placebo (2 participants) in the 52-week safety trial (SKYLIGHT 4).[44] This difference was deemed by the EMA to be likely due to chance. Although the CS did not raise this as a possible issue, a summary of the neoplasm data reported across tables in the CS is presented in Table 11. This is notably different to the EMA data.

Table 11: Endometrial hyperplasia, cancer or disordered proliferative endometrium events reported in the CS for DAYLIGHT and SKYLIGHT 1, 2 and 4

Trial	Placebo	Fezolinetant 45mg
DAYLIGHT	2/226	1/226
SKYLIGHT 1	0/175	0/173
SKYLIGHT 2	0/167	0/167
SKYLIGHT 4	0/186	1/203 (simple hyperplasia without atypia)

Source: Table 43 of the CS appendices, and Tables 44 and 53 of the CS.AEs were coded using MedDRA.

A letter published in The Lancet[45] in late 2023 meta-analysed detailed neoplasm data from SKYLIGHT 1, 2 and 4, from www.clinicaltrials.gov records, which showed a statistically significant increased risk of neoplasms with fezolinetant 45mg when compared with placebo. The Peto odds ratio (OR) for fezolinetant 45mg versus placebo was 4.55 (95% CI: 1.70, 12.15) for the analysis of randomised data only (without extension data). A major limitation of this meta-analysis was the pooling of studies with different follow up periods (the double blind period of the SKYLIGHT 4 trial was 52 weeks, but only 12 weeks for the SKYLIGHT 1 and 2 trials). Therefore, the EAG carried out a Bayesian pairwise meta-analysis on the Douxfils data[45] using a complementary log-log link[46] which estimates hazard ratios (HR), to allow for the different follow up periods (Appendix 1). The fixed effect model yielded an HR of 8.22 (95% credible interval [CrI]: 2.18, 61.36) for fezolinetant versus placebo. Because of the low number of studies and the low event rate, the uncertainty of this estimate is high. However, the probability that the HR is greater than 1 (i.e. that there is an increased risk of neoplasms) is 99.95%, which supports the conclusions of Douxfils et al.[45]. The fixed effect model fitted the data well, with a median total residual deviance less than the number of unconstrained datapoints (see Appendix Appendix 1 for details).

Douxfils et al.[45] noted that of the 25 fezolinetant events in total, across 23 patients, only three were benign, and concluded that "*The higher incidence of neoplasms with fezolinetant requires an in-depth evaluation of all cases observed to further assess associated baseline risk factors and propose risk minimisation strategies*.". Table 12 summarises the neoplasm event data from this letter relating to the randomised phases of the SKYLIGHT 1, 2 and 4 trials (i.e. without inclusion of data from any extension periods); the EAG notes this includes two extra fezolinetant 45mg patients with events in SKYLIGHT 4, compared to the EMA data.

Table 12 Number of patients with a neoplasm event reported in the randomised phases of the SKYLIGHT 1, 2 and 4 trials as extracted from the www.clinicaltrials.gov website[45]

Trial	Placebo	Fezolinetant 45mg
SKYLIGHT 1 (12 weeks)	0/175	3/173
SKYLIGHT 2 (12 weeks)	0/167	2/167
SKYLIGHT 4 (52 weeks)	2/610	9/609

The company were asked to comment on the findings of the Douxfils 2023 letter[45] and to provide any additional evidence (clarification question A14). The company reiterated its stance that the evidence base did not present a signal for an increased risk of neoplasms. In their clarification response, the company's arguments against such a signal were:

1. The US Food and Drug Administration (FDA) has concluded that 50% of malignancy events for fezolinetant patients were likely due to pre-existing malignancy that was undiagnosed at baseline, which would make the rate of events attributable to fezolinetant within the expected background rate.

In the 120 day safety update of clinical safety (p196 of the FDA report.[47]) the reviewer does indeed comment that "*six of the participants (8/16 malignancy TEAEs in treatment arms) were likely to have had preexisting malignancy that were undiagnosed at the time of enrolment. This brings the incidence rate of malignancy TEAEs within the normal background rate of cancer for the age group of this population*". This is based upon the reviewer's rationale that cancers were "*sporadically distributed without patterns related to duration of drug exposure, body system and/or cancer type. Most cancers were singular in nature and occurring in disparate systems and varied cancer types except for colon cancer (2 cases), squamous cell skin cancer (3 cases) and endometrial cancer (3 cases)*". However, the rationale is somewhat speculative, and does not properly consider the dose response effect occurring across the 30mg and 45mg fezolinetant groups, nor the low probability of such a differential effect across the 30mg, 45mg and placebo groups occurring by chance. Importantly, the above statement was made in the absence of input from the FDA Division of Oncology on the studies included in the CS (CL 0301, CL 0302 and CL0304, p155 of the FDA report[47]), which was described as '*pending*' (p196 of the FDA report.[47]). The delayed report from the Division of Oncology was appended to the FDA report, and the division's statements implied that they considered it difficult to make a decision about a neoplasm risk signal in the CS studies based on uncertainties in the data.

2. The Peto method used by Douxfils 2023 letter[45] is not appropriate for rare events.

The EAG agrees that the Peto OR is not without its drawbacks, but also notes that other methods of pooling carried out by Douxfils were reported to lead to a similar result.[48] Moreover, the

complementary log-log model used by the EAG (which allows for differing follow up times) used a standard HR, which demonstrated a plausible additional effect of fezolinetant.

3. The NK3R antagonism mechanism for neoplasm development is not supported by the literature.

The EAG notes that even if a plausible mechanism is currently unavailable this does not constitute evidence for no effect.

4. Out of the DAYLIGHT, SKYLIGHT and MOONLIGHT trials, malignant neoplasms were only observed in SKYLIGHT 4.

The EAG do not find this surprising because SKYLIGHT 4 is the only long-term (52 week) randomised trial evaluating the 45mg dose of fezolinetant, and longer trials should provide longer latencies for neoplasm development. MOONLIGHT 3 was also 52 weeks but was a non-comparative trial using the lower dose of 30mg. Although the company claims that there were no malignant neoplasms in MOONLIGHT 3, 1/150 developed endometrial adenocarcinoma.[49] In the 52 week SKYLIGHT 1 extension, 1.32% of participants using fezolinetant 45mg experienced a malignant neoplasm event,[45] which is close to the proportion in the fezolinetant 45mg group with a neoplasm event in the SKYLIGHT 4 trial (1.48%).[45] However, in the 52 week SKYLIGHT 2 extension, there was just one benign neoplasm and no malignant events in the fezolinetant 45mg group.[45]

5. A thorough re-analysis of all neoplasm events in phase 2 and 3 studies suggested that a drug effect was unlikely because of short latencies, tumour type heterogeneity, risk factor history and plausible alternative aetiologies (i.e. pre-existing conditions).

The EAG note that no reference is given for this analysis, and so the EAG is unable to check the validity of this claim.

Taking all information into consideration, the EAG considers that the safety concerns raised by Douxfils et al. 2024[48] may be valid until further data suggest otherwise. The neoplasm safety data should be taken into account by clinicians and patients when making treatment choices.

Liver test elevations

The CS reports that liver test elevation findings, based on alkaline phosphatase (ALP), alanine aminotransferase (ALT), aspartate aminotransferase (AST) and total bilirubin (TBL), were similar across groups (Table 13). The thresholds for abnormality for these data were not reported in the CS, and do not appear to be available in the trial CSRs.

Table 13: Liver test elevations in the 24 week DAYLIGHT trial and the 12 week SKYLIGHT 1 and 2 trials

Trial	Placebo	Fezolinetant 45mg
DAYLIGHT	6/226 (2.7%)	10/226 (4.4%)
SKYLIGHT1	5/175 (2.9%)	7/173 (4.0%)
SKYLIGHT2	0/167	3/167 (1.8%)

Source: Tables 44 and 53 of the CS

The CS also reports on 'special liver tests' carried out in the DAYLIGHT and SKYLIGHT 1, 2 and 4 trials. These appear to be the same tests as above, but with a defined, and presumably higher, threshold for abnormality. In the DAYLIGHT trial, 3/224 patients in the fezolinetant 45 mg group had ALT or AST levels that were three times the upper limit of normal (ULN) compared to none in the placebo group. In addition, 3/224 patients in the fezolinetant group had alkaline phosphatase levels that were more than 1.5 times ULN, compared to 1/220 patients in the placebo group. Similarly, in the SKYLIGHT 2 trial, 3 fezolinetant 45mg patients had ALT or AST levels >3 times ULN during the 12 week double blind period, compared to one patient in the placebo group. Likewise, in the SKYLIGHT 4 trial, ALT increases of > 3 times ULN occurred at a higher rate in the fezolinetant group than the placebo group (1.9% compared to 0.9%) over the course of the study.

The FDA issued a warning on 12th September 2024 about a rare occurrence of serious liver injury with the use of fezolinetant for VMS.[50] This was based on a review of a post-marketing case of a patient with elevated ALT, AST, ALP and total bilirubin. This patient also had signs and symptoms of liver injury. These events had occurred within 40 days of starting the treatment. Clinical improvement and normalisation of test results eventually occurred after cessation of the treatment. The FDA provides advice on methods to reduce risk in clinical practice[50], which includes baseline and subsequent liver function tests.

The FDA warning, as well as the data provided by the CS on liver test elevations and special liver tests, indicate that this may be important. The company were asked during clarification whether similar recommendations are expected from the MHRA in the near future, and how they would be implemented in the NHS. The company responded by stating that, "*discussions around hepatic laboratory tests for patients receiving fezolinetant are ongoing with the EMA and shortly Medicines and Healthcare Products Regulatory Agency (MHRA). The expected result of these discussions is currently unknown. However, outcomes could be expected within the next year*." Therefore, there is potential for additional liver function tests to be required for patients taking fezolinetant in the future.

Further to this, previous trials[51, 52] of another NK3R antagonist, pavinetant (MLE4901), were discontinued due to transient elevations of liver transaminase enzymes in some postmenopausal

women. Although these events were attributed to the chemical structure of pavinetant rather than the class of NK3R antagonists, this evidence alongside the results of the fezolinetant trials demonstrates a need to carefully monitor liver function over time in people receiving fezolinetant for VMS.

3.2.3.1 Overall, drug-related, and serious TEAEs, TEAEs of special interest and special tests undertaken

TEAEs were defined as adverse events observed after starting administration of study intervention and either up to 21 days after the last dose of study intervention, or first dose of study intervention during the active treatment extension period, whichever occurred first.

Overall TEAEs

Overall TEAEs were similar across fezolinetant 45mg and placebo groups in the DAYLIGHT and three SKYLIGHT trials. See tables 41, 48 and 49 in the CS and Table 40 in the CS Appendix F.

Drug-related TEAEs

Although there were few group differences in the SKYLIGHT 1 and SKYLIGHT 4 trials, there was a notably greater number of drug-related TEAEs in the fezolinetant 45mg group than the placebo group in DAYLIGHT (39/226 compared to 25/226) and SKYLIGHT 2 (25/167 compared to 11/167). See tables 41, 48 and 49 in the CS and Table 40 in the CS Appendix F.

Serious TEAEs

Serious TEAEs were at comparable levels in the fezolinetant 45mg and placebo groups for the DAYLIGHT trial (10/226 compared to 8/226), and the SKYLIGHT 1 (2/173 compared to 1/175) and SKYLIGHT 2 trials (2/167 compared to 0/167). However, in the SKYLIGHT 4 trial, serious TEAEs were experienced by more people in the fezolinetant 45mg group than the placebo group (23/609 compared to 14/610). Across all trials, no deaths were reported. Further details are provided in Tables 41, 43,45, 48 and 49 in the CS and Table 40 in the CS Appendix F.

TEAEs of special interest

TEAEs of special interest during the DAYLIGHT trial and the SKYLIGHT 1 and 2 trials are summarised below (Table 14). There were no clear differences between groups. Note that this does not include neoplastic or hepatic TEAEs, which are reported separately in section 3.2.6.1.

Table 14: TEAEs of special interest in the 24 week DAYLIGHT trial and the 12 week SKYLIGHT 1 and 2 trials

TEAE of special interest	Trial	Placebo	Fezolinetant 45mg
Uterine bleeding	DAYLIGHT	10/226 (4.4%)	6/226 (2.7%)
	SKYLIGHT1	2/175 (1.1%)	2/173 (1.2%)
	SKYLIGHT2	1/167 (0.6%)	1/167 (0.6%)
Thrombocytopenia	DAYLIGHT	1/226 (0.4)	0/226
	SKYLIGHT1	1/175 (0.6%)	1/175 (0.6%)
	SKYLIGHT2	0/167	0/167
Bone fractures	SKYLIGHT1	0/175	1/173 (0.6%)
	SKYLIGHT2	1/167 (0.6%)	0/167
Potential abuse liability	SKYLIGHT1	0/175	0/173
	SKYLIGHT2	1/167 (0.6%)	0/167
	SKYLIGHT4	0/610	1/609 (0.2%)
Depression	SKYLIGHT1	2/175 (1.1%)	3/173 (1.7%)
	SKYLIGHT2	4/167 (2.4%)	1/167 (0.6%)
	SKYLIGHT4	13/610 (2.1%)	11/609 (1.8%)
Wakefulness	SKYLIGHT1	1/175 (0.6%)	1/173 (0.6%)
	SKYLIGHT2	1/167 (0.6%)	1/167 (0.6%)
	SKYLIGHT4	4/610 (0.7%)	4/609 (0.7%)
Effect on memory	SKYLIGHT1	0/175	0/173
	SKYLIGHT2	0/167	0/167
	SKYLIGHT4	1/610 (0.2%)	1/609 (0.2%)

Source: Adapted from Tables 44 and 53 of the CS, and Table 47 of the CS appendices. AEs were coded using MedDRA.
Abbreviations: TEAE: treatment emergent adverse event.

Special tests

Special tests were carried out for endometrial thickness and central nervous system (CNS) effects. There were no clear differences across groups for endometrial thickness in any of the trials (see tables 46 and 54 in the CS and Table 44, CS Appendix F). Few differences were observed for CNS effects in any of the trials, although in SKYLIGHT 4 there were higher rates of insomnia in the fezolinetant group (24/609 compared to 11/610) (see section B.2.10.2.9 in the CS and Table 47, CS Appendix F).

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

3.3.1 Company's exploratory NMA

In response to the EAG's first clarification question, the company stated that UK clinical experts emphasised that non-hormonal treatments, which are mainly prescribed off-label (except clonidine), are suboptimal in terms of their effectiveness and have unpleasant side effects. The company then presented results from an exploratory NMA to generate comparative clinical evidence for fezolinetant versus non-hormonal alternatives.

Eligibility criteria for the NMA were largely based on the protocol for the company SLR, although the only outcomes considered were VMS severity and frequency, and there were additional requirements that a drug had to be licenced for any indication, and the doses had to be at levels in line with UK clinical practice. Although the fezolinetant data in the NMA were from trials that were limited to people who were HRT-unsuitable (or limited to the sub-groups of people who were HRT-unsuitable) this was not the case for the non-fezolinetant trials. The NMA considered 15 trials, 14 of which were from the company SLR,[32-34, 53-62] with the additional trial from DAYLIGHT.[31, 63] Of these 15 trials identified by the SLR, only three fezolinetant versus placebo trials[31-33] and two paroxetine versus placebo trials (N30-003 and N30-004)[34] were included in the exploratory NMA. Thus, 10 trials were excluded.

Two of the 10 exclusions appeared to have been appropriately excluded, as they involved outcomes[61] or follow up times[62] that were not in line with the protocol and other studies, and which could therefore lead to inconsistency in the network. However, the other eight exclusions were not based on such considerations of exchangeability.

Five of these were desvenlafaxine versus placebo trials,[53-55, 59, 60, 64] which were excluded because desvenlafaxine is not licensed in the UK. Licence status is not a valid reason for exclusion because a licence is not a prerequisite for clinical use in the NHS,[65] and, in any event, most of the non-hormonal pharmacological comparator treatments listed in the NICE scope are not licensed for VMS anyway. However, if desvenlafaxine has not been licensed for *any* indication in the UK it may be difficult or impossible to prescribe, and thus could be regarded as being of limited relevance. Nevertheless, the EAG considers that it could have been included in the NMA, to represent SNRIs as a class, along with evidence supporting a class effect. It is also worth pointing out that although the company had stated that licensing was a pre-requisite inclusion criterion for the NMA, the company included these studies in their SLR.

The remaining three exclusions were gabapentin versus placebo studies,[56-58] which were excluded for having a daily dose of 1200mg, which exceeds the maximum 'UK clinical practice' dose of 900mg daily. Although the British National Formulary (BNF) suggests a daily dose of 900mg for VMS, doses up to 3600mg are allowed for other conditions, suggesting that 1200mg is not considered unsafe in an absolute sense.

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

Company's exploratory NMA

The methods described by the company for the exploratory NMA,[30] appear appropriate, although the NMA report is incomplete[30]. For example, code provided in response to clarification question A8 for

the fixed effect model cannot be validated as some components appear to be missing. No code is provided for the random effects model and no details on the prior distributions used for the heterogeneity parameter are given. Detailed convergence checks and model fit results (e.g. residual deviance, heterogeneity estimate) are also not provided. Since no executable code files were provided, the EAG cannot fully verify the results presented.

In the company's exploratory NMA,[30] the fixed effect (FE) models fit the data marginally better than the random effects (RE) models. The FE result suggested that fezolinetant 45mg was more effective than paroxetine 7.5mg for change in frequency of moderate/severe VMS (mean difference [redacted]), although the 95% CrI was wide and included no effect using the RE model (mean difference [redacted]). The EAG believes that the results of the NMA are incomplete because of the inappropriate exclusions of other comparators. However, given that all studies of other comparators were placebo controlled, addition of other comparators to the network would not affect the comparison of fezolinetant to paroxetine.

3.5 Additional work on clinical effectiveness undertaken by the EAG

Given that the company provided NMA efficacy results for only one comparator, the EAG sought to identify any other relevant systematic reviews and NMAs which compared fezolinetant to other non-hormonal treatments covered by the appraisal's scope. A relevant study by Morga et al 2023[25] was cited in the CS (on p148).

Methods of Morga et al 2023

This study was sponsored by the company (Astellas Pharma) and several authors (including the first author) are employees of the company. The review protocol was not registered on PROSPERO nor on any other registry. The review compared the efficacy of fezolinetant 45 mg with hormone therapy and non-hormonal therapies for VMS in postmenopausal women.

A range of appropriate bibliographic databases were searched, although the searches were conducted in June 2021, so the review is somewhat out of date. This meant that the review included (pooled) data from the SKYLIGHT 1 and 2 trials, but not from the DAYLIGHT trial, as it had not begun. Eligible trials were phase 3 or 4 RCTs in postmenopausal women with ≥7 moderate to severe VMS per day or ≥50 VMS per week. Eligible non-hormonal therapy treatments were fezolinetant, SSRIs/SNRIs, clonidine, gabapentin extended release (ER), oxybutynin and elinzanetant. Trials had to report one or more of the following outcomes:

- Mean change from baseline to week 12 in frequency of moderate to severe VMS
- Mean change from baseline to week 12 in VMS severity

- VMS frequency as a binary response: proportion of participants with ≥ 75% reduction from baseline in VMS frequency at 12 weeks

Studies reporting outcome definitions comparable to those used in the SKYLIGHT trials (see CS p51) were deemed appropriate for inclusion in the NMA. Results were reported as posterior mean treatment differences for continuous outcomes and posterior mean ORs for binary outcomes. The EAG considers that the use of posterior median ORs would have been better for reporting, given that the forest plots for VMS frequency response look skewed.

The use of JAGS via R – code from NICE DSU TSD 2[66] was appropriate, although the code was only reported for the fixed-effect model, not for the random effects model, so the prior distributions used for heterogeneity are unclear. I^2 and Cochran's Q were used to identify heterogeneity for pairwise comparisons, but the main paper presented only fixed-effect results.

Results of Morga et al 2023

Although 23 comparator publications were included in the NMA, most related to hormonal therapies; the results reported here focus only on the comparisons between fezolinetant 45mg and non-hormonal therapies (and not placebo).

Risk of bias

The EAG disagrees with the risk of bias judgements made in the review, which were presented in a supplementary file; no supporting information was provided to justify how decisions were arrived at. The authors considered that trials generally had a low risk of bias. For the fezolinetant SKYLIGHT trials, issues with bias due to missing data have already been discussed in Section 3.2.1.1.

For the comparator trials, the EAG also noted bias problems arising from missing data. The following desvenlafaxine trials all used the last-observation-carried-forward (LOCF) approach for handling missing data: Pinkerton et al 2012,[59] Speroff et al 2008,[64] Bouchard et al 2012,[55] and Archer et al 2009.[53] Pinkerton et al 2014 also used LOCF to impute missing data in their trial of gabapentin vs placebo.[58] Given these methods, and given that data in these trials were missing at rates typically between around 15% to 35% – mostly due to adverse events – the EAG considers that results for the continuous outcomes from these trials are at a high risk of bias due to missing data. The paroxetine trial by Simon et al 2013[34] used mean VMS data from the previous week to impute missing data. This trial is also judged by the EAG to have a high risk of bias due to missing outcome data. All of these non-hormonal therapy trials were rated by Morga et al as being at low risk of attrition bias (i.e. bias due to missing outcome data) and were included in one or more of the NMAs.

Methods for how missing data were handled for the binary responder outcome were either unclear or not reported for all the non-hormonal trials included in the NMA, except for the analyses in the

SKYLIGHT trials which the EAG judged to be at low risk of bias. In the absence of such details the EAG assumes LOCF was also used, so the comparator trial results are judged to be at high risk of bias due to missing outcome data for the binary outcomes.

NMA results

Estimates of heterogeneity were not provided, so it was not possible to judge whether this was appropriately estimated, but since there are more than three studies for some pairwise comparisons it is assumed that it was. However, the heterogeneity for the severity network was primarily estimated based on the comparisons of particularly placebo vs desvenlafaxine (6 studies) and placebo vs tibolone (2 studies), which were identified as having a large heterogeneity in pairwise comparisons, but it is unclear whether this heterogeneity would also apply to other comparisons, such as fezolinetant 45mg vs placebo (which cannot be reliably estimated as there is only one data point in the network comprised of the pooled SKYLIGHT 1 and 2).

To check for inconsistency, an inconsistency model was compared to the NMA model. However comparison of the residual deviances for each data point[67] were not provided or commented on, so the inconsistency assessment is incomplete. However, there is only limited indirect evidence so finding evidence of inconsistency in this network is unlikely (although this does not mean inconsistency is not present). The model fit statistics suggested that the random effects model fitted better than the fixed effect (FE) model (also termed common effect model) for both VMS frequency (Deviance information criterion (DIC) RE=149, FE=164) and VMS severity (DIC RE=-36, FE=-29), but there was a similar fit for the binary VMS frequency response outcome (DIC RE=234, FE=237).

The results for the RE model indicated no evidence of differences between fezolinetant 45mg and non-hormonal therapies for all three outcomes. The comparators to fezolinetant 45mg included in the NMAs were:

- *Change in frequency of moderate to severe VMS*: paroxetine 7.5mg (1 trial), desvenlafaxine (6 trials: results for 4 different doses, ranging between 50mg and 200mg) and gabapentin ER 1800mg (1 trial)
- *Change in severity of moderate to severe VMS*: desvenlafaxine (5 trials: results for 4 different doses, ranging between 50mg and 200mg) and gabapentin ER 1800mg (1 trial)
- *VMS frequency as a binary response*: desvenlafaxine (5 trials: results for 4 different doses, ranging between 50mg and 200mg)

For the FE model (which had a worse model fit) the results suggested that fezolinetant 45mg was statistically significantly better than all non-hormonal comparators for change in frequency of moderate to severe VMS. However, the clinical significance of the difference appears limited, being

between 1-2 events (depending on the comparator). This interpretation is supported by the limited evidence of concurrence of results when VMS frequency was analysed using a more clinically important binary responder outcome: the only meaningful difference seen was between fezolinetant 45mg and the lowest desvenlafaxine dose (50mg); relevant data were unavailable for paroxetine and gabapentin.

Furthermore, the EAG considers the fezolinetant SKYLIGHT trials to be the only trials in the responder network with results which have a low risk of bias; the EAG considers the comparator trial results to be at high risk of bias, so results of the NMA should be interpreted with caution. No differences between treatments were seen for change in severity of moderate to severe VMS using the fixed effect model.

The EAG concludes that the results from this NMA indicate that there is little evidence to suggest that fezolinetant 45mg has a clinically-meaningful benefit over SSRIs, SNRIs and gabapentin for reducing the frequency of moderate to severe vasomotor symptoms, and that there is no evidence to suggest any benefit for reducing the severity of vasomotor symptoms. Although the EAG also noted some methodological and reporting limitations with the NMA, its results are somewhat uncertain, primarily due to the high risk of bias due to missing outcome data seen across trials.

3.6 Conclusions of the clinical effectiveness section

The evidence presented on the efficacy and safety of fezolinetant 45mg was based on the results of four placebo-controlled randomised trials. The EAG considers that the efficacy results should not be considered as reliable estimates of the treatment effects expected to be seen in the NHS setting. The main trial limitations were the use of over-optimistic methods when dealing with missing data, and the recruitment of selected cohorts which all had large numbers of moderate to severe VMS events per day. The latter is important because a subgroup analysis indicated that trial efficacy estimates for improvements in VMS frequency appear to be driven by patients with many (~10 or more) VMS events per day. These factors may have biased the VMS frequency results in favour of fezolinetant. Although statistically significant improvements in VMS severity were also reported across the three RCTs, the clinical significance of the results is questionable. Fezolinetant did not improve quality of life by a clinically significant amount.

The EAG is also concerned by the company's exclusion of non-hormonal treatments as relevant comparators; the company's submission did not present a network meta-analysis comparing fezolinetant with a range of non-hormonal treatments. The EAG identified a relevant NMA, published in 2023, which was conducted by the company. The EAG's interpretation of this NMA's results is that there is little evidence to suggest that fezolinetant has a clinically-meaningful benefit over SSRIs,

SNRIs and gabapentin for reducing the frequency of moderate to severe vasomotor symptoms, and there is no evidence to suggest any benefit for reducing the severity of vasomotor symptoms.

Fezolinetant's safety profile appears somewhat uncertain with respect to the risk of developing neoplasms and serious liver injuries.

4 COST EFFECTIVENESS

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

The company conducted a SLR to identify economic evaluations for the treatment of VMS associated with menopause. The original search was performed in June 2023 and updated in April 2024. See Appendix G of the CS for a detailed description of the searches and results of the review. One publication was identified through the SLR update that compared the cost-effectiveness of fezolinetant with no pharmacologic treatment for The Institute for Clinical and Economic Review in the US. [68] Table 55 of the CS summarises the cost-effectiveness model, patient population, quality-adjusted life years (QALYs), costs, and comparative cost-effectiveness of fezolinetant compared to no pharmacologic treatment. The resulting incremental cost per QALY gained was $390,000.

Points for critique

The literature searching for the company's review of cost-effectiveness evidence appears to have been conducted to a high standard and is well reported – See Appendix 2 for details. The EAG considers that all relevant publications are likely to have been identified for the HRT-unsuitable population. The publication identified by SLR is a cost-effectiveness study of fezolinetant with no pharmacologic treatment for The Institute for Clinical and Economic Review in the US. The EAG notes several key features which differ from the company's cost-effectiveness analysis. These features pertain to:

- Baseline population characteristics: Beaudoin et al. study[68] used SKYLIGHT 2 as the source of baseline population characteristics rather than DAYLIGHT. The mean age at baseline was 54.3 years, which is similar to the mean age of 54.5 from DAYLIGHT.
- Comparators: The base-case comparison in Beaudoin et al.[68] focused on no pharmacologic treatment as estimated by the placebo arm of the fezolinetant clinical trials (SKYLIGHT 1 and 2). For the purposes of adding context to the fezolinetant versus no pharmacologic treatment comparison, treatments currently being used for reducing VMS associated with menopause (HRT) were also compared to no pharmacologic treatment in a scenario analysis. The resulting cost per QALY gained was $16,000.
- Model structure: The study has three mutually exclusive and exhaustive health states: 1) on treatment: responding or not responding (those that discontinued due to the intervention not improving symptoms); 2) off treatment (discontinued due to symptom resolution); and 3) all-cause death. It did not categorize health states by frequency of moderate to severe VMS per day.
- Time horizon: Outcomes were estimated over a lifetime time horizon rather than a 10-year time horizon used in the company's model.

- Transition probabilities: The study used Avis et al., 2015[5] to inform the transition probabilities for moving from on treatment to off treatment. The Study of Women's Health Across the Nation (SWAN) was a longitudinal study spanning 17 years with objectives of determining total duration of frequent VMS (defined as symptoms on ⩾6 days in the last 2 weeks) during the menopausal transition and to quantify how long frequent VMS persist after the final menstrual period. In the CS, the model used efficacy from the DAYLIGHT and SKYLIGHT trials to inform the transition probabilities for movement across different moderate to severe VMS frequency health states.
- Treatment effect: The key treatment effects from the fezolinetant trials (SKYLIGHT 1 and 2) included reduction in frequency and severity of VMS as compared to placebo using the 12-week endpoint in the trials. Average VMS frequency per cycle was summed over the model time horizon. The difference in VMS episodes on average per cycle (annual) was compared between fezolinetant and placebo and then divided by the average number of VMS episodes per day to come to an equivalent estimate of the total number of symptom-free days. In the CS, the treatment effect included a reduction in moderate to severe VMS frequency per day from DAYLIGHT up to week 24, pooled SKYLIGHT 1 and 2 from week 24-52 and a constant average based on pooled SKYLIGHT 1 and 2 from week 52 onwards for fezolinetant, while the placebo effect from DAYLIGHT was applied to week 12 only.
- Discontinuation: Treatment-specific discontinuation due to lack of efficacy was applied in the first year. Beyond the first year, discontinuation was associated with resolution of symptoms (people who responded to therapy were assumed to remain on treatment until the natural cessation of VMS). The company's model applied treatment-specific discontinuation rates to all model cycles for fezolinetant over the full model duration and for no active treatment until week 12.
- Utility values: HRQoL was derived using a mapping algorithm between the MENQOL and EQ-5D questionnaires, due to no direct utility scores. The mapping instrument derived a linear relationship of EQ-5D = 0.992 – 0.042*MENQOL. However, MENQOL includes more health domains related to menopause and symptoms associated with menopause, which means the utility values were associated not only with VMS, but also other symptoms correlated with VMS. In the company's model, EQ-5D data from DAYLIGHT, with adjustment based on clinical input, were used.

The longest placebo-controlled trials available were 12 weeks for fezolinetant in Beaudoin et al. study[68] (SKYLIGHT 1 and 2). Thus, the study lacked long-term placebo-controlled efficacy data, which created uncertainties in the long-term effectiveness of fezolinetant. Furthermore, without direct utility scores, the study used a linear mapping algorithm between MENQOL and EQ-5D, which captured not only menopause, but also other symptoms correlated with menopause. Due to the

limitations of the evidence available at the time of Beaudoin et al, the EAG considers that this study is not closely related to the company's decision problem.

The EAG notes that there is a cost-effectiveness evaluation of two continuous-combined hormone therapies for moderate to severe VMS which was not identified in the company's SLR.[69] However, the interventions considered were HRT, which is not related to the company's decision problem in an HRT-unsuitable population, and the model structure consisted of more general health states, which did not distinguish different severity levels and/or frequencies of VMS.

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

The company submitted a *de-novo* model to compare the cost-effectiveness of fezolinetant with no active treatment in menopausal people with moderate to severe VMS frequency of ≥ 7 events per day, who are deemed unsuitable for HRT.

A Markov cohort model structure is used to estimate long-term health outcomes and costs based on menopausal people transitioning between four VMS frequency health states representing different frequencies of moderate to severe VMS over a 10-year time horizon. All menopausal people have a probability of natural cessation of VMS and risk of all-cause mortality, which is assumed not to be affected by treatment. When people discontinue fezolinetant, they follow the natural history of moderate to severe VMS in the absence of treatment, while those on no active treatment follow the natural history.

Fezolinetant is modelled to affect QALYs by increasing the proportion of people who transition from a high moderate to severe VMS frequency of ≥ 7 events per day to low moderate to severe VMS frequencies of between 0 and 2, between 2 and 7 and between 7 and 9 (from frequency of ≥ 9) events per day, which is associated with improved HRQoL, compared to no active treatment which is modelled to have a larger percentage of people in the high moderate to severe VMS frequency health states (i.e., ≥ 7 events per day).

Fezolinetant is modelled to affect costs by having higher drug acquisition costs compared to no active treatment but a higher percentage of menopausal people in the low moderate to severe VMS frequency health states, which are associated with lower resource use and costs than the high moderate to severe VMS frequency health states.

The appropriateness of the data and assumptions used in the model are discussed in the relevant sections below.

4.2.1 NICE reference case checklist

The model submitted by the company is assessed in relation to the NICE reference case in Table 15.

Table 15 NICE reference case checklist

Element of health technology assessment	Reference case	EAG comment on company's submission
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	The CS is appropriate.
Perspective on costs	NHS and PSS	The CS is appropriate.
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	The CS is appropriate.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	The CS is not appropriate. The time horizon in the model is 10 years but median VMS cessation in the model is 7.4 years, with approximately 40% of postmenopausal people still experiencing VMS by year 10 and a small proportion still receiving treatment with fezolinetant.
Synthesis of evidence on health effects	Based on systematic review	The CS is partly appropriate. The company included DAYLIGHT and SKYLIGHT 1 & 2 trials in the model. However, the company applied the natural history study for the comparator arm which was based on structured expert elicitation (SEE) estimates, without a SLR on studies of natural history of VMS.
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults.	The CS is not appropriate. HROoL was measured with EQ-5D-5L and valued using the UK tariff. The EQ-5D-5L was converted to EQ-5D-3L using appropriate algorithms by Hernández Alav.[70] However, the EQ-5D values used in the model were modified based on clinical opinion.
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	The CS is appropriate.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	The CS is appropriate.
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	The CS is appropriate.
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	The CS is appropriate.

Discounting	The same annual rate for both costs and health effects (currently 3.5%)	The CS is appropriate.
CS: company submission; PSS: personal social services; QALYs: quality-adjusted life years; HRQoL, health-related quality of life; EQ-5D: standardised instrument for use as a measure of health outcome.		

4.2.2 Model structure

4.2.2.1 Summary of company submission

The model used in the CS is a Markov cohort model that consists of six health states: Four health states representing different frequencies of moderate to severe VMS per day (0 ≤ VMS Frequency < 2, 2≤ VMS Frequency < 7, 7 ≤ VMS Frequency < 9, and VMS Frequency ≥ 9) and two remaining health states representing natural VMS cessation and all-cause death (see Figure 2). The company justified the use of VMS frequency for the health states based on discussions with UK clinical experts and four patient experts who considered VMS frequency to be a more objective measure of the health effects of treatment compared to VMS severity; however, it is also stated in the CS that some clinical experts suggested that VMS severity should be captured alongside VMS frequency in the model. Furthermore, the CS states that published literature and clinical guidelines, such as NICE NG23, support the decision to use VMS frequency as a proxy for VMS severity.[71-73] The DAYLIGHT and SKYLIGHT trials that are used as the evidence in this submission evaluated frequency and severity of VMS as two separate trial outcomes. The CS states that capturing both frequency and severity outcomes in the model may be feasible in theory, but it would not be possible to obtain precision estimates, which are needed for a probabilistic analysis. Consequently, the company deemed it appropriate to model VMS health states based on frequency of moderate to severe VMS rather than severity. In response to EAG clarifications (question B1), the company included additional justification for this decision by providing a correlation analysis of frequency and severity outcomes from the SKYLIGHT trials. The correlation coefficients were between [redacted] and [redacted] for SKYLIGHT 1 and between [redacted] and [redacted] for SKYLIGHT 2, while the coefficients for pooled SKYLIGHT 1 and 2 were between [redacted] and [redacted]. The company states that this analysis demonstrates that VMS frequency and severity are correlated, indicating that the frequency-based health states in the model implicitly captures severity.

The cut-off thresholds used to define the moderate to severe VMS frequency health states were based on a statistical analysis of the distribution of average daily moderate to severe VMS frequency in DAYLIGHT (Figure 26 in CS) and utility values associated with each average daily moderate to severe VMS frequency health state (Table 57 in CS). These cut-off thresholds were validated by UK clinical experts for clinical plausibility. The detailed rationale for choice of each cut-off is provided in CS Section B.3.2.2.1 and in the appendix of the Structured Expert Elicitation (SEE) report.[74] In

summary, the company ran several generalized estimating equations (GEE) models, which were based on a weekly average frequency of moderate to severe VMS per day collected from the DAYLIGHT and SKYLIGHT trials with EQ-5D utility index being the continuous dependent variable and health state defined by moderate to severe VMS frequency cut-offs as the categorical independent variable. EQ-5D scores based on the pooled data of fezolinetant and placebo from all included participants at baseline, week 4, week 12, week 16 and week 24 were used to model non-treatment specific EQ-5D utilities. In the analysis, VMS frequency of seven per day and above was based on the DAYLIGHT eligibility criteria (i.e., a minimum average of seven moderate to severe VMS per day, within the last 10 days prior to randomization). The other cut-offs for other frequency categories were based on statistically significant differences in utility values among individuals experiencing below two moderate to severe VMS per day and those who experience between two and six moderate to severe VMS per day. The company ran a total of seven models based on the DAYLIGHT trial alone to explore different cut-off thresholds, and after choosing the model that was assessed as the most relevant (model 7 in Table 16), it was validated and supported by the pooled data of DAYLIGHT and SKYLIGHT 1 and 2. Table 16 presents each model explored by the company, the assumed cut-offs and utility estimates based on DAYLIGHT.

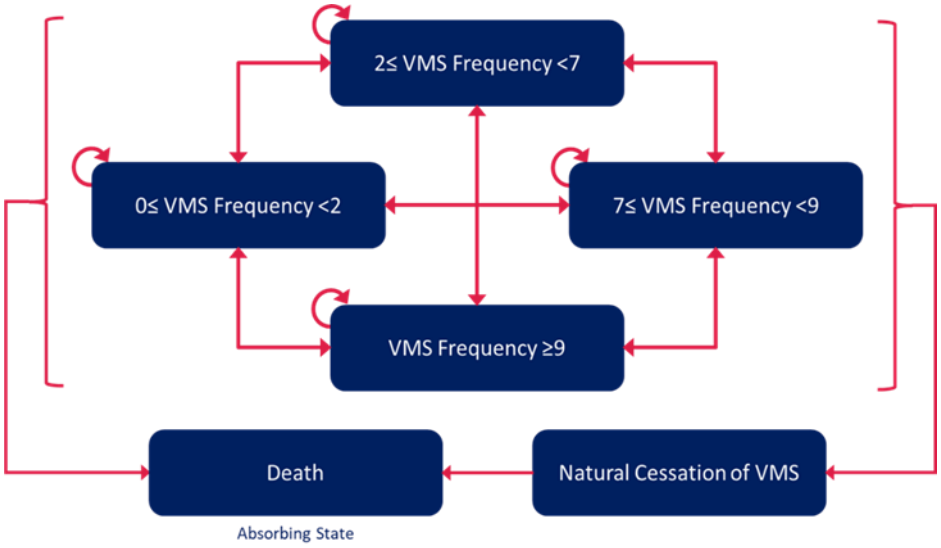
Table 16 GEE model results based on DAYLIGHT trial alone for all models explored by the company, including assumed cut-offs and utility estimates

Covariate	Parameter Estimate	Standard Error	Mean Estimate (95% CI)	Chi-Square	Pr > ChiSq*
Model 0					
Intercept	[redacted]	[redacted]	-	-	
=0	[redacted]	[redacted]	[redacted]	-	-
> 0 to < 1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
1 to < 2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2 to < 3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
3 to < 4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
4 to < 5	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
5 to < 6	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
6 to < 7	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
7 to < 8	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
8 to < 9	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
9 to < 10	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
10 to < 11	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
11 to < 12	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12 to < 13	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
13 to < 14	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥ 14	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Model 1					
Intercept	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
0 to < 1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
1 to < 2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2 to < 3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
3 to < 4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
4 to < 5	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
5 to < 6	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
6 to < 7	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
7 to < 8	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
8 to < 9	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
9 to < 10	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
10 to < 11	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
11 to < 12	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12 to < 13	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
13 to < 14	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥ 14	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Model 2					
Intercept	[redacted]	[redacted]	-	-	-
0 to < 2	[redacted]	[redacted]	[redacted]	-	-
2 to < 3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
3 to < 7	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
7 to < 12	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥ 12	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Model 3					
Intercept	[redacted]	[redacted]			
0 to < 2	[redacted]	[redacted]	[redacted]	-	-
2 to < 5	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
5 to < 7	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
7 to < 9	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥ 9	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Model 4					
Intercept	[redacted]	[redacted]			
0 to < 2	[redacted]	[redacted]	[redacted]	-	-
2 to < 7	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
7 to < 11	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥ 11	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Model 5					
Intercept	[redacted]	[redacted]			
0 to < 2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2 to < 7	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
7 to < 12	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥ 12	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Model 6					
Intercept	[redacted]	[redacted]			
0 to < 2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2 to < 7	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
7 to < 13	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥ 13	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Model 7: Final model					
Intercept	[redacted]	[redacted]			
0 to < 2	[redacted]	[redacted]	[redacted]	-	-
2 to < 7	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
7 to < 9	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
≥ 9	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Figure 2 Health state transition diagram (from Figure 25 in the CS)



Abbreviations: VMS: vasomotor symptoms

Two cohorts of menopausal people were modelled: one cohort for fezolinetant and the other for no active treatment. Upon entering the model, each cohort was distributed across the four moderate to severe VMS frequency health states using the baseline starting distribution from DAYLIGHT. A scenario analysis was performed that applies the baseline distribution from pooled DAYLIGHT and SKYLIGHT 1 and 2 trials.

As presented in Figure 2, in every 4-weekly cycle, the modelled individuals can remain in the same health state, transition to one of the other VMS frequency health states, experience natural cessation of VMS or die due to any cause. The probability of natural cessation of VMS is the same for both

treatment arms and is independent of the frequency of moderate to severe VMS and the time since entering the model.

In both treatment arms, the modelled cohort can be either 'on-treatment' or 'off-treatment'. In the fezolinetant arm, the modelled cohort 'on-treatment' is distributed across the health states in each 4-weekly cycle according to the transition probabilities reflecting the efficacy of fezolinetant. In each 4-weekly cycle, a proportion of the cohort may discontinue treatment and move to 'off-treatment,' where they are assigned transition probabilities reflecting the natural history of VMS in the absence of treatment. In the no active treatment arm, the modelled cohort 'on-treatment' is based on the transition probabilities of the placebo arm of the trials; however, the placebo effect is assumed only for the first 12 weeks in the model. Thereafter, the no active treatment cohort moves to the 'off-treatment' health states, where the transition probabilities follow the natural history of VMS. A description of the transition probabilities is provided in 4.2.6.1.

Points for critique

The EAG has several concerns about the model structure. The primary concern relates to the use of VMS frequency as a proxy for VMS severity. The company justified the use of frequency of moderate to severe VMS per day because it was the primary endpoint of the DAYLIGHT and SKYLIGHT trials, and due to the difficulty in capturing changes in the severity of moderate to severe VMS. Clinical advice to the EAG expressed concern about using VMS frequency as a proxy for VMS severity because a menopausal person with two severe night sweats in a 24-hour period, requiring a change of bed sheets and clothing, is not comparable to someone with a frequency of greater than seven per day where the severity of symptoms of each event are less severe. Here, it is important to highlight that the definition of moderate to severe VMS used in the design of DAYLIGHT and SKYLIGHT 1 and 2 did not include any requirement of VMS frequency. The definition used to design the studies were consistent with the FDA and EMA guidance and included the following: 1) mild: sensation of heat without sweating; 2) moderate: sensation of heat with sweating, able to continue activity; 3) severe: sensation of heat with sweating, causing cessation of activity.[21, 75] The eligibility criteria of the DAYLIGHT and SKYLIGHT trials included a requirement of having at least seven moderate to severe VMS per day, within the last 10 days prior to randomization. In NHS clinical practice, VMS frequency is not commonly used to define moderate to severe VMS. In the model, seven moderate to severe VMS per day represents the baseline cut-off to start receiving fezolinetant (see Section 3.2.2.2). At the same time, it is assumed in the model that women who experience a decrease in the frequency of moderate to severe VMS after treatment keep receiving the treatment (except for those who discontinue treatment based on all-cause discontinuation rates from the trials), even if they experience only 0 to <2 moderate to severe VMS per day (note that treatment discontinuation is considered independent of VMS frequency health states). It remains unclear how

the use of moderate to severe VMS frequency reflects the changes in VMS severity. Although the primary endpoint of DAYLIGHT was frequency of moderate to severe VMS, the EAG notes that a key secondary endpoint of DAYLIGHT was severity of moderate to severe VMS, while the co-primary endpoints of SKYLIGHT 1 and SKYLIGHT 2 were frequency of moderate to severe VMS and severity of moderate to severe VMS. In the CS, it is stated that developing a model structure that would incorporate both frequency and severity of moderate to severe VMS may be feasible in theory but was not done because calculation of precision estimates for these input parameters would not be possible for a probabilistic analysis. The EAG notes that the precision estimates for both the mean frequency of moderate to severe VMS (Figure 7 in the CS, p. 74) and the mean weighted severity of VMS experienced per 24 hours (Figure 8 in the CS, p. 76) are provided. Thus, the precision estimates needed to inform probability distributions for parameters to inform a probabilistic analysis should be possible. The company further justified their approach 1) by noting that a combined frequency and severity 'scoring scale' would lack credibility due to lack of validation and absence from common clinical practice, 2) with a correlation analysis of frequency and severity from the trials. However, no details were provided on the type of correlation analysis and only results for SKYLIGHT 1 and SKYLIGHT 2 were provided despite the fact that DAYLIGHT was used to define the health states in the model and estimate the health state utility values. Furthermore, when analysing the results of this correlation analysis, the EAG notes that the correlation estimates were all under the value of [redacted] with some as low as [redacted]. The EAG considers these results indicative of a moderate to weak correlation. Therefore, the EAG is concerned that the model structure based solely on frequency of moderate to severe VMS per day is not adequately capturing the impact of fezolinetant on the severity of moderate to severe VMS.

item 1. Frequency of moderate to severe VMS is not commonly used to define severity of VMS in NHS clinical practice.

The EAG's second concern relates to the definition of the health states. The assumptions underlying the cut-off thresholds were based on statistical analysis of the distribution of average daily moderate to severe VMS frequency in DAYLIGHT and utility values associated with each average daily moderate to severe VMS frequency health state. The company provided a rationale behind the cut-off thresholds, which was mostly based on a statistically significant difference between the frequency categories, a statistically significant difference in utility values associated with different frequency levels and validation with clinical experts on clinical plausibily. However, the EAG expresses several concerns with this approach. First, the utility values for health state 7≤ VMS Frequency <9 and the utility values for VMS frequency health state 2≤ VMS Frequency <7 differ only slightly (a difference of [redacted]) and the confidence intervals between these values overlap substantially. The difference between the other categories of VMS frequency differ more notably (a difference in utility of 0.04

between health states 0≤ VMS Frequency <2 and 2≤ VMS Frequency <7 and a difference of 0.38 between health states 7≤ VMS Frequency <9 and VMS Frequency ≥9). The EAG assessed the information about GEE model 7, which was provided in the appendix of the SEE report. The company assumed a frequency of seven moderate to severe VMS per day, which was used in the trial eligibility criteria, to be a fixed cut-off threshold. The company stated that this assumption was supported by the use of frequency of seven VMS per day as a cut-off in clinical trials to define moderate to severe VMS. The EAG notes that the utility values for frequency between 7 to < 8 and 8 to < 9 are higher than the utility values between 6 to < 7. Consequently, it seems that the company aimed to identify a GEE model that would provide utility values for the frequency below seven VMS per day to be lower than the utility values for the frequency of 7 to < 8 VMS per day. The EAG considers the justification for the choice of final model to be uncertain. Based on the description provided in the CS, the final cut-off thresholds were based on a statistically significant difference comparing each health state to the previous one and feedback from the clinical experts. However, the health state with VMS frequency 7 to < 9 is not statistically significantly different from the previous health state of VMS frequency 2 to < 7 and the p-value is high (p-value=0.356). The EAG notes that among the explored models, several other models appear to be associated with more favorable p-values (e.g., model 5). However, the provided information about the GEE models is too scarce for the EAG to make a more detailed assessment. The EAG notes that the company did not analyse other cut-off thresholds for defining moderate to severe VMS other than seven per day, which was based on FDA recommendations for VMS tral designs. In general, the EAG is also concerned about the use of utility values to define VMS frequency health states because the EQ-5D utility values are likely to reflect changes in quality of life due to other menopause symptoms over and above VMS.

item 2. Arbitrary cut-off thresholds are used to define moderate to severe VMS frequency health states in the model

4.2.3 Population

4.2.3.1 Summary of company submission

The patient population considered in the model is menopausal people with moderate to severe VMS for whom HRT is deemed unsuitable, in line with the population included in the DAYLIGHT trial, but narrower than the licenced indication 'for the treatment of moderate to severe VMS associated with menopause'.[76] The definition of mild, moderate and severe VMS in DAYLIGHT episodes is consistent with EMA[21] and FDA guidance[75] (i.e., moderate: sensation of heat with sweating, able to continue activity; severe: sensation of heat with sweating, causing cessation of activity) but eligibility criteria for DAYLIGHT required participants to have a minimum average of seven moderate to severe VMS per day, within the last 10 days prior to randomization. The definition of HRT-unsuitable used in the company's base case analysis includes:

- HRT-contraindicated (due to venous thromboembolism, cardiovascular disease, metabolic syndrome, severe hypertension, uncontrolled/complex diabetes mellitus, porphyria, etc);
- HRT-caution (medical risk assessment of the specific caution concluded that the risk of HRT outweighs the likely benefit);
- HRT-stoppers (menopausal people who have previously received HRT but no longer take HRT); and
- HRT-averse (menopausal people who have made an informed choice not to take HRT following a consultation on the benefits and risks of HRT).

A separate scenario analysis excludes HRT-averse people.

The baseline patient characteristics used in the model are based on the moderate to severe VMS frequency distribution of DAYLIGHT at randomization (42% with VMS frequency per day of greater than 7 and less than 9; 58% with VMS frequency per day of greater than 9; see Table 59, p152 of CS), while the mean starting age is 51 years based on the average age of menopause onset in the UK.

No separate subgroup populations are considered in the company's base case analysis.

Points for critique

The EAG's primary concern in relation to the population used in the model is whether it reflects menopausal people likely to receive fezolinetant in NHS primary care. Firstly, the EAG notes that both the licenced wording for fezolinetant and the company's target population as defined in the decision problem (Table 1, p15 of CS) does not specifically define moderate to severe VMS in terms of 'frequency' of moderate to severe VMS per day as used in the model. The EAG notes the lack of consistency in how VMS is classified in the literature, with frequency, severity, intensity, or a composite of these outcomes being the most reported vasomotor-related outcomes in studies, but with substantial variation in the definition of each outcome reported and measured in clinical trials. A systematic review of variation in menopausal VMS outcomes in clinical trials reported that across 214 trials, 49 different primary outcomes were identified for VMS and 16 different tools had been used to measure these outcomes, with little consistency in how the frequency and severity/intensity of VMS were defined and/or assessed.[77] In the absence of a standardised outcome measure, moderate to severe VMS is likely to be classified in different ways by different GPs in NHS clinical practice. Heterogeneity in moderate to severe VMS can also be expected in terms of how menopausal people classify their own symptoms. The EAG's clinical advisor strongly supported the view that VMS frequency per day is not an accurate proxy for the severity of VMS, with other measures (such as the MENQOL questionnaire and the Greene Climacteric Scale) available to assess the burden of symptoms associated with menopause for treatment. Importantly, the evidence used in the model to assess the cost-effectiveness of fezolinetant is for a DAYLIGHT eligible population and therefore

based on a population with a frequency of seven or more moderate to severe VMS, but this frequency limit may not be appropriate for assessing the need for treatment in the NHS.

The second point is that while the baseline patient characteristics for the population with moderate to severe VMS is based on the frequency distribution of VMS per day from DAYLIGHT at randomization (i.e., a frequency limit of greater than 7 per day), the mean starting age used in the model is based on the average age of menopause onset in the UK (51 years) rather than the mean age of 54.5 years from DAYLIGHT. The EAG considers it reasonable to reflect the mean age of menopause onset in the UK as the most relevant to NHS practice but notes the mismatch between an average population at the start of menopause and the population used in DAYLIGHT who are likely to have had VMS for several years before recruitment into the trial due to the higher mean age. To be eligible for DAYLIGHT, participants had to have a baseline minimum average of seven moderate to severe events of VMS per day, which the EAG considers to be a very high baseline VMS frequency when the company's SEE exercise used to estimate the natural history of postmenopausal women in the UK in the absence of treatment suggests a much lower proportion of ≥7 moderate to severe VMS per day at years 1, 3 and 6 (see Section 3.2.1.1). Therefore, the EAG considers the mean age of 51 years for menopause onset in the UK to be an appropriate baseline age in the model, but the EAG does not agree with the baseline distribution of a minimum average of seven moderate to severe events of VMS per day in the absence of treatment to be an appropriate baseline to represent menopausal people likely to receive fezolinetant in NHS primary care (see Section 3.2.1.2). The EAG believes that the baseline distribution should be closer to the elicited values at one year for the natural history of postmenopausal women in the UK because a significant reduction in VMS frequency would not be expected after one year in the absence of treatment. The proportions at one year for postmenopausal women with moderate to severe VMS and not receiving treatment elicited through an SEE study (without and with an addition validation by three UK clinical experts) are significantly different from the baseline distribution of DAYLIGHT (see Table 17). If such a rapid change in VMS were to be expected between baseline and one year in the absence of treatment, it would suggest a lack of need for treatments for moderate to severe VMS because the symptoms would resolve rapidly by themselves over time. This suggests that either the results of the company's SEE exercise are unreliable for assessing the natural history in the absence of treatment (see details in Section 4.2.6.2) or the company's baseline VMS frequency distribution derived from DAYLIGHT is more representative of a very high frequency of moderate to severe VMS and where participants are likely to have had symptoms for several years before recruitment into the trial and, therefore, not representative of moderate to severe VMS in the NHS in the absence of treatment. In response to EAG clarifications (question B11b), the company acknowledges that there is a disconnect between the baseline distribution and the natural history used in the model because the SEE exercise was not

anchored on the trial baseline distribution when eliciting values of natural history. The EAG considers this a major limitation of the SEE study.

Table 17 Baseline distribution estimates from DAYLIGHT in the absence of treatment and Year 1 estimates elicited for natural history in the absence of treatment

Source of estimates	0 ≤ VMS-F < 2	2 ≤ VMS-F < 7	7 ≤ VMS-F < 9	VMS-F ≥ 9
Baseline distribution from DAYLIGHT	0.00%	0.22%	41.81%	57.96%
Year 1 SEE natural history estimates	16.91%	47.13%	17.16%	18.80%
Year 1 clinician natural history estimates*	10%	30%	40%	20%

*Clinician natural history estimates are outputs of the SEE study that were additionally validated by three UK clinical experts in order to generate natural history inputs for the model base case.

The company considered alternative scenarios using mean age and baseline distribution from pooled SKYLIGHT and DAYLIGHT trial data, but these scenarios had minimal impact on the base case assessment of cost-effectiveness because the mean age and baseline distribution from the pooled analysis was very similar to DAYLIGHT.

In addition, although the product licence for fezolinetant is not restricted to postmenopausal people, no data is available for perimenopausal people (see Sections 2.3 and 3.1.2). The population of DAYLIGHT and the population used in the company's SEE exercise is menopausal people who met at least one of the following criteria:

- Spontaneous amenorrhea for ≥ 12 consecutive months;
- Spontaneous amenorrhea for ≥ 6 months with biochemical criteria of menopause (FSH > 40 IU/L);
- Had bilateral oophorectomy ≥ 6 weeks before the screening visit (with or without hysterectomy).

In response to EAG clarifications (question B11a), the company stated that they do not intend to restrict the population in the decision problem to only postmenopausal people because the physiological mechanism underlying VMS is consistent across menopause transition from perimenopause to postmenopause; however, the company recognises that the evidence from the DAYLIGHT and SKYLIGHT trials is only for postmenopausal people with VMS associated with menopause (see Section 2.3). The EAG further highlights that the natural history used in the model based on the SEE exercise is for postmenopausal people only. Therefore, the cost-effectiveness of fezolinetant is only assessed in a target population of postmenopausal people, with no evidence of clinical effectiveness or natural history provided for perimenopausal people.

item 3. The baseline distribution with a moderate to severe VMS frequency limit of greater than seven per day from the DAYLIGHT trial is unlikely to be appropriate for assessing the need for treatment in the NHS and is unlikely to match that seen in UK clinical practice for postmenopausal people.

4.2.4 Intervention and comparator

4.2.4.1 Summary of company's submission

The intervention assessed in the model is fezolinetant 45 mg once daily, while the comparator is no active treatment. The company considered no active treatment to be the only relevant comparator because the target population in the model is menopausal people with moderate to severe VMS for whom HRT is deemed unsuitable. In addition, UK clinical experts advising the company confirmed that no active treatment represents the sole relevant comparator to fezolinetant due to the lack of effective and safe treatment options available in UK clinical practice.

In response to EAG clarifications (question A1) on the potential consideration of non-hormonal pharmacological treatments listed in the NICE scope as relevant comparators, the company maintains its position that no active treatment represents the sole relevant comparator in the decision problem. However, the company presents an exploratory analysis comparing fezolinetant with paroxetine (SSRI).

Points for critique

The intervention is in line with the licensed dose for fezolinetant in moderate to severe VMS associated with menopause, and the dosing schedule for fezolinetant used in the model is consistent with the fezolinetant 45 mg arms of DAYLIGHT and SKYLIGHT 1 and 2. Therefore, the EAG has no concerns regarding the intervention dose used in the model.

As discussed in Section 2.2.5, the EAG is concerned that no active treatment may not be the sole relevant comparator to fezolinetant. For people with moderate to severe VMS deemed suitable for HRT, HRT would be the first-line treatment, and the company foresees that fezolinetant would be the only active treatment provided to those that have stopped HRT for any reason. For people with moderate to severe VMS deemed unsuitable for HRT, the company suggests that fezolinetant would be the only active treatment provided. Consequently, the effectiveness evidence informing the model is from DAYLIGHT in an HRT-unsuitable population and from pooled SKYLIGHT 1 and 2 in the subpopulation who were deemed HRT-unsuitable. Furthermore, the estimates of natural history elicited for the trajectory of moderate to severe VMS in the absence of treatment specifically considered a population unsuitable for HRT. However, other non-hormonal pharmacological treatments (such as anti-depressants, clonidine or gabapentiniods) are considered as comparators in

the NG23 guideline, the NICE scope and, according to EAG clinician advice, are still part of NHS clinical practice.

More specifically, in NG23,[20] HRT represents the mainstay treatment for the management of symptoms associated with menopause, including VMS, while it does not recommend the use of SSRIs, SNRIs or clonidine as first-line treatments for VMS alone. However, the guideline does not provide recommendations for treatment use specifically in people for whom HRT is not considered suitable. For this subpopulation, the NICE scope includes non-hormonal pharmacological treatments, including anti-depressants (such as SSRIs and SNRIs), clonidine, and anti-convulsants (such as gabapentin and pregabalin), as well as no pharmacological treatment and non-pharmacological treatments (such as CBT). In addition to no active treatment, the EAG considers the non-hormonal pharmacological treatments to be relevant comparators for the reasons outlined in Section 2.2.5.

The company presented an exploratory NMA comparing fezolinetant to paroxetine (SSRI) in response to EAG clarification questions. However, the company noted caution in the interpretation of the cost-effectiveness results using this analysis because (i) the NMA was only exploratory in nature (see Section 3.3 EAG critique); and (ii) the singular treatment effect estimate from the exploratory NMA had to be mapped to the four-by-four transition probability matrices (VMS frequency categories) at different 4-weekly cycles creating a significant source of uncertainty, since the model structure did not allow adequate use of a relative treatment effect from the NMA; and (iii) week 48-52 transition probabilities had to be extrapolated beyond week 52 given the lack of efficacy data beyond this point. The EAG agrees with the caution surrounding the company's exploratory comparison of fezolinetant with paroxetine because the company did not adapt the model with sufficient functionality to enable appropriate application of estimates of relative treatment effect applied to natural history as requested in EAG clarifications (question B3). Therefore, the EAG does not consider the results of the company's exploratory cost-effectiveness comparison of fezolinetant with paroxetine (Table 2 of company's response to EAG clarifications) to be reliable and is not considered further in this report.

> ***item 4. No active treatment is unlikely to be the sole relevant comparator to fezolinetant because menopausal people presenting with moderate to severe VMS and deemed HRT-unsuitable are more likely to be offered non-hormonal pharmacological treatments as opposed to 'nothing' in NHS clinical practice.***

4.2.5 Perspective, time horizon and discounting

4.2.5.1 Summary of company's submission

The analysis is conducted from the perspective of the NHS and Personal Social Services (PSS) in England and Wales over a 10-year time horizon. A 3.5% annual discount rate is used for both costs and health effects.

Points for critique

While the EAG considers the perspective and discounting to be in line with the NICE health technology evaluations manual [78], the time horizon of 10 years used in the company's base case analysis is not sufficient to capture the natural history of VMS, which is used for no treatment after week 12 in the model and applied in the post-discontinuation health states for fezolinetant. The EAG notes that using the median moderate to severe VMS cessation of 7.4 years used in the company's base case analysis, approximately 40% of postmenopausal people still experience moderate to severe VMS by year 10 in the model. Moreover, the clinical experts in the SEE estimated that moderate to severe VMS would cease (in all patients) by an aggregate median of 14.5 years (interquartile range 11.25, 30.5) or a mean of 20.5 years (standard deviation (SD) 13.6). Details about VMS duration is discussed in Sections 2.2.1 and 4.2.6.4. In addition, although the median treatment duration for fezolinetant is modelled to be approximately [redacted] years, a very small proportion of the cohort is still receiving fezolinetant by year 10. Thus, the EAG considers a 10-year time horizon insufficient to capture all the long-term effects of fezolinetant in the company's base case analysis.

4.2.6 Treatment effectiveness and extrapolation

The model includes four elements relating to treatment effectiveness and extrapolation of effects over the long-term, which are discussed below in turn: (i) short-term efficacy from clinical trials up to week 52; (ii) natural history in the absence of treatment; (iii) long-term treatment effect and treatment discontinuation rates; and (iv) VMS cessation and mortality.

4.2.6.1 Short-term efficacy from clinical trials up to week 52

The effectiveness evidence in the company's base case is sourced from DAYLIGHT up to week 24 for fezolinetant and up to week 12 for no active treatment (placebo arm of DAYLIGHT). The model applies four-weekly (per cycle) transition probability matrices for transitions between VMS frequency health states at each model cycle, derived using week 0–24 DAYLIGHT efficacy data for fezolinetant (Table 78, Appendix N of CS) and week 0–12 for placebo (Table 79, Appendix N of CS). At the end of week 12, people in the no active treatment arm revert to the natural history of VMS in the absence of treatment and follow the natural history transitions (Section 4.2.6.2 below).

The company states that using placebo data from DAYLIGHT up to week 12 only for the no active treatment arm is conservative because participants in the placebo arm experienced notable improvements in moderate to severe VMS frequency from baseline through to week 12 in DAYLIGHT, which is unlikely to be observed in clinical practice as people will not have additional follow-up or GP interactions associated with partaking in a clinical trial. Note that placebo data are available from DAYLIGHT up to week 24 and the company did not state explicitly why placebo data up to week 24 were not used in the base case analysis other than stating that the company deemed the

benefits of no active treatment to be likely overestimated in the model. In addition, the company states that UK clinical experts agreed that a placebo effect observed in menopause trials is unlikely to persist over the long-term and people receiving no active treatment would be expected to follow natural history more closely. At EAG clarifications (question B3) the EAG requested estimates of the relative treatment effect for fezolinetant compared to placebo from the trials for each VMS frequency health state in each 4-weekly cycle, up to week 24 from DAYLIGHT and up to week 12 from SKYLIGHT 1 and 2 (HRT-unsuitable), and to apply the relative effects to natural history in the absence of treatment (no active treatment arm). The company did not provide the requested estimates because they did not deem them the most appropriate source of efficacy in the model for two reasons: (i) applying relative treatment effects to the SEE estimates for natural history would lead to a loss of granularity in the modelling of fezolinetant efficacy compared with using observed trial data; and (ii) there is considerable uncertainty in the SEE estimates for the natural progression of VMS.

At the end of week 24 for fezolinetant, efficacy evidence from the pooled SKYLIGHT 1 and 2 trials with a longer follow-up of 52 weeks from the subpopulation of participants who were deemed HRT-unsuitable is used in the model. For weeks 24–52, four-weekly (per cycle) transition probability matrices are applied to each model cycle, derived using pooled week 24–52 SKYLIGHT 1 and SKYLIGHT 2 efficacy data (Table 78, Appendix N of CS).

Table 18 summarises the proportion of the cohort in each of the VMS frequency health states while on-treatment for fezolinetant and no active treatment (placebo) up to week 52 in the company's base case analysis.

Table 18 Proportion of cohort in the on-treatment health states up to week 52 in the model

	Fezolinetant				No active treatment (placebo)			
Time point	**0 ≤ VMS-F < 2**	**2 ≤ VMS-F < 7**	**7 ≤ VMS-F < 9**	**VMS-F ≥ 9**	**0 ≤ VMS-F < 2**	**2 ≤ VMS-F < 7**	**7 ≤ VMS-F < 9**	**VMS-F ≥ 9**
Baseline	0%	0%	42%	58%	0%	0%	42%	58%
4 weeks	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
8 weeks	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12 weeks	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
24 weeks*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
52 weeks	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

*The corresponding 24-week values based on the placebo arm of DAYLIGHT are: [redacted] for 0 ⩽ VMS-F < 2; [redacted] for 2 ⩽ VMS-F < 7; [redacted] for 7 ⩽ VMS-F < 9; and [redacted] for VMS-F ⩾ 9, but these are not used in the company's base case analysis because it is assumed that no active treatment reverts to the natural history of VMS after week 12.

Points for critique

Although, the company judged the SKYLIGHT 1 & 2 and DAYLIGHT trials to be at low risk of bias (CS, Table 16), the EAG identified two issues with missing data, which it believes place the results of all three trials at high risk of bias (due to missing data) with respect to continuous outcomes (this includes the VMS frequency and VMS severity outcomes). Details are provided in Section 3.2.1.1.

As discussed in Section 4.2.2.1, one of the EAG's primary concerns relating to the treatment effectiveness evidence used in the model is the focus on VMS frequency only, with no consideration given to the relative importance of a reduction in the severity of VMS compared to a reduction in the frequency of moderate to severe VMS. Figure 7 of the CS shows how the mean frequency of moderate to severe VMS per day changed from baseline to week 24 in DAYLIGHT, while Figure 8 of the CS shows how the mean weighted severity of VMS experienced per day (incorporating the incidence of mild, moderate and severe hot flushes per 24-hour period) changed from baseline to week 24 in DAYLIGHT. Both outcomes showed a statistically significant LS mean difference for fezolinetant versus placebo. This was further confirmed in the NMA by Morga et al., (2023) based on data from the two SKYLIGHT trials up to week 12 (double-blind randomised period) which estimated a relative mean difference of 2.78 (95% CrI, 1.93 to 3.62) for the comparison of fezolinetant 45 mg at 12 weeks versus placebo for change from baseline in daily frequency of moderate to severe VMS (see Section 3.5) and a mean difference of 0.26 (95% CrI, 0.15 to 0.37) for the comparison of fezolinetant 45 mg versus placebo for reduction in VMS severity in the fixed-effect model (however, note that no meaningful difference between fezolinetant 45 mg and placebo for reduction in VMS severity was estimated in the random effects model).

Another primary concern relates to how the treatment effectiveness estimates for the reduction in frequency of moderate to severe VMS per day is implemented in the model after week 12. The company has applied the absolute changes in 4-weekly cycles, observed in the separate fezolinetant and placebo arms of DAYLIGHT up to week 12, for the on-treatment VMS frequency health states, but after week 12 the placebo effects are lost in the no active treatment arm, where this arm reverts to the natural history of VMS, while fezolinetant continues to receive the absolute changes in 4-weekly cycles observed in the separate fezolinetant arm of DAYLIGHT up to week 24 and the pooled absolute changes observed in the fezolinetant arm of SKYLIGHT 1 and 2 from week 24 to week 52. The EAG's concern relates to the fact that the *relative* differences between fezolinetant and placebo from the trials are not implemented in the model after week 12. This effectively means that the placebo effect observed in the trials is lost after 12 weeks. The EAG considers it more appropriate to adjust the absolute changes observed in the fezolinetant arm for the placebo effect observed in the 24-week period of DAYLIGHT. The EAG agrees with the company that the notable improvements in moderate to severe VMS frequency observed from baseline through to week 24 in the placebo arm of

DAYLIGHT are unlikely to be observed in clinical practice outside the clinical trial setting, but the EAG notes that any improvement from additional follow-up or GP interactions associated with partaking in a clinical trial would be expected to be applied to both the fezolinetant and placebo arms of the trials, and should not affect how the treatments perform relative to one another. Therefore, the EAG considers it more appropriate to derive estimates of relative treatment effect for fezolinetant vs. placebo (i.e., adjusting the absolute changes observed in the fezolinetant arm for the placebo effect), with data available on the relative difference up to week 24 from DAYLIGHT and up to week 12 from SKYLIGHT 1 and 2, and then apply the estimates of relative treatment effect to the natural history of moderate to severe VMS in the absence of treatment. This approach would provide an estimate of the transitions between VMS frequency health states over time for fezolinetant, while using the estimates of natural history to represent the no active treatment comparator, rather than the transition probabilities from the placebo arm of the trials.

The implications of the company's base case approach are that the "placebo effect" is lost arbitrarily from week 12 in the model. This can be observed in the proportions of the cohort in the VMS frequency health states between week 12 and week 52 in the model (Table 19), where there is a notable abrupt shift for no active treatment in the proportion in the lower VMS frequency categories ($0 \leqslant$ VMS-F < 2 and $2 \leqslant$ VMS-F < 7) at week 12 to the higher VMS frequency categories ($7 \leqslant$ VMS-F < 9 and VMS-F $\geqslant 9$) after week 12. Consequently, this means that the cost-effectiveness of fezolinetant relative to no active treatment is highly sensitive to the estimates of natural history of moderate to severe VMS frequency that is assumed in the model. It also means that the base case analysis is favouring fezolinetant because the absolute changes observed in the fezolinetant arm of the trials is implemented without any deduction of the placebo effect from week 12 onwards, but the relative difference between fezolinetant and placebo appears fairly constant from about week 6 to week 24 in mean frequency of moderate to severe VMS per day from DAYLIGHT (see Figure 7, p74 of the CS).

The EAG considers the use of estimates of the relative treatment effect for fezolinetant compared to placebo from the trials for each VMS frequency health state in each 4-weekly cycle, up to week 24 from DAYLIGHT and up to week 12 from SKYLIGHT 1 and 2 (HRT-unsuitable), which were requested at EAG clarifications (question B3), and the application of the relative effects to natural history in the absence of treatment (no active treatment arm). The EAG considers the company's justification to not provide the requested estimates as weak because the SEE estimates are already used in the model for the natural progression of VMS from week 12 for no active treatment and for fezolinetant once participants come off-treatment, and the estimates of relative treatment effect would be based on the differences observed in the trial data between the fezolinetant and placebo arms at any given timepoint.

Table 19 Proportion of cohort in VMS frequency health states for fezolinetant and no active treatment between week 12 and week 52 in the model.

	Fezolinetant							
	On-treatment				Off-treatment			
Time point	0 ≤ VMS-F < 2	2 ≤ VMS-F < 7	7 ≤ VMS-F < 9	VMS-F ≥ 9	0 ≤ VMS-F < 2	2 ≤ VMS-F < 7	7 ≤ VMS-F < 9	VMS-F ≥ 9
12 weeks	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
24 weeks	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
52 weeks	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	No active treatment							
	On-treatment				Off-treatment			
Time point	0 ≤ VMS-F < 2	2 ≤ VMS-F < 7	7 ≤ VMS-F < 9	VMS-F ≥ 9	0 ≤ VMS-F < 2	2 ≤ VMS-F < 7	7 ≤ VMS-F < 9	VMS-F ≥ 9
12 weeks*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
24 weeks	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
52 weeks	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

* After week 12, the proportion of the cohort on-treatment (placebo) transitions to the natural history of moderate to severe VMS frequency in the absence of treatment, which results in an abrupt shift for no active treatment from the lower VMS frequency categories (VMS-F < 7) at week 12 to the higher VMS frequency categories (7 ⩽ VMS-F) after week 12.

The company applies efficacy data for fezolinetant from DAYLIGHT up to week 24 and pooled SKYLIGHT 1 and 2 from week 24 to week 52 for the subpopulation of participants who were deemed HRT-unsuitable. The EAG considers the subpopulation data from SKYLIGHT to be appropriate given the population of the decision problem. However, the company have not assessed the impact of switching from efficacy data from DAYLIGHT at week 24 to efficacy data from pooled SKYLIGHT 1 and 2 from week 24 to week 52 for fezolinetant. The EAG considers the data to be reasonably consistent at week 24 where the mean change from baseline to week 24 in mean frequency of moderate to severe VMS per day from DAYLIGHT was -8.15 (SD = 4.43) based on a sample size of 226, while the corresponding mean change from baseline to week 24 from pooled SKYLIGHT 1 and 2 was [redacted] (SD = [redacted]) based on a sample size of 231. Therefore, it would seem appropriate to also consider the efficacy data from the pooled SKYLIGHT trials in the first 24 weeks of the model, which would create consistency with the longer term data from week 24 to week 52 from the SKYLIGHT trials. The EAG is unclear why pooled DAYLIGHT and SKYLIGHT 1 and 2 (HRT-unsuitable) is not considered in the first 24 weeks of the model for fezolinetant, particularly given the company's base case approach where there is no adjustment for placebo effect incorporated in the efficacy estimates for fezolinetant after week 12.

item 5. The absolute changes from baseline in the frequency of moderate to severe VMS observed in the fezolinetant arm of the trials is not adjusted for the placebo effect from week 12 onwards in the model.

4.2.6.2 Natural history in the absence of treatment

Once menopausal people in the on-treatment VMS frequency health states transition to the off-treatment VMS frequency health states, they are modelled to follow natural history of moderate to severe VMS in the absence of treatment until VMS cessation, death, or the end of the model time horizon. The company did not identify any relevant studies on the natural history of VMS from the SLR conducted to identify clinical evidence on the comparative efficacy and safety of active prescription treatments for postmenopausal women with moderate to severe VMS. As a result, the company conducted a SEE exercise to derive estimates of natural history of moderate to severe VMS based on the judgements of six clinical experts from England and Wales with experience in the care of menopausal women, using Structured Expert Elicitation Resources (STEER) guidance.[79] The outputs of the SEE exercise were validated by three UK clinical experts to generate natural history inputs for the model base case, with the resulting estimates further validated by four patient experts. Details on the methodology and outputs from the company's SEE exercise are presented in a separate report accompanying the CS.[74]

In the SEE exercise, experts were asked to consider a hypothetical cohort of 1,000 postmenopausal women in the UK who were not currently receiving any treatment for moderate to severe VMS and who were not suitable for HRT and not experiencing any ongoing clinical benefit from previous treatments. Experts were asked to estimate the proportion (i.e., the number of women out of 1,000) experiencing different frequencies of daily moderate to severe VMS across three timepoints (after 1 year, after 3 years, and after 6 years), where the daily moderate to severe VMS frequencies were categorized into those that matched the health states of the model, so that a trajectory over time could be mapped out. Experts provided their individual-level judgements via an online survey, after which data from the group of experts were aggregated using unweighted linear opinion pooling. The bisection method was used to elicit the 1st percentile (lower plausible limit), 99th percentile (plausible upper limit), 50th percentile (median), 25th percentile (lower quartile) and 75th percentile (upper quartile) for each quantity. Table 20 summarises the natural history health state distribution proportions elicited at the three timepoints. The company noted that there was considerable uncertainty in the SEE estimates for the natural progression of VMS (see Figure 28, p156 of CS showing the corresponding uncertainty in the elicited proportions).

Table 20 SEE natural history health state distribution proportions (reproduced from Table 61 of CS)

Time point	0 ≤ VMS-F < 2	2 ≤ VMS-F < 7	7 ≤ VMS-F < 9	VMS-F ≥ 9
Year 1	16.91%	47.13%	17.16%	18.80%
Year 3	36.77%	40.34%	13.36%	9.53%
Year 6	48.12%	33.33%	12.39%	6.15%

The company states that further discussions with clinical experts suggested that the high uncertainty captured in the estimates is representative of the highly variable experience of menopausal people with VMS and reflects the counterfactual nature of trying to estimate the VMS frequency in a cohort that would rarely exist in reality – clinicians would not leave people experiencing moderate to severe VMS untreated. Furthermore, when the three elicited timepoints (1, 3 and 6 years) were applied to the baseline VMS frequency distribution from DAYLIGHT, a very rapid rate of VMS frequency decline was observed (see Figure 29, p157 of CS). Consequently, the company sought further input, with three UK clinical experts (including one expert involved in the SEE exercise) and four patient experts to validate the SEE estimates used in the model base case. The experts agreed that the timeline, especially the one-year timepoint, was unrealistically fast for the resolution of VMS based on typical clinical practice. Three alternative approaches were proposed to adjust the SEE results to better reflect UK clinical practice:

1. Only use year 6 proportions from the SEE;
2. Use the SEE year 1 estimates at year 2;
3. Only use year 3 and year 6 estimates.

The clinical experts acknowledged the challenges in deciding a preferred option and patient experts considered the year 6 estimates to provide a more realistic reflection of the natural progression of VMS; therefore, the company chose to only use year 6 proportions from the SEE in their base case analysis and presented scenario analyses based on: (i) using year 1 estimates at year 2; and (ii) using only use year 3 and year 6 estimates. The base case model calculates the distributions at each four-week cycle from baseline up to year 6 by assuming a linear change from the baseline distribution of DAYLIGHT to year 6 in each VMS frequency health state (Figure 30, p158 of CS).

Points for critique

The EAG has several concerns regarding the natural history in the absence of treatment. The EAG first notes that the company's SLR conducted to identify clinical evidence on the comparative efficacy and safety of active prescription treatments for postmenopausal women with moderate to severe VMS is unlikely to be sufficient to identify studies of natural history in the absence of treatment. At EAG clarifications, the EAG requested that the company perform a SLR specifically to identify studies of natural history of VMS, to report the findings and consider the use of alternative estimates of natural history in the model (question B12). The company responded by clarifying that a SLR was not specifically conducted to identify studies on the natural history of VMS, but given the heterogeneity and the individualised nature of VMS, a SLR to provide data on the specific outcome of moderate to severe VMS frequency per day for use in the model is unlikely to resolve uncertainty observed in the SEE estimates surrounding the natural progression of VMS.

The EAG's primary concern relates to the credibility of the results of the SEE for informing the natural history of moderate to severe VMS over time when a baseline VMS frequency distribution with a minimum average of seven moderate to severe VMS per day is used in the model. As noted previously, the EAG believes that the baseline distribution should be closer to the elicited values at one year for the natural history of postmenopausal women in the UK who do not receive any active treatment for moderate to severe VMS because such a rapid reduction in the frequency of moderate to severe VMS between baseline and one year in the absence of treatment would suggest a lack of need for treatments for moderate to severe VMS because the symptoms would resolve rapidly by themselves over time. The implausibility of the significant reduction in symptoms occurring within the first year was also confirmed by the clinical experts informing the company. The approach used by the company to overcome this concern involved discarding the SEE estimates at years 1 and 3 and only use the estimates elicited at year 6 in the base case analysis, with a linear change from the baseline distribution of DAYLIGHT up to year 6 applied. The EAG does not consider this approach satisfactory because it suggests that either the SEE estimates are unreliable or that the baseline VMS frequency distribution is not representative of moderate to severe VMS observed in UK clinical practice in the absence of treatment. The EAG notes that severity of VMS used in the SEE study was defined per EMA and FDA guidance (i.e., severe: sensation of heat with sweating causing cessation of activity; moderate: sensation of heat with sweating, but able to continue activity; mild: sensation of heat without sweating), but that the SEE exercise was not anchored on the baseline VMS frequency distribution used in the model with a minimum average of seven moderate to severe VMS per day. This was confirmed by the company in response to EAG clarifications (question B11b), where the company acknowledges that the SEE exercise was not anchored on the trial baseline distribution when elicited. Therefore, it is possible that the experts who provided proportions experiencing different frequencies of daily moderate to severe VMS at years 1, 3 and 6 had a very different starting distribution of moderate to severe VMS frequencies per day in mind.

The approach to elicitation was in line with the STEER guidance.[79] This method aims to minimise bias associated with eliciting subjective judgements. Several principles, suggested by the STEER framework, were adhered to in the company's elicitation exercise. Specifically: only substantive experts were recruited using pre-specified recruitment and selection criteria, the experts were trained prior to the task, including awareness on forms of bias, a SEE protocol was developed, and experts were frequently asked to reflect on and revise their answers as needed following each quantity elicited.

Several observations can be made regarding the methods employed in the SEE and the likely credibility of the resulting elicited estimates:

- Judgements were elicited individually from six experts using an online survey. The remote survey did not allow for interaction and clarification opportunities with the facilitator, which may lead to internal inconsistencies and low quality of responses. Out of six experts, five had judgements that were inconsistent where some or all the elicited percentiles were not in order of magnitude. In addition, quantities from one expert were deemed to have reflected patient variability rather than subjective uncertainty surrounding the median. One other expert answered in terms of proportions instead of choosing a number out of 1,000. These inconsistencies may suggest that the training provided to experts was insufficient and that a draft elicitation should have been piloted with a substantive expert before eliciting judgements from all six experts. To address the internal inconsistencies, a follow-up email allowed the experts to amend their answers when the qualitative rationale they provided to justify their individual judgments did not align with the estimate, and all provided new and consistent estimates. However, no consensus workshop was conducted where experts discuss, review and refine their judgments for quantities of interest where individual distributions have limited overlap; this approach may have been useful for experts to share the rationale and evidence considered with each other to produce a single distribution that summarises the range of expert opinions. Instead, judgments were aggregated through linear pooling with equal weighting given to all experts.
- The number of questions was high due to the need to elicit four different VMS frequency cut-offs at three different time points.
- The moderate to severe VMS frequency categories may have appeared arbitrary to the experts because the categorical assumptions introduced to the continuous data may not have been consistent with the experience of UK clinicians, despite being based on statistical analysis of trial data.
- The bisection method was used to determine uncertain judgements. The bisection approach uses a variable interval method; 1%, 99%, 50%, 25% and 75% percentiles. This is an alternative to the fixed interval methods. The STEER guidance concludes that either approach is appropriate for SEE to inform health care decision making.
- Individual judgements were pooled using unweighted linear opinion pooling. Linear opinion pooling is an approach recommended for quantities where independence can be assumed. The quantities elicited over time are unlikely to be independent. For each expert the proportion expressed at each of the three time points (year 1, 3 and 6) will be contingent (conditional) on the proportion expressed at other time points. It is not clear that this relationship between time points was made explicit in the questions posed to experts, i.e., phrasing the questions such that "conditional on the previous time point what percentage of postmenopausal women with moderate to severe VMS frequency of $\geqslant$ 9 is still maintained".

There was considerable uncertainty in the resulting SEE estimates and between-expert variation in the estimates. The company states that discussion with experts suggested that the high uncertainty captured in the estimates is representative of the highly variable experience of menopausal people with VMS and reflects the counterfactual nature of trying to estimate the VMS frequency in a cohort that would rarely exist in reality because clinicians would not leave people experiencing moderate to severe VMS untreated. However, clinicians also acknowledged that many menopausal people deemed unsuitable for HRT are not currently receiving active treatment due to the limitations of available options. The EAG notes that this suggests that no active treatment is not the relevant comparator in this target population who are unsuitable for HRT because such a population rarely exists in NHS practice. This further supports the view that no active treatment is unlikely to be the sole relevant comparator to fezolinetant because menopausal people presenting with moderate to severe VMS and unsuitable for HRT are more likely to be offered non-hormonal pharmacological treatments as opposed to 'nothing' in NHS clinical practice.

In the absence of alternative estimates of natural history of moderate to severe VMS, the model relies on the estimates from the SEE study. The EAG notes, however, that clinical opinion is unlikely to be a reliable source to elicit how VMS progresses over time in the absence of treatment because menopausal people are unlikely to report back to their GP on how their VMS are reducing over time, if there are no suitable treatment options available, i.e., in the population where HRT is unsuitable (HRT-contraindicated, -caution, -stoppers and -averse).

When presented with alternative options for using the SEE estimates in the model (i.e., use the year 1 estimates at year 2; only use the year 3 and year 6 estimates; or only use the year 6 estimates), the clinical experts acknowledged the challenges in deciding a preferred option; two experts selected to use the year 1 estimates at year 2, while the third expert selected to only use the year 6 estimates. The company states that patient experts unanimously chose the option to use only the year 6 estimates but also acknowledged the challenges in deciding a preferred option. This highlights the uncertainty associated with the use of the SEE estimates for informing the natural history of moderate to severe VMS in postmenopausal people. The company chose to only use year 6 estimates in the base case analysis and presented scenario analyses for the alternative options. The cost-effectiveness results for the scenario analyses demonstrate that the alternative options have a material impact on the ICER for fezolinetant relative to no active treatment (see Table 76, p182 of CS), which highlights how sensitive the cost-effectiveness results are to the natural history assumed in the model.

item 6. The credibility of the SEE estimates for informing the natural history of moderate to severe VMS over time in NHS clinical practice for postmenopausal people who are deemed HRT-unsuitable is unclear.

4.2.6.3 Long-term treatment effect and treatment discontinuation rates

From week 52 onwards, the effectiveness evidence for fezolinetant in the company's base case is sourced from the pooled SKYLIGHT 1 and 2 trials, where an average 4-weekly transition probability matrix was derived based on week 24 to week 52 efficacy data from pooled SKYLIGHT 1 and 2 (HRT-unsuitable subpopulation) and applied to each four-weekly cycle from week 52 onwards (Table 78, Appendix N of CS). For no active treatment, the placebo effect observed in DAYLIGHT is lost after week 12 and the cohort follows the natural history transitions informed by the SEE study, where a linear change from the baseline distribution of DAYLIGHT up to year 6 is assumed for each VMS frequency health state. For menopausal people who discontinue treatment with fezolinetant at any time point, they are modelled to transition to the off-treatment VMS frequency health states, in which case they revert to the natural history distribution of VMS and follow natural history transitions until VMS cessation, death, or the end of the model time horizon, whichever comes first.

All-cause treatment discontinuation rates are applied to the on-treatment VMS frequency health states to reflect the proportion of menopausal people returning to their natural history distribution in each model cycle. For fezolinetant, the per cycle probability of discontinuation of 2.43% was derived from week 0–24 DAYLIGHT data and applied in each model cycle up to week 24. From week 24 onwards, the per cycle probability of discontinuation of [redacted] was sourced from pooled SKYLIGHT 1 and 2 week 24–52 trial data (HRT-unsuitable). For no active treatment, the per cycle probability of discontinuation of 4.17% was sourced from DAYLIGHT week 0–24 placebo arm and applied in each model cycle up to week 12.

Points for critique

The EAG considers the approach used by the company to model long-term treatment effectiveness and discontinuation for fezolinetant to be reasonable, except for the concerns raised previously that the absolute changes observed in the fezolinetant arm of the trials are not adjusted for the placebo effect from week 12 onwards. For no active treatment, although the placebo effect is applied up to week 12 only, the per cycle probability of discontinuation in the first 12 weeks is sourced from week 0 to week 24 data of DAYLIGHT. The EAG would not expect to observe a material difference in the discontinuations rates for week 0-12 and week 0-24 for the placebo arm of the trial. The EAG notes that there was a greater, albeit marginal, proportion of participants who withdrew from treatment due to a TEAE in the placebo arm compared to the fezolinetant arm (6.2% versus 4.9%, respectively) of DAYLIGHT, while the proportion of participants where treatment withdrawal was attributed to the study drug was equal for both arms of the trial (3.1% each) – see Table 41, p122 of CS. This suggests

that the per cycle probability of discontinuation should be similar for fezolinetant and placebo. The difference of 1.74% (i.e., 4.17% - 2.43%) between arms in the first 12 weeks is unlikely to have a material impact on the cost-effectiveness results due to the short time period of 12 weeks for which it is applied; however, the EAG considers it more appropriate to use the same per cycle discontinuation rate between treatments in the first 12 weeks.

The implications of assuming long-term treatment effects for fezolinetant that are not adjusted for the placebo effect from week 12 onwards, while the placebo effect in the comparator of no active treatment is short lived and disappears after 12 weeks can be observed in the proportion of the cohort distributed across the 'on-treatment' and 'off-treatment' (natural history) VMS frequency health states over time for fezolinetant and no active treatment (Table 21), while the corresponding Markov trace showing the proportion of the cohort distributed across the modelled health states over time is presented in Figure 3 for fezolinetant and Figure 4 for no active treatment.

For fezolinetant, the absolute changes observed in the fezolinetant arm of DAYLIGHT up to week 24 and pooled SKYLIGHT 1 and 2 from week 24 to week 52 are applied, while the average week 24-52 data for the fezolinetant arm from pooled SKYLIGHT 1 and 2 are extrapolated from year 1 onwards for the proportion of the cohort on-treatment. The corresponding median duration on fezolinetant is 28.5 months (2.4 years) in the company's base case analysis. Table 21 shows that fezolinetant has a high proportion of the cohort in the low VMS frequency health states (VMS-F < 7) and a very low proportion in the high VMS frequency health states (7 ⩽ VMS-F) over time; this is a result of absolute treatment effects that shift the high baseline VMS frequency distribution (7 ⩽ VMS-F) to low VMS frequency health states while on-treatment, and after coming off treatment the natural history is closer to the SEE estimates for year 6, where 81% of the cohort is estimated to be in the low VMS frequency health states at year 6.

For no active treatment, the absolute changes observed in the placebo arm of DAYLIGHT up to week 12 are applied, followed by movement onto the 'off-treatment' (natural history) VMS frequency health states, which results in an abrupt shift from the lower VMS frequency categories at week 12 to higher VMS frequency categories after week 12. No active treatment is associated with a significantly larger proportion of the cohort in the high VMS frequency health states (7 ⩽ VMS-F) compared to fezolinetant. The EAG considers the projections shown in Figure 3 and Figure 4 to be:

- Optimistic for fezolinetant because the absolute changes observed in the fezolinetant arm are not adjusted for the placebo effect in the trials from week 12 onwards, where notable improvements were observed in moderate to severe VMS frequency from baseline through to week 24 in the placebo arm of DAYLIGHT;

- Pessimistic for no active treatment because the baseline starting distribution had a high VMS frequency distribution (7 ⩽ VMS-F) that did not align with the proportions elicited at one year from the SEE exercise for postmenopausal women with moderate to severe VMS in the absence of treatment, where 64% had low moderate to severe VMS frequency of less than seven per day. Although the year 6 SEE estimates were used in the company's base case analysis rather than the year 1 estimates, this results in a notable abrupt shift from low VMS frequency health states to high VMS frequency health states from week 12, which would not be expected in NHS clinical practice.

Table 21 Proportion of cohort distributed across the 'on-treatment' and 'off-treatment' (natural history) VMS frequency health states over time in the company's base case analysis

	Fezolinetant							
	On-treatment				**Off-treatment**			
Time point	**0 ≤ VMS-F < 2**	**2 ≤ VMS-F < 7**	**7 ≤ VMS-F < 9**	**VMS-F ≥ 9**	**0 ≤ VMS-F < 2**	**2 ≤ VMS-F < 7**	**7 ≤ VMS-F < 9**	**VMS-F ≥ 9**
Baseline	0%	0%	42%	58%	0%	0%	0%	0%
4 weeks	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
8 weeks	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12 weeks	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
24 weeks	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
1 year	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
5 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
6 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
10 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
15 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
20 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	No active treatment							
	On-treatment				**Off-treatment**			
Time point	**0 ≤ VMS-F < 2**	**2 ≤ VMS-F < 7**	**7 ≤ VMS-F < 9**	**VMS-F ≥ 9**	**0 ≤ VMS-F < 2**	**2 ≤ VMS-F < 7**	**7 ≤ VMS-F < 9**	**VMS-F ≥ 9**
Baseline	0%	0%	42%	58%	0%	0%	0%	0%
4 weeks	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
8 weeks	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12 weeks*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
24 weeks	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
1 year	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
5 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
6 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
10 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
15 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
20 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

* After week 12, the proportion of the cohort on-treatment (placebo) transitions to the 'off-treatment' (natural history of moderate to severe VMS frequency in the absence of treatment).

Figure 3 Markov trace showing the proportion of the cohort distributed across the modelled health states for fezolinetant over time in the company's base case analysis

Figure 4 Markov trace showing the proportion of the cohort distributed across the modelled health states for no active treatment over time in the company's base case analysis

4.2.6.4 VMS cessation and mortality

The model includes a VMS cessation health state to capture the natural cessation of VMS in menopausal people, which is independent of treatment. People in any VMS frequency health state (whether on-treatment or off-treatment) can experience cessation of VMS, after which VMS cannot return, and treatment is discontinued if people were previously receiving treatment. Transitions to the VMS cessation health state in each model cycle is based on a per cycle probability of VMS cessation, which was derived from the median duration of VMS of 7.4 years reported by Avis et al., 2015[5] and assumed to be constant over the modelled time horizon.

The model assumes that moderate to severe VMS does not have an impact on the risk of mortality. Therefore, the modelled cohort is assumed to have the same mortality rate as the general female population in the UK, which was derived from age-specific mortality rates for females from the National UK lifetables for 2019.[80]

Points for critique

The EAG considers the approach used by the company to capture the natural cessation of VMS in menopausal people to be appropriate in general; but the EAG highlights two concerns: (i) the median duration of moderate to severe VMS of 7.4 years is too high for the company's modelled base case population of postmenopausal people (starting age 51 years at the onset of menopause); and (ii) the assumption of a constant rate of VMS cessation over time.

The source used by the company to derive the median duration of VMS is based on the SWAN, which is a multiethnic observational study of the menopausal transition among 3302 women enrolled at seven sites in the US.[5] The study is one of the largest longitudinal studies spanning 17 years that reports the findings of the duration of VMS as women transition from pre-menopause or early perimenopause into late postmenopause. The study determined: (i) the total duration of frequent VMS, defined as ≥6 days in the previous two weeks, during menopause transition (the entire follow-up period from pre-menopause or early perimenopause and as far into postmenopause as each woman was observed), and (ii) quantified how long frequent VMS persists starting from the last menstrual period date and extending as far into postmenopause as each woman was observed; in addition to identifying risk factors for longer VMS duration. The baseline age for eligibility into the study was between 42 and 52 years and with the absence of hormone therapy, amongst other eligibility criteria. The EAG considers this source of evidence suitable for informing the median duration of VMS, but has several concerns about the median duration of 7.4 years used in the company's base case analysis and the assumption of a constant rate over time:

- The median total VMS duration of 7.4 years is for the entire follow-up period in the study of women pre-menopause or early perimenopause through to postmenopause, whereas the

company's modelled population is postmenopausal people with a mean starting age of 51 years at the onset of menopause because both the effectiveness evidence for fezolinetant is based only on postmenopause data and the natural history of VMS used in the model is elicited for postmenopause women. Therefore, the median duration of VMS is not reflecting the correct modelled population. The EAG notes that the study reports separately a median VMS duration of 3.4 years in the subset of women who were postmenopause at the onset of VMS (20.2% of the total cohort of 3302 women), which is more reflective of the company's modelled base case population of postmenopausal people.

- The reported median duration of VMS does not consider the severity of the symptoms, only the frequency defined as ≥ 6 days in the previous two weeks, whereas the company's modelled population is specific to the frequency of moderate to severe VMS.
- The study reported that younger age at first report of VMS was a risk factor for duration of VMS, including showing a difference in the risk of longer VMS duration for ages 50 to 54.9 compared to age ≥ 55 years, which suggests that the assumption of a constant rate over time as menopausal people age in the model may not be appropriate.

The company applied an exponential distribution to the median VMS duration of 7.4 years as shown in Figure 5. The EAG notes that this approach results in a probability of VMS cessation for menopausal people of approximately 60% by the end of the modelled time horizon of 10 years. Therefore, the EAG does not consider the median VMS duration of 7.4 years and the assumption of constant rate of VMS to be consistent with the modelled time horizon of 10 years because approximately 40% of postmenopausal people are still experiencing VMS by year 10 in the model. The EAG considers the median VMS duration of 3.4 years from Avis et al., 2015[5] based on the postmenopause subpopulation of SWAN to be more consistent with the modelled population of postmenopausal people with an average age of 51 years at menopause onset, which results in approximately 12% of people still experiencing VMS by year 10 (Figure 6). The EAG also considers that it may be appropriate to extend the modelled time horizon to 20 years to ensure that the model captures the natural cessation of VMS in most postmenopausal people; although the EAG does not expect an extension of the modelled time horizon from 10 to 20 years to have any material impact on the cost-effectiveness results because the majority of people have discontinued treatment with a duration of treatment of 2.4 years in the company's base case analysis.

Figure 5 Probability of VMS cessation used in the company's base case analysis based on a median VMS duration of 7.4 years with an exponential distribution

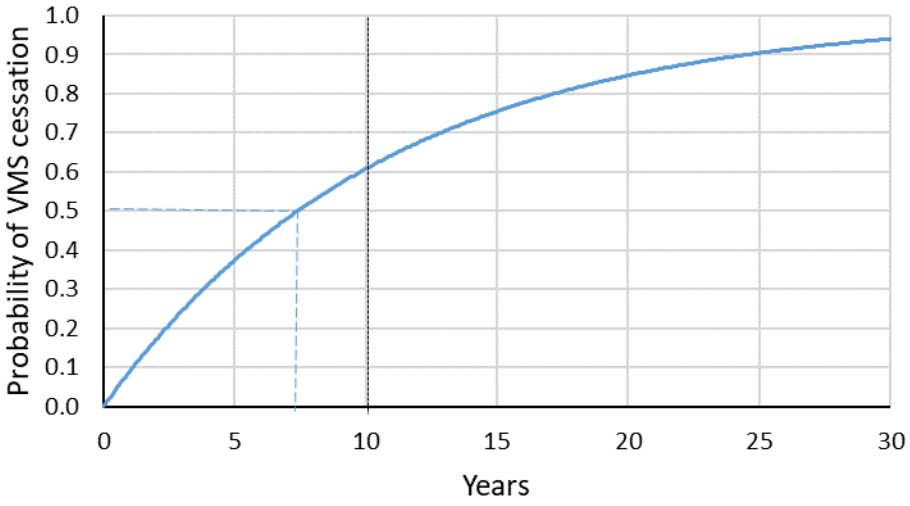
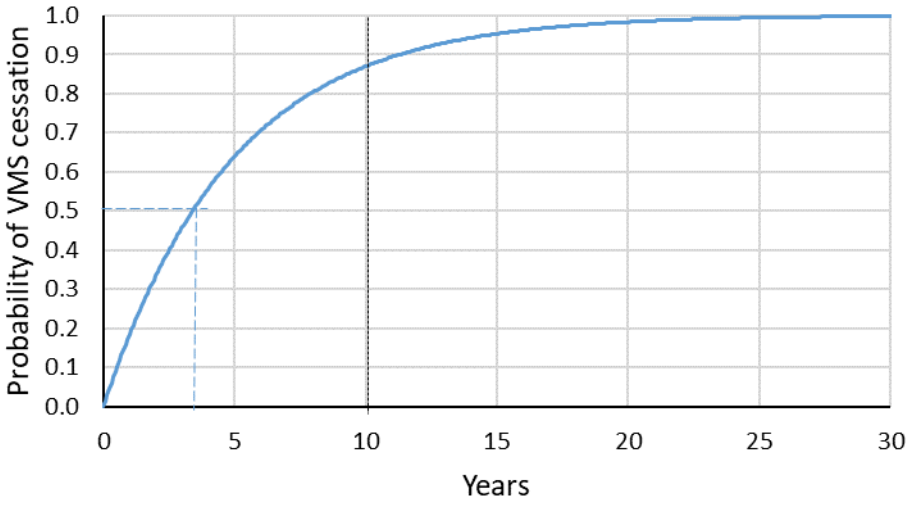


Figure 6 Probability of VMS cessation based on a median VMS duration of 3.4 years with an exponential distribution



All-cause mortality is estimated based on National Life Tables for the UK. The model assumes menopausal people do not have an increased risk of death due to moderate to severe VMS, which the EAG considers to be appropriate.

item 7. The median total VMS duration of 7.4 years that includes women pre-menopause or early perimenopause through to postmenopause does not align with the modelled population of postmenopausal people with a mean starting age of 51 years at the onset of menopause and the modelled time horizon of 10 years.

4.2.7 Adverse events

4.2.7.1 Summary of company submission

Due to the low incidence of TEAEs and serious adverse events (SAEs) of <5% observed for fezolinetant and placebo in the DAYLIGHT and SKYLIGHT trials, no adverse events were considered in the model.

Points for critique

The EAG considers the approach used by the company to be reasonable as the low incidence of events would have minimal impact on the cost-effectiveness results. However, the EAG has expressed concerns about an increased risk of neoplasms (benign, malignant and unspecified [including cysts and polyps]) and a rare occurrence of serious liver injury with the use of fezolinetant for VMS (see Section 3.2.3).

4.2.8 Health-related quality of life

4.2.8.1 Summary of company submission

Health state utility values are applied to time spent in health states in the model, in order to calculate QALYs that reflect the improvement in HRQoL associated with treatment. The company did not undertake a separate literature review to identify studies assessing the HRQoL of menopausal people but instead referred to the SLR used to identify published evidence on previous economic evaluations of fezolinetant. Utility values from the one identified previous economic evaluation[68]were not used in the cost-effectiveness model because EQ-5D-5L data were directly available from the DAYLIGHT and SKYLIGHT 1 and 2 trials.

The data from the DAYLIGHT trial were used to estimate the base case utility values in the model. Importantly, in the DAYLIGHT trial, the utility values for the placebo arm were higher than the utility values for the fezolinetant arm across all VMS frequency health states (see Table 22). The pooled treatment arm mean EQ-5D-5L scores were converted into utility values for each model health state. As described in section 4.2.2.1, these utility values were also used to define the threshold cut-offs for each VMS frequency health state. Due to the inclusion criteria of the DAYLIGHT trial, only people with a mean daily frequency of VMS of at least seven episodes per day were included at baseline, while the utility values for the health states of $0 \leqslant$ VMS frequency < 2 and $2 \leqslant$ VMS frequency < 7 were based on the post-baseline measures.

Table 22 Treatment specific health-state utilities DAYLIGHT (reproduced from Table 47 in the company's responses to points for clarifications, p56)

Health state	HRT-unsuitable			
	Fezolinetant		Placebo	
	Mean	**SE**	**Mean**	**SE**
Cessation of VMS	[redacted]	[redacted]	[redacted]	[redacted]
0 ≤ VMS Frequency < 2	[redacted]	[redacted]	[redacted]	[redacted]
2 ≤ VMS Frequency < 7	[redacted]	[redacted]	[redacted]	[redacted]
7 ≤ VMS Frequency < 9	[redacted]	[redacted]	[redacted]	[redacted]
VMS Frequency ≥ 9	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: HRT: hormone replacement therapy; SE: standard error; VMS: vasomotor symptoms.

For each VMS frequency health state, the utility values were estimated using a GEE linear mixed effects regression model. In the regression model, the dependent variable was the utility index score derived from EQ-5D-5L data mapped to EQ-5D-3L using the algorithm provided by Hernandez Alava et al., (2023), and the independent variable was the VMS frequency health states, which were defined based on the EQ-5D data and clinical expert opinion. The statistical fit of the GEE utilities was assessed using Quasi Information Criterion (QIC), which was provided for the final model (Table 46 and 49, company response to EAG clarifications) but not for the other models that were considered.

Utility values for VMS cessation were calculated from the average EQ-5D-5L data of participants who reported a VMS frequency of 0 during any visit, including at baseline. No imputation was performed for missing EQ-5D data. Patients who did not have an evaluation on a scheduled visit were excluded from the analysis for the timepoint of that visit.

Importantly, the company then adjusted the EQ-5D utility values from DAYLIGHT to reflect feedback from UK clinical experts. The company justified this approach by referring to a previous study that have indicated that generic measures, such as EQ-5D, are insufficiently sensitive to the improvements in symptoms among menopausal women.[81] This study included 3,721 women aged 50–69 treated with HRT and reported statistically significantly alleviated VMS, improved sexual functioning, and reduced sleep problems but no statistically significant benefits in overall HRQoL measured by the EQ-5D. Consequently, the company assumed that the results of DAYLIGHT showing no significant improvement in EQ-5D/VAS scores were not surprising and are explained by the challenges associated with trying to capture HRQoL benefits from treatments for menopause. As a result, the company adjusted the EQ-5D utility values from DAYLIGHT with the opinion from UK clinical experts. More specifically, one clinical expert suggested that the difference in utility between VMS cessation and 0 ⩽ VMS Frequency < 2 would be greater in UK clinical practice than that seen in DAYLIGHT. This expert estimated a utility value of 0.810 for the 0 ⩽ VMS Frequency < 2 health state. For the higher-frequency health states (7 ⩽ VMS Frequency < 9 and VMS Frequency ⩾ 9), the

company adjusted the utility values from DAYLIGHT by assuming a 5% reduction, which was justified based on clinical expert feedback suggesting that the trial values seemed too optimistic compared to what they have observed in their clinical experience. Table 23 summaries the health state utility values used in the model.

The company provided a scenario analysis with utility values based on the SKYLIGHT trials.

Table 23 Health state utility values used in the base case and scenario analysis (reproduced from Table 64 of CS, p162 and the company cost-effectiveness model updated in response to the EAG clarification questions)

Health state	UK clinician adjusted DAYLIGHT data (base case)		DAYLIGHT (scenario)		Pooled DAYLIGHT & SKYLIGHT 1 & 2 (HRT-unsuitable; (scenario)		Pooled SKYLIGHT 1 & 2 (HRT-unsuitable; scenario)	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Cessation of VMS	0.843	0.012	0.843	0.012	0.852	0.008	[redacted]	[redacted]
0 ≤ VMS Frequency < 2	0.810	0.010	0.833	0.010	0.841	0.007	[redacted]	[redacted]
2 ≤ VMS Frequency < 7	0.793	0.010	0.793	0.010	0.810	0.006	[redacted]	[redacted]
7 ≤ VMS Frequency < 9	0.746	0.011	0.785	0.011	0.793	0.007	[redacted]	[redacted]
VMS Frequency ≥ 9	0.710	0.012	0.747	0.012	0.773	0.007	[redacted]	[redacted]

Estimates in bold denote where estimates have been adjusted based on UK clinician estimates compared with the mean values in DAYLIGHT and DAYLIGHT (excluding HRT-averse population). Clinicians suggested the utilities for the 7 ≤ VMS-F < 9 and VMS-F ≥ 9 health states would be lower; these were therefore adjusted assuming a 5% reduction in utility compared to the utility reported in the DAYLIGHT trial (excluding HRT-averse population) for the corresponding VMS frequency health state. Utilities apply to both treatment arms of the model. Abbreviations: SE: standard error; UK: United Kingdom; VMS: vasomotor symptoms.

In response to EAG clarifications (question B8), the company stated that the treatment effect on utility values was different in the DAYLIGHT and SKYLIGHT 1 and 2 trials. Utility values for all health states, which are provided in Tables 47 and 48 in the company's response to EAG clarifications, are notably higher from the SKYLIGHT 1 and 2 trials compared to DAYLIGHT ([redacted] vs. [redacted] for VMS frequency 0 to ⩾9 in SKYLIGHT 1 and 2 and DAYLIGHT, respectively), and were relatively close to the age-adjusted utility values for women in the general population (0.86-0.82 for ages 51-61). The company performed additional analyses in which the treatment-specific utilities from DAYLIGHT were adjusted for baseline utility and an interaction term between baseline utility and treatment. The company stated in their response to clarifications that based on these additional analyses, the discrepancy between the utility values in DAYLIGHT and SKYLIGHT trials is in part due to the differences in the baseline utility of patients in each treatment arm. When comparing the utility values for each health state, the difference between the utility values for health state 2 ⩽ VMS Frequency < 7 and health state 7 ⩽ VMS Frequency < 9 was relatively minor in

DAYLIGHT. In SKYLIGHT 1 and 2, the difference between the utility values for health state 7 ⩽ VMS Frequency < 9 and health state ⩾ 9 was minor.

Points for critique

The EAG are concerned about several assumptions the company made in their estimation of the utility values used in the model. First, although the company stated that they applied GEE linear regression using the DAYLIGHT data to estimate the utility values, scarce information was provided on how the final model was selected. As described in Section 4.2.2.1, the appendix of the SEE report provides some information on the utility values and on how they were used to define the health states in the model. It also included information on several GEE models which the company explored. The EAG considers the company's justification for the choice of final model to be weak (see section Model Structure for the description and EAG's critique). In response to EAG clarifications (question B7), the company provided QIC for the final model, but it was not provided for the other models that were explored. Consequently, given the scarce information provided, the EAG could not assess the appropriateness of the methods applied.

Second, the EAG are concerned that the utility values from the DAYLIGHT trial for all health states of the model were higher in the placebo arm compared to the fezolinetant arm (see Table 21). Although the company used a mean estimate of utility values across the fezolinetant and placebo arms, the improvement in quality of life associated with fezolinetant is unclear. Furthermore, these values were used to define the health states used in the model. In contrast, the utility values from the SKYLIGHT 1 and 2 trials were higher in the fezolinetant arm than in the placebo arm. However, the EAG notes that the utility data from SKYLIGHT 1 and 2 were available only up to week 12, after which placebo participants could cross-over to the fezolinetant arm. This makes it more difficult for the EAG to interpret the utility values from SKYLIGHT 1 and 2 relative to DAYLIGHT because the utility values per specific timepoint are not available. Furthermore, the EAG notes that all utility values in SKYLIGHT 1 and 2 were relatively close to the age-adjusted utility values for women in the general population (Section 3.2.2.5).

Finally, the EAG considers that the adjustment of the utility values with clinical expert opinion is not an appropriate approach. The company stated that the EQ-5D measure is not sufficient to capture changes in quality of life due to reduced frequency and/or severity of VMS and used this as a justification for the finding that there were no statistically significant differences between the fezolinetant and placebo arms, including the fact that the values in the placebo arm were higher than in the fezolinetant arm for all health states. Consequently, the company decided to adjust these values based on the opinion of two clinical experts. This adjustment was validated by one clinical expert and 4 patient experts. The EAG considers this method to adjust utility values to be highly uncertain and

not appropriate. The EAG refers to the NICE health technology evaluations manual 2022, which provides recommendations for circumstances when the EQ-5D is considered to not be an appropriate measure. Specifically, the manual requires a case to be made on why the EQ-5D is inappropriate including the need to provide qualitative empirical evidence on the lack of content validity of the EQ-5D, showing that key dimensions of health are missing. A flowchart (Figure 4.1 of the NICE manual) details the hierarchy of preferred HRQoL methods. Further detailed recommendations are provided in Rowen, et al. 2020.[82] The adjustment method applied in the CS does not follow these requirements outlined in the NICE manual and is considered arbitrary.

item 8. The modelled health state utility values are highly uncertain.

4.2.9 Resource use and costs

4.2.9.1 Summary of company's submission

The company conducted a SLR to identify any relevant studies reporting healthcare resource use for VMS associated with menopause. Full details of the SLR are reported in Appendix I of CS. In total, 24 publications reporting on 23 unique studies met the inclusion criteria of Table 65, Appendix I. However, due to the health state definitions used in the model and issues with the generalisability of the included studies to UK clinical practice, none of these studies were considered by the company to be suitable to inform the base case analysis. As a result, the company used UK clinical expert opinion to inform health care resource use rates by health state in the model.

The CS includes costs related to (i) treatment acquisition and (ii) health-state resource use. No administration costs were considered for fezolinetant as it is self-administered orally by patients. Resource use associated with adverse events were not included in the model. Unit costs were informed by national public sources, inflated to 2021/22 prices where appropriate and discounted at an annual rate of 3.5%. Table 24 provides a summary of the costs used in the company's base case analysis.

Table 24 Costs used in the company's base case analysis

Item	Model input	Source
Drug acquisition costs per year		
Fezolinetant	£ 584	The dosage of fezolinetant is 45 mg orally per day. List price per pack of fezolinetant 45 mg is £44.80 (28 tablets per pack, per 4-weekly model cycle). [83]
No active treatment	£ 0	No drug acquisition costs were considered for no active treatment.
Health state costs per year		
Cessation of VMS	£55.00	The model assumed resource use rates to be health-state dependent and independent of treatment received. Healthcare resource use (HCRU)
0 ⩽ VMS Frequency < 2	£55.00	

2 ⩽ VMS Frequency < 7	£110.00	inputs in the base case were derived using estimates from UK clinical experts [84] and validated by patient experts [85], which included the use of GP visits, emergency room visits, and specialist visits as seen in UK clinical practice. The details of HCRU by health state (including physician first and follow-up visits, ER visit, and specialist first and follow-up visits) are summarised in Table 66 of CS. The unit cost associated with the HCRU by health state were obtained using the Personal Social Services Research Unit (PSSRU) Costs of Health and Social Care 2022 [86] and the NHS Schedule of Costs 2021/22.[87] The details of healthcare resource unit costs were summarised in the Table 67, CS.
7 ⩽ VMS Frequency < 9	£323.47	
VMS Frequency ⩾ 9	£538.66	

The model assumed resource use rates to be health-state dependent and independent of treatment received. Inputs below in the base case were derived using estimates from UK clinical experts [84] and validated by patient experts,[85] which included the use of GP visits, emergency room visits and specialist visits as seen in UK clinical practice (Table 25).

Table 25 Healthcare resource use by health state per year (reproduce from Table 66, CS)

Procedures and Monitoring Services	Cessation of VMS[a]	0 ≤ VMS Frequency < 2	2 ≤ VMS Frequency < 7	7 ≤ VMS Frequency < 9	VMS Frequency ≥ 9
GP visits (first visit)	1.00	1.00	1.00	1.00	1.00
GP visits (follow-up visit)	0.00	0.00	1.00	2.00	2.00
ER visits	0.00	0.00	0.00	1.00	1.00
Specialist visit (first visit)	0.00	0.00	0.00	0.00	1.00
Specialist visit (follow-up visit)	0.00	0.00	0.00	0.00	0.00

Abbreviations: VMS: vasomotor symptoms.
Source: UK clinician estimates.[84]

Points for critique

In general, the EAG considers the approach used by the company to derive resource use and costs to be appropriate. The drug acquisition and administration costs are aligned with the licensed posology for fezolinetant in moderate to severe VMS.

For the health state costs, the higher cost associated with the higher frequency health state of VMS ⩾ 9 per day is due to an additional specialist first visit, which incurs a unit cost of £215.19 (higher than an ER visit of £158.47). The company also provided a scenario analysis using the healthcare resource use sourced from Whiteley et al. 2013[16]. This source was based on a US study due to limited UK studies reporting on resource use rates for VMS. In this source, there are no emergency room visits and specialist visits included. The EAG considers the healthcare resource use in the company's base case to be reasonable.

The company did not include any adverse events costs in the model because of the low incidence of TEAEs and SAEs observed for fezolinetant and placebo in the DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 trials. However, a recent Lancet publication[88] has drawn attention to the significantly higher risk of neoplasm in the fezolinetant groups than the placebo group. EAG asked the explanation and evidence of the risk at the clarification stage. The company stated in their response to clarifications that the data assessed, including nonclinical and clinical safety data from the fezolinetant development programme, the structural properties of fezolinetant, literature on key characteristics of carcinogens and epidemiological data, and post-marketing data, do not indicate an increased risk of neoplasms. Further details are discussed in Section 3.2.3.

In the response to clarifications, the company provided costs of a hepatic laboratory test for people taking fezolinetant (£1.61). This test cost is applied for patients receiving fezolinetant in cycle 0 (baseline), 1, 2, 3, 6 and 9, in line with the FDA recommendations for liver function tests in the first year of treatment. The discussions with EMA and MHRA around additional testing in patients receiving fezolinetant are currently ongoing. Consequently, costs of a hepatic laboratory test should not be included in the base case analysis.

5 COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results

5.1.1 Summary of company's submission

All analyses presented in the CS use the list price for fezolinetant 45mg. A summary of the inputs and variables used in the company's base case analysis is presented in Table 69 of the CS, while the assumptions used in the model are summarised in Table 70 of the CS.

Table 26 shows the company's base case deterministic and probabilistic cost-effectiveness results. The probabilistic ICER for fezolinetant relative to no treatment is £10,355/QALY, and the deterministic ICER is £10,364/QALY. Table 27 presents the incremental net health benefit (INHB) under deterministic and probabilistic base case analysis. Fezolinetant is shown to have a positive INHB compared to no treatment.

Table 26: Deterministic and probabilistic base-case results (reproduced from Table 74 and Table 71, CS)

Deterministic							
Technologies	**Total**			**Incremental**			**ICER (£/QALY)**
	Costs (£)	**LYG**	**QALYs**	**Costs (£)**	**LYG**	**QALYs**	
Fezolinetant 45 mg	£2,803.18	8.449	6.684	£1,199.49	0.000	0.116	£10,364
No Treatment	£1,603.69	8.449	6.568	-	-	-	-
Probabilistic							
Technologies	**Total**			**Incremental**			**ICER (£/QALY)**
	Costs (£)	**LYG**	**QALYs**	**Costs (£)**	**LYG**	**QALYs**	
Fezolinetant 45 mg	£2,796.94	8.434	6.659	£1,194.98	0.000	0.115	£10,355
No Treatment	£1,601.95	8.434	6.543	-	-	-	-

Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life years gained; QALYs: quality-adjusted life years.

Table 27: Deterministic and probabilistic incremental net health benefit (reproduced from Table 75 and Table 72, CS)

Deterministic						
Technologies	**Total**		**Incremental**		**INHB**	
	Costs (£)	**QALYs**	**Costs (£)**	**QALYs**	**At £20,000**	**At £30,000**
Fezolinetant	£2,803.18	6.684	£1,199.49	0.116	0.06	0.08
No Treatment	£1,603.69	6.568	-	-	-	-
Probabilistic						
Technologies	**Total**		**Incremental**		**INHB**	
	Costs (£)	**QALYs**	**Costs (£)**	**QALYs**	**At £20,000**	**At £30,000**
Fezolinetant	£2,796.94	6.659	£1,194.98	0.115	0.06	0.08
No Treatment	£1,601.95	6.543	-	-	-	-

Abbreviations: QALYs: quality-adjusted life years; INHB: incremental net health benefit.

In the company response to EAG clarifications, an additional comparator (SSRI – paroxetine 7.5 mg) was included in a scenario using evidence from an exploratory NMA.[30] The results of the NMA were used to generate transition probabilities for paroxetine and the methodology of which is summarised in Section 3.3. The cost-effectiveness results of the pairwise comparisons between fezolinetant and SSRIs (paroxetine) are presented in Table 28. The ICER is £18,554/QALY. Table 29 presents the incremental net health benefit (INHB).

Table 28 Summary of cost-effectiveness results for fezolinetant versus paroxetine (SSRI)

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
Fezolinetant	£2,803.18	8.449	6.684	£1,366.40	0.000	0.074	£18,554
paroxetine (SSRI)	£1,436.79	8.449	6.610	-	-	-	-

Abbreviations: ICER: incremental cost-effectiveness ratio; LYG: life years gained; QALYs: quality-adjusted life years; SSRI: selective serotonin reuptake inhibitor.

Table 29 Incremental net health benefit for fezolinetant versus paroxetine (SSRI)

Technologies	Total		Incremental		INHB	
	Costs (£)	QALYs	Costs (£)	QALYs	At £20,000	At £30,000
Fezolinetant	£2,803.18	6.684	£1,366.40	0.074	0.01	0.03
paroxetine (SSRI)	£1,436.79	6.610	-	-	-	-

Abbreviations: QALYs: quality-adjusted life years; INHB: incremental net health benefit; SSRI: selective serotonin reuptake inhibitor.

Points for critique

To aid understanding of the key drivers of the cost-effectiveness results, Table 30, Table 31 and Table 32 below provide a summary of the disaggregated costs and QALYs respectively, based on the company's deterministic cost-effectiveness results.

Table 30: Summary of QALY gain by health state (reproduced from Table 68, Appendix J, CS)

Health state	QALY fezolinetant	QALY no active treatment	Increment	Absolute increment	% absolute increment
0 ≤ VMS Frequency < 2	2.21	1.35	0.86	0.86	38%
2 ≤ VMS Frequency < 7	1.19	0.88	0.31	0.31	14%
7 ≤ VMS Frequency < 9	0.50	0.99	−0.48	0.48	22%
VMS Frequency ≥ 9	0.48	1.05	−0.57	0.57	26%
Natural cessation of VMS	2.30	2.30	0.00	0.00	0%
Total	6.68	6.57	0.12	2.23	100%

Abbreviations: QALY, quality-adjusted life year; VMS: vasomotor symptoms.

Table 31: Summary of costs by health state (reproduced from Table 69, Appendix J, CS)

Health state	Cost fezolinetant	Cost no active treatment	Increment	Absolute increment	% absolute increment
0 ≤ VMS Frequency < 2	£1,171.18	£93.75	£1,077.43	£1,077.43	51%
2 ≤ VMS Frequency < 7	£699.55	£123.90	£575.65	£575.65	27%
7 ≤ VMS Frequency < 9	£309.70	£431.72	−£122.03	£122.03	6%
VMS Frequency ≥ 9	£468.90	£800.46	−£331.56	£331.56	16%
Natural cessation of VMS	£153.85	£153.85	£0.00	£0.00	0%
Total	£2,803.18	£1,603.69	£1,199.49	£2,106.67	100%

Abbreviations: VMS: vasomotor symptoms.

Table 32: Summary of predicted costs by cost category (reproduced from Table 70, Appendix J, CS)

Cost category	Cost fezolinetant	Cost no active treatment	Increment	Absolute increment	% absolute increment
Drug Costs	£1,743.74	£0.00	£1,743.74	£1,743.74	76%
Healthcare resource costs	£1,059.45	£1,603.69	−£544.25	£544.25	24%
Total	£2,803.18	£1,603.69	£1,199.49	£2,287.98	100%

Abbreviations: QALY, quality-adjusted life year.

The additional costs of fezolinetant compared to no active treatment are driven by the drug acquisition costs of fezolinetant, with some percentage of this cost offset by lower healthcare resource costs associated with less people in the (more costly) higher VMS frequency states of 7 ⩽ VMS Frequency < 9 and VMS Frequency ⩾ 9. The QALY gain for fezolinetant is driven by the gains in HRQoL associated with more people in the (better HRQoL) lower VMS frequency states 0 ⩽ VMS Frequency < 2 and 2 ⩽ VMS Frequency < 7 over time.

5.2 Company's sensitivity analyses

5.2.1 Summary of company's submission

The company conducted univariate deterministic sensitivity analysis (DSA) on key model inputs: natural VMS cessation, resource use, treatment discontinuation, utilities and patient characteristics. The company plotted the ten most influential parameters on a tornado plot (see Figure 34, CS). The DSA results showed that the INHB was most sensitive to the health state utility values for the VMS frequency health states. None of the ten variables crossed the point of indifference (i.e., when INHB is zero) for either their upper bound or lower bound values or had any substantial impact on the cost-effectiveness estimates.

The company also reported 18 scenario analyses, and the results from these analyses are presented in Table 76, CS. The results of the scenario analyses show that in all scenarios, the ICERs are below a £30,000/QALY cost-effectiveness threshold.

No subgroup analyses were conducted by the company.

5.3 Model validation and face validity check

5.3.1 Summary of company submission

The company undertook internal and external validation of the model. The internal validation of the model was conducted by independent health economist prior to the submission and made use of a

checklist[89] incorporating key elements of the TECH-VER checklist.[90] This is described in detail in Section B.3.14.1 of CS.

For external validation, expert input from health economists and clinicians was sought during the development of the cost-effectiveness model, and subsequently validated by patient experts.[84]

Points for critique

The EAG considers that the company's validation procedure was appropriate. The EAG reviewed the company's model in detail. In the probabilistic sensitivity analysis (PSA), the company set a maximum limit of 1000 simulations. The EAG changed the limit to assess convergence of the simulations; however, the model could not generate simulations of costs and QALYs for the health states of VMS frequency 0 to 2 and natural cessation of VMS in the fezolinetant arm above 1000 simulations. The EAG notes that the PSA results from 1000 simulations are similar to the deterministic results (see Table 26); the EAG does not consider this model error to have an impact on the assessment of cost-effectiveness.

6 EXTERNAL ASSESSMENT GROUP'S ADDITIONAL ANALYSES

6.1 Exploratory and sensitivity analyses undertaken by the EAG

A summary of the key issues identified and critiqued in Section 4, along with the scenario where the EAG addresses each issue in its additional analyses, is shown in Table 33. The EAG identified a number of limitations and areas of uncertainty in the company's cost-effectiveness analysis. Where possible, the EAG explored alternative assumptions to the company's base-case analysis, focusing on those issues that are expected to have the most impact (EAG Scenarios 1-5). A description of the EAG scenario analyses is presented in Section 6.1.1, while the impact on the cost-effectiveness results is presented in Section 6.2.

In Section 6.3, the EAG presents a restricted base case, with caveats. The EAG is concerned about several key structural uncertainties that have not been adequately addressed in the company's cost-effectiveness analysis. The effect of making changes simultaneously on elements that are considered to form part of the EAG's restricted base case are presented in Section 6.3 and the impact of alternative scenarios explored.

Table 33 Summary of the main issues identified by the EAG in Section 4 and EAG scenarios

Critique item and description The EAG considers that:		Dealt with in the		Area of remaining uncertainty	Significant impact on ICER
		EAG Scenarios	EAG restricted base-case		
1	***Frequency of moderate to severe VMS is not commonly used to define severity of VMS in NHS clinical practice.***	No	No	Yes	Unclear
2	***Arbitrary cut-off thresholds are used to define moderate to severe VMS frequency health states in the model.***	No	No	Yes	Unclear
3	***The baseline distribution with a moderate to severe VMS frequency limit of greater than seven per day from the DAYLIGHT trial is unlikely to be appropriate for assessing the need for treatment in the NHS and is unlikely to match that seen in UK clinical practice for postmenopausal people.***	Yes Sc. 1a-1b	Yes	Yes	Yes
4	***No active treatment is unlikely to be the sole relevant comparator to fezolinetant because menopausal people presenting with moderate to severe VMS and deemed HRT-unsuitable are more likely to be offered non-hormonal pharmacological treatments as opposed to 'nothing' in NHS clinical practice.***	No	No	Yes	Yes
5	***The absolute changes from baseline in the frequency of moderate to severe VMS observed in the fezolinetant arm of the trials is not adjusted for the placebo effect from week 12 onwards in the model.***	Yes Sc. 2a-2d	Partially	Yes	Yes
6	***The credibility of the SEE estimates for informing the natural history of moderate to severe VMS over time in NHS clinical practice for postmenopausal people who are deemed HRT-unsuitable is unclear.***	Yes Sc. 3a-3d	Yes	Yes	Yes
7	***The median total VMS duration of 7.4 years that includes women pre-menopause or early perimenopause through to postmenopause does not align with the modelled population of postmenopausal people with a mean starting age of 51 years at the onset of menopause and the modelled time horizon of 10 years.***	Yes, Sc. 4a-4b	Yes	No	No
8	***The modelled health state utility values are highly uncertain.***	Yes Sc. 5a-5b	Yes	Yes	Yes

6.1.1 Issues explored by the EAG in additional analyses

6.1.1.1 Scenarios 1: Baseline distribution

As discussed in section 4.2.3.1, the baseline distribution is assumed to be of a minimum average of seven moderate to severe events of VMS per day in the absence of treatment, which the EAG does not consider be an appropriate baseline to represent menopausal people likely to receive fezolinetant in NHS primary care. The EAG believes that the baseline distribution should be closer to the elicited values at one year for the natural history of postmenopause women in the UK because a significant reduction in VMS frequency would not be expected after one year in the absence of treatment. The proportions elicited at one year for postmenopause women with moderate to severe VMS and not receiving treatment were: 17% for 0 ≤ VMS Frequency < 2; 47% for 2 ≤ VMS Frequency < 7; 17% for 7 ≤ VMS Frequency < 9; and 19% for VMS Frequency ≥ 9, which is significantly different from the baseline distribution of DAYLIGHT (42% for 7 ≤ VMS Frequency < 9; 58% for VMS Frequency ≥ 9). If such a rapid change in VMS were to be expected between baseline and one year in the absence of treatment, it would suggest a lack of need for treatments for moderate to severe VMS because the symptoms would resolve rapidly by themselves over time. This suggests that either the results of the company's SEE exercise are unreliable for assessing the natural history in the absence of treatment, or the company's baseline VMS frequency distribution derived from DAYLIGHT is representative of a population with seven or greater moderate to severe VMS and where participants are likely to have had symptoms for several years before recruitment into the trial and, therefore, not representative of moderate to severe VMS in the NHS in the absence of treatment. In response to EAG clarifications (question B11b), the company acknowledges that there is a disconnect between the baseline distribution and the natural history used in the model because the SEE exercise was not anchored on the trial baseline distribution when eliciting values of natural history.

Scenario 1 is split into two scenarios 1a and 1b, where alternative assumptions for the baseline distribution are assumed, while keeping the year 6 SEE estimates for the natural history as per the company's base case:

- Scenario 1a assumes Year 1 SEE estimates for the baseline distribution.
- Scenario 1b assumes Year 1 clinician natural history distribution estimates for baseline distribution.

Table 34 Baseline distribution estimates

Source of estimates	0 ≤ VMS-F < 2	2 ≤ VMS-F < 7	7 ≤ VMS-F < 9	VMS-F ≥ 9
SEE natural history estimates	16.91%	47.13%	17.16%	18.80%
Clinician natural history estimates	10%	30%	40%	20%

6.1.1.2 Scenario 2: Placebo effect from trials

As discussed in Section 4.2.6.1, the company's base case approach assumes no adjustment for placebo effect from the trials after week 12. The implication of this assumption is that the placebo effect is lost arbitrarily from week 12 in the model for no active treatment and reverts to the natural history of VMS, while fezolinetant continues to receive the absolute changes in 4-weekly cycles observed in the separate fezolinetant arm of DAYLIGHT up to week 24 and the pooled absolute changes observed in the fezolinetant arm of SKYLIGHT 1 and 2 from week 24 to week 52, without adjustment for placebo effect in the trials. This can be observed in the proportions of the cohort in the VMS frequency health states between week 12 and week 52 in the model (Table 19), where there is a notable abrupt shift for no active treatment in the proportion in the lower VMS frequency categories ($0 \leqslant$ VMS-F < 2 and $2 \leqslant$ VMS-F < 7) at week 12 to the higher VMS frequency categories ($7 \leqslant$ VMS-F < 9 and VMS-F $\geqslant 9$) after week 12. Consequently, this means that the cost-effectiveness of fezolinetant relative to no active treatment is highly sensitive to the estimates of natural history of moderate to severe VMS frequency that is assumed in the model. It also means that the base case analysis is favouring fezolinetant because the absolute changes observed in the fezolinetant arm of the trials is implemented without any deduction of the placebo effect from week 12 onwards, but the relative difference between fezolinetant and placebo appears fairly constant from about week 6 to week 24 in mean frequency of moderate to severe VMS per day from DAYLIGHT (see Figure 7, p74 of the CS).

Scenario 2 is split into four separate Scenarios, 2a, 2b, 2c and 2d, where alternative assumptions for placebo effect are assumed:

- Scenario 2a assumes placebo effect until week 24 using data from the DAYLIGHT trial.
- Scenario 2b assumes no placebo effect in the no active treatment arm of the model.
- Scenario 2c assumes placebo effect up to week 52 using data from the DAYLIGHT trial for week 0-24 and for weeks 25-52, an average of DAYLIGHT transition probabilities for week 0-24.
- Scenario 2d assumes the placebo effect from the trials is extrapolated over the time horizon of the model in the same way as the effect for fezolinetant is extrapolated over the time horizon of the model, using data from the DAYLIGHT trial for week 0-24 and an average of DAYLIGHT week 0-24 transition probabilities for the longer-term extrapolation.

6.1.1.3 Scenario 3: Natural history in the absence of treatment

As discussed in Section 4.2.6.3, the EAG is concerned about the credibility of the results of the SEE exercise for informing the natural history of moderate to severe VMS over time when a baseline VMS frequency distribution with a minimum average of seven moderate to severe VMS per day is used in the model. As noted in Scenario 1, the EAG believes that the baseline distribution should be closer to

the elicited values at one year for the natural history of postmenopause women in the UK and the rapid reduction in the frequency of moderate to severe VMS between baseline and one year in the absence of treatment leads to the cost-effectiveness of fezolinetant relative to no active treatment being highly sensitive to the estimates of natural history of moderate to severe VMS frequency that is assumed in the model. This was also confirmed by the clinical experts informing the company who supported the view that such a significant reduction in symptoms is unlikely to occur within the first year. The approach used by the company to overcome this concern involved discarding the SEE estimates at years 1 and 3 and only use the estimates elicited at year 6 in the base case analysis, with a linear change from the baseline distribution of DAYLIGHT up to year 6 applied. The EAG does not consider this approach satisfactory because it highlights a loss of credibility of the SEE estimates by suggesting that either the SEE estimates are unreliable or that the baseline VMS frequency distribution is not representative of moderate to severe VMS observed in UK clinical practice in the absence of treatment. Furthermore, there was considerable uncertainty in the resulting SEE estimates and between-expert variation in the estimates.

Scenario 3 is split into four separate scenarios, 3a, 3b, 3c and 3d, where alternative assumptions used to inform the natural history of moderate to severe VMS are used:

- Scenario 3a assumes Year 1, 3 and 6 SEE estimates for natural history.
- Scenario 3b assumes Year 3 and 6 SEE estimates for natural history.
- Scenario 3c assumes Year 1, 3 and 6 clinician natural history estimates.
- Scenario 3d assumes Year 3 and 6 clinician natural history estimates.

Table 35 SEE natural history health state distribution proportions (reproduced from Table 61 of CS)

Time point	0 ≤ VMS-F < 2	2 ≤ VMS-F < 7	7 ≤ VMS-F < 9	VMS-F ≥ 9
SEE natural history estimates				
Year 1	16.91%	47.13%	17.16%	18.80%
Year 3	36.77%	40.34%	13.36%	9.53%
Year 6	48.12%	33.33%	12.39%	6.15%
Clinician natural history estimates				
Year 1	10%	30%	40%	20%
Year 3	20%	40%	30%	10%
Year 6	50%	40%	5%	5%

6.1.1.4 Scenario 4: VMS duration

As discussed in Section 4.2.6.4, the EAG considers the approach used by the company to capture the natural cessation of VMS in menopausal people to be appropriate, but the EAG is concerned that the median duration of moderate to severe VMS of 7.4 years is too high for the company's modelled base

case population of postmenopausal people (starting age 51 years at the onset of menopause). This median estimate represents the entire follow-up period in the study of women pre-menopause or early perimenopause through to postmenopause, whereas the company's modelled population is postmenopausal people with a mean starting age of 51 years at the onset of menopause because both the effectiveness evidence for fezolinetant is based only on postmenopause data and the natural history of VMS used in the model is elicited for postmenopause women. Therefore, the median duration of VMS is not reflecting the modelled population. The EAG notes that the Avis et al., 2015[5] study reports separately a median VMS duration of 3.4 years in the subset of women who were postmenopause at the onset of VMS (20.2% of the total cohort of 3302 women), which is more reflective of the company's modelled base case population of postmenopausal people. Furthermore, the company applied an exponential distribution to the median VMS duration of 7.4 years, which results in a probability of VMS cessation for menopausal people of approximately 60% by the end of the modelled time horizon of 10 years. Consequently, the EAG considers the modelled time horizon to be too short to capture the natural cessation of VMS in most postmenopausal people when assuming a median VMS duration of 7.4 years.

Scenario 4 is split into two scenarios 4a and 4b, where alternative assumptions for the duration of VMS are used:

- Scenario 4a assumes a median duration of VMS of 3.4 years for postmenopausal people.
- Scenario 4b assumes an increased time horizon from 10 to 20 years when using median duration VMS of 7.4 years.

6.1.1.5 Scenario 5: Health-related quality of life

As discussed in Section 4.2.8.1, the EAG is concerned about several assumptions the company made in their estimation of the utility values used in the model. First, although the company stated that they applied GEE linear regression using the DAYLIGHT data to estimate the utility values, scarce information was provided on how the final model was selected making it difficult to assess the appropriateness of the final model. Second, the EAG is concerned that the utility values from the DAYLIGHT trial for all health states of the model were higher in the placebo arm compared to the fezolinetant arm (see Table 22). In contrast, the utility values from the SKYLIGHT 1 and 2 trials were higher in the fezolinetant arm than in the placebo arm. However, given that the utility data from SKYLIGHT 1 and 2 were available only up to week 12, after which placebo participants could cross-over to the fezolinetant arm, it is difficult for the EAG to interpret the utility values from SKYLIGHT 1 and 2 relative to DAYLIGHT. Furthermore, the EAG notes that all utility values in SKYLIGHT 1 and 2 were relatively close to the age-adjusted utility values for women in the general population. Finally, the EAG considers that the adjustment of the utility values with clinical expert opinion is not an appropriate approach. The company stated that the EQ-5D measure is not sufficient to capture

changes in quality of life due to reduced frequency and/or severity of VMS, which was used as justification for the finding that there were no statistically significant differences in EQ-5D between the fezolinetant and placebo arms in DAYLIGHT, including the fact that the values in the placebo arm were higher than in the fezolinetant arm for all health states. Thus, the company decided to adjust these values based on the opinion of two clinical experts. This adjustment was validated by one clinical expert and 4 patient experts. The EAG considers this method to adjust utility values to be highly uncertain and not appropriate. Consequently, given these concerns, the modelled health state utility values are highly uncertain.

Scenario 5 is split into three scenarios, 5a, 5b and 5c, where alternative assumptions for the utility values were used:

- Scenario 5a assumes utility values based on the DAYLIGHT data (UK) without adjustment with clinical expert opinion.
- Scenario 5b assumes utility values based on the pooled DAYLIGHT & SKYLIGHT 1 & 2 (HRT-unsuitable) data (UK).
- Scenario 5c assumes utility values based on the pooled SKYLIGHT 1 & 2 (HRT-unsuitable) data.

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

Table 36 shows the results of the EAG scenarios (Scenarios 1-5). The scenario with the largest impact on the ICER is Scenario 1a, where the baseline VMS frequency distribution reflects that of the Year 1 SEE estimates for the UK natural history in the absence of treatment rather than the high VMS frequency starting distribution used in the trials to demonstrate a treatment effect. This is driven by a substantial decrease in costs for no treatment arm but also an increase in QALYs associated with no treatment compared to fezolinetant.

The other EAG scenarios that have a large impact on the company's base case results are Scenario 2d, placebo effect extrapolated over the modelled time horizon, Scenario 3a, use of Year 1, 3 and 6 SEE estimates for natural history, Scenario 1b, use of Year 1 clinician natural history distribution estimates for baseline distribution, and Scenario 5c, utility values based on data from the pooled SKYLIGHT 1 and 2 trials. Scenario 2d, which assumes placebo effect for the modelled time horizon worsens the cost-effectiveness of fezolinetant compared to no treatment because it decreases the costs of no treatment arm and increases the QALYs associated with this arm while keeping the costs and QALYs associated with fezolinetant arm unchanged. These changes in costs and QALYs for no treatment arm

are caused by postponing the abrupt shift from the lower moderate to severe VMS frequency categories to higher moderate to severe VMS frequency and keeping a large proportion of the cohort in the lower frequency health states (⩽ 7 VMS-F), which are associated with lower costs and higher utility values, until week 52 instead of week 12 in the company's base case analysis. While the EAG considers this scenario to be unrealistic in NHS practice, it demonstrates the impact of the placebo effect from the trials on the estimate of cost-effectiveness of fezolinetant. The company should have incorporated a relative treatment effect from the trials in its base case analysis to derive the transition probabilities at different time points, instead the company have extrapolated the absolute effect from the fezolinetant arm of the trials over the modelled time horizon without any adjustment for the placebo effect observed in the trials from week 12 onwards (see Section 4.2.6.1). The other scenarios on placebo effect also show a significant impact on the cost-effectiveness results (see Scenarios 2c and 2d). Scenario 3a, which assumed Year 1, 3 and 6 SEE estimates for natural history, leads to a decrease in total costs and an increase in QALYs for both fezolinetant and no treatment because a substantially larger proportion of the cohort stays in the lower frequency moderate to severe VMS health states, which are associated with lower costs and higher utility values, for a longer period. Moreover, the change in the no treatment arm was greater than in the fezolinetant arm (i.e., more substantial decrease in costs and increase in QALYs associated with no treatment) because of the company's approach to remove the placebo effect after week 12, which results in an abrupt shift of the cohort to the natural history distribution. Thus, a large proportion of patients remain in lower moderate to severe VMS frequency categories instead of moving to high moderate to severe VMS frequency categories after week 12. The change in costs and QALYs for Scenario 1b (use of Year 1 clinician natural history distribution estimates for baseline distribution) leads to a decrease in costs and an increase in QALYs for both fezolinetant and no treatment arms but the change is greater in the no treatment arm (i.e., greater decrease in costs and increase in QALYs associated with no treatment arm), resulting in higher incremental costs and lower incremental QALYs for no treatment compared to fezolinetant. The cause of these changes is the same as for Scenario 1a and 3a. Finally, in Scenario 5b (utility values based on the pooled DAYLIGHT and SKYLIGHT 1 and 2 data), while costs of both fezolinetant and no treatment arms remained unchanged, QALYs associated with both these arms increased substantially because the utility values derived from SKYLIGHT 1 and 2 were substantially higher than in DAYLIGHT and were relatively close to the age-adjusted utility values for women in the general population. The increase in QALYs associated with no treatment arm was greater leading to smaller incremental QALYs associated with no treatment compared to fezolinetant than in the base case analysis. This is because the utility values derived from DAYLIGHT were higher for health states with higher moderate to severe VMS frequency than the utility values that were derived from DAYLIGHT and adjusted with clinical opinion.

Table 36 Cost-effectiveness results of the EAG scenario analyses

Scenario #		Name	Option	Costs	QALYs	Inc. Costs	Inc. QALYs	ICER, /QALY	% change of ICER
		Company's base-case results	Fezolinetant	£2,803.18	6.68	-	-	-	-
			No Treatment	£1,603.69	6.57	£1,199.49	0.116	£10,364.17	
Baseline distribution	1a	Use Year 1 SEE estimate for baseline distribution	Fezolinetant	£2,629.86	6.72	-	-	-	-
			No Treatment	£1,084.87	6.67	£1,545.00	0.045	£34,308.17	231%
	1b	Use Year 1 Clinician Natural History Distribution Estimates for baseline distribution	Fezolinetant	£2,672.25	6.71	-	-	-	-
			No Treatment	£1,213.54	6.65	£1,458.71	0.064	£22,641.88	118%
Placebo effect	2a	Increase the placebo effect from week 12 to week 24	Fezolinetant	£2,803.18	6.68	-	-	-	-
			No Treatment	£1,561.31	6.58	£1,241.87	0.107	£11,621.78	12%
	2b	No placebo effect at all	Fezolinetant	£2,803.18	6.68	-	-	-	-
			No Treatment	£1,653.46	6.56	£1,149.72	0.126	£9,119.42	-12%
	2c	Placebo effect up to week 52, probability of discontinuation of 4.17% per 4-week (company's base case)	Fezolinetant	£2,803.18	6.68	-	-	-	-
			No Treatment	£1,471.56	6.60	£1,331.62	0.088	£15,126.91	46%
	2d	Placebo effect over modelled time horizon, probability of discontinuation of 4.17% per 4-week (company's base case)	Fezolinetant	£2,803.18	6.68	-	-	-	-
			No Treatment	£1,321.69	6.63	£1,481.50	0.054	£27,191.69	162%
	3a		Fezolinetant	£2,644.89	6.72	-	-	-	-

Natural history study		Use Year 1, 3 and 6 SEE estimates for natural history	No Treatment	£1,156.59	6.66	£1,488.30	0.057	£26,171.46	153%
	3b	Use Year 3 and 6 SEE estimates for natural history	Fezolinetant	£2,689.18	6.71	-	-	-	-
			No Treatment	£1,330.64	6.62	£1,358.54	0.083	£16,333.99	58%
	3c	Use Year 1,3 and 6 clinician natural history estimates	Fezolinetant	£2,653.00	6.71	-	-	-	-
			No Treatment	£1,243.70	6.64	£1,409.30	0.075	£18,718.94	81%
	3d	Use Year 3 and 6 clinician natural history estimates	Fezolinetant	£2,683.45	6.71	-	-	-	-
			No Treatment	£1,363.37	6.61	£1,320.08	0.093	£14,246.24	37%
VMS duration	4a	Median duration of VMS of 3.4 years	Fezolinetant	£2,216.72	6.78	-	-	-	-
			No Treatment	£1,342.11	6.69	£874.61	0.096	£9,138.54	-12%
	4b	Increase model time-horizon to 20 years when using median duration VMS of 7.4 years	Fezolinetant	£3,250.41	11.09	-	-	-	-
			No Treatment	£1,967.42	10.97	£1,283.00	0.115	£11,118.53	7%
Utility values	5a	Utility values-DAYLIGHT data (UK)	Fezolinetant	£2,803.18	6.80	-	-	-	-
			No Treatment	£1,603.69	6.71	£1,199.49	0.085	£14,188.67	37%
	5b	Utility values-Pooled DAYLIGHT & SKYLIGHT 1 & 2 (HRT-unsuitable) pooled data (UK)	Fezolinetant	£2,803.18	6.89	-	-	-	-
			No Treatment	£1,603.69	6.82	£1,199.49	0.073	£16,359.38	58%
	5c		Fezolinetant	£2,803.18	6.97	-	-	-	-

		Utility values- Pooled SKYLIGHT 1 & 2 (HRT-unsuitable)	No Treatment	£1,603.69	6.91	£1,199.49	0.064	£18,686.20	80%

6.3 EAG's preferred assumptions

The EAG's presents a restricted base case that includes the following changes from the company's base case assumptions:

- An increase in the placebo effect from DAYLIGHT from week 12 to week 24 for no treatment.
- Median duration of VMS of 3.4 years for postmenopausal people.
- Utility values based on the pooled DAYLIGHT and SKYLIGHT 1 & 2 (HRT-unsuitable) data (UK).

Table 37 shows the cumulative impact of the EAG's restricted base case assumptions on the ICER.

The selection of changes made to the EAG's restricted base case comes with the caveat that the EAG is unable to address the key structural uncertainties in the company's base case analysis because the company have not addressed the EAG's concerns at points for clarification, namely:

- The need to incorporate relative treatment effects in the model that adjust for the placebo effect observed in the trials.
- The need to incorporate a baseline distribution of moderate to severe VMS frequency that reflects that seen in the NHS for menopausal people who are likely to receive treatment with fezolinetant.
- Structural concerns regarding the health states used in the model, where arbitrary cut-off thresholds are used to define moderate to severe VMS frequency health states in the model and the concern that frequency of moderate to severe VMS is not commonly used to define severity of VMS in NHS clinical practice.
- No active treatment is unlikely to be the sole relevant comparator to fezolinetant in the NHS.
- Uncertainty about the natural history in the absence of treatment, which is a key driver of the cost-effectiveness of fezolinetant relative to no treatment.
- Absence of data for perimenopausal people, with cost-effectiveness of fezolinetant only assessed in postmenopausal people.

Therefore, a number of important uncertainties remain. The EAG is particularly concerned that the baseline distribution with a moderate to severe VMS frequency limit of greater than seven per day from the DAYLIGHT trial is unlikely to be appropriate for assessing the need for treatment in the NHS and is unlikely to match that seen in UK clinical practice for postmenopausal people because the trial starting distribution is very different from the elicited Year 1 Clinician Natural History Distribution for natural history in the absence of treatment. The EAG is also concerned about uncertainty in the natural history estimates from the SEE exercise. The EAG presents two additional scenarios on the EAG's restricted base case:

- Use Year 1 clinician natural history distribution estimates for baseline distribution.
- Use Year 3 and 6 clinical validation estimates for natural history.

Table 38 shows the impact of the alternative assumptions on the EAG's restricted base case.

Table 37 Cumulative cost-effectiveness results for the EAG's restricted base case

Scenario #	Name	Option	Costs	QALYs	Inc. Costs	Inc. QALYs	ICER, /QALY	% change of company's base-case ICER
	Company's base case results	Fezolinetant	£2,803.18	6.68	-	-	-	-
		No Treatment	£1,603.69	6.57	£1,199.49	0.116	£10,364.17	-
2a	Increase the placebo effect from week 12 to week 24	Fezolinetant	£2,803.18	6.68	-	-	-	-
		No Treatment	£1,561.31	6.58	£1,241.87	0.107	£11,621.78	12%
2a+4a	Median duration of VMS of 3.4 years	Fezolinetant	£2,216.72	6.78	-	-	-	-
		No Treatment	£1,301.47	6.69	£915.25	0.087	£10,496.77	1%
2a+4a+5b (EAG restricted base case)	Utility values based on pooled DAYLIGHT and SKYLIGHT 1 & 2 (HRT-unsuitable) data (UK)	Fezolinetant	£2,216.72	6.95	-	-	-	-
		No Treatment	£1,301.47	6.89	£915.25	0.056	£16,470.01	59%

Table 38 Cost-effectiveness results for alternative assumptions on the EAG's restricted base case

Scenario #	Name	Option	Costs	QALYs	Inc. Costs	Inc. QALYs	ICER, /QALY
	EAG restricted base case (2a+4a+5b)	Fezolinetant	£2,216.72	6.95	-	-	-
		No Treatment	£1,301.47	6.89	£915.25	0.056	£16,470.01
EAG restricted base case + 1b +3d							
1b	Use Year 1 clinical validation estimates for baseline distribution	Fezolinetant	£2,113.65	6.96	-	-	-
		No Treatment	£1,002.43	6.92	£1,111.22	0.035	£31,304.74
1b+3d	Use Year 3 and 6 clinical validation estimates for natural history	Fezolinetant	£2,093.06	6.96	-	-	-
		No Treatment	£973.60	6.92	£1,119.46	0.037	£30,433.64

6.4 Conclusions of the cost effectiveness section

The company submitted a decision model to compare the cost-effectiveness of fezolinetant with no active treatment in menopausal people with moderate to severe VMS frequency of ⩾7 per day who are deemed unsuitable for HRT. In general, the EAG is concerned that the company's base case analysis does not adequately reflect menopausal people who are likely to receive fezolinetant treatment in the NHS. The EAG has several major concerns in relation to the modelled population, model structure, baseline starting distribution, method used to incorporate treatment effects from the trials, natural history estimates used in the model and adjusted modelled health state utility values.

The company makes it clear in their response to EAG clarifications that it does not intend to restrict the population covered in the decision problem to only postmenopausal people, i.e., it includes both perimenopause and postmenopausal people. However, the evidence used in the cost-effectiveness analysis is predominantly for postmenopausal people as the treatment effects are from the DAYLIGHT and SKYLIGHT trials in postmenopausal people and the natural history estimates used in the model were elicited for postmenopausal people. Thus, no evidence of clinical effectiveness or natural history has been provided for perimenopausal people.

Another key concern is in relation to the use of moderate to severe VMS frequency as a proxy for VMS severity in the model. The company states that this assumption was made in the model because: 1) it is aligned with the DAYLIGHT and SKYLIGHT trials; 2) calculations incorporating both frequency and severity of moderate to severe VMS would not be possible for a probabilistic analysis; and 3) correlation between frequency and severity from the trials is demonstrated. The definition of moderate to severe VMS used in the design of DAYLIGHT and SKYLIGHT 1 and 2 trials did not include any requirement of VMS frequency, which is consistent with the US FDA and EMA guidance about definition of moderate to severe VMS. In NHS clinical practice, VMS frequency is not commonly used to define moderate to severe VMS. Also, the EAG notes that the precision estimates needed to inform probability distributions for parameters to inform a probabilistic analysis should be possible, based on the mean frequency of moderate to severe VMS and the mean weighted severity of VMS experienced per 24 hours. Moreover, the correlation analysis provided by the company are indicative of moderate to weak correlation between frequency and severity (estimated values are under [redacted]).

The EAG is also concerned about the cut-off thresholds used to define moderate to severe VMS frequency health states in the model. The company provided a rationale behind the cut-off thresholds, which was mostly based on a statistically significant difference between the frequency categories, and a statistically significant difference in utility values associated with different frequency levels. However, utility values between the frequency categories differ slightly and the final cut-off

thresholds were based on a statistically significant difference comparing each health state to the previous one and feedback from the clinical experts. The EAG notes that among the explored models, several other models appear to be associated with more favourable p-values. Thus, the EAG considers the justification for the choice of the cut-off thresholds to be uncertain.

The company compared fezolinetant to no active treatment in the model, and the EAG is concerned that no active treatment may not be the sole relevant comparator to fezolinetant. Menopausal people presenting with moderate to severe VMS and deemed HRT-unsuitable are more likely to be offered non-hormonal pharmacological treatments as opposed to 'nothing' in NHS clinical practice.

A further concern from the EAG relates to how the company have incorporated treatment effects in the model, which are based on absolute effects from the fezolinetant arm of the trials without adjustment for the placebo effect from week 12 onwards. The company's base case approach (without adjustment for placebo effect) is that the placebo effect is lost arbitrarily from week 12 in the model. The EAG agrees that the notable improvements in moderate to severe VMS frequency observed from baseline through to week 24 in the placebo arm of DAYLIGHT are unlikely to be observed in clinical practice outside the clinical trial setting, but any improvement associated with partaking in a clinical trial would be expected to be applied to both the fezolinetant and placebo arms of the trials and therefore adjustment should be made for these effects. The appropriate method to incorporate this adjustment is to use relative treatment effects for fezolinetant in the model rather than absolute effects observed in the trials.

The EAG is also concerned about the natural history estimates used in the model in the absence of treatment, which is a key driver for the cost-effectiveness of fezolinetant relative to no treatment. The credibility of the SEE estimates for informing the natural history of moderate to severe VMS over time is unclear, when a baseline VMS frequency distribution with a minimum average of seven moderate to severe VMS per day is used in the model. There was considerable uncertainty in the resulting SEE estimates and between-expert variation in the estimates. The SEE exercise was not anchored on the trial baseline VMS frequency distribution, so it is possible that the experts who provided proportions experiencing different frequencies of daily moderate to severe VMS at years 1, 3 and 6 had a very different starting distribution of moderate to severe VMS frequencies per day in mind. The EAG notes that clinical opinion is unlikely to be a reliable source to elicit how VMS progresses over time in the absence of treatment because it represents a cohort that would rarely exist since clinicians would not leave people experiencing moderate to severe VMS untreated and people are unlikely to report back when there are no suitable treatment options available.

Furthermore, the median total VMS duration of 7.4 years that includes women pre-menopause or early perimenopause through to postmenopause does not align with the modelled population of

postmenopausal people with a mean starting age of 51 years at the onset of menopause and the modelled time horizon of 10 years. The EAG notes that using the median VMS duration of 7.4 years from Avis et al., (2015) study in the model results in approximately 40% of postmenopausal people are still experiencing VMS by year 10. The EAG considers the median VMS duration of 3.4 years from the same study to be more consistent with the modelled population of postmenopausal people, which results in approximately 12% of people still experiencing VMS by year 10.

The EAG's final concern is the substantial uncertainty in the modelled health state utility values. The scarce information was provided on how the final utility estimation model was selected. The EAG could not assess the appropriateness of the methods applied. The EAG is concerned about the utility data from DAYLIGHT and SKYLIGHT 1&2 trials, which does not reflect the consistent direction of changing in two arms (all utility values were higher in the placebo arm than in the fezolinetant arm in DAYLIGHT, but the opposite applies in the SKYLIGHT 1 & 2 trials). More importantly, the EAG considers that the adjustment of the utility values with the opinion of two clinical experts is not an appropriate approach.

The modelled assumptions with the largest impact on the cost-effectiveness results are those relating to: (i) baseline distribution, (ii) unadjusted long-term placebo effect, (iii) natural history estimates, and (iv) clinician opinion adjusted utility values. Since there are many critical issues and uncertainties in the company's model, the restricted EAG assumptions include the following changes from the company's base case: (i) an increase in the placebo effect from DAYLIGHT from week 12 to week 24 for no treatment; (ii) use of median duration of VMS of 3.4 years for postmenopausal people, and (iii) use of pooled DAYLIGHT and SKYLIGHT 1 & 2 (HRT-unsuitable) health state utility values. The resulting ICER increases from the company's base case ICER of £10,364 per QALY gained for fezolinetant relative to no treatment to £16,470 per QALY gained under the EAG's restricted base case. The EAG has critical concerns about the baseline distribution and natural history estimates used in the model, which have a significant impact on the ICER. However, due to the insufficient evidence provided by the company and inflexible model settings, it is challenging for the EAG to select any of the scenarios as EAG's preferred base case. Thus, the EAG provides additional exploratory scenarios on the EAG's restricted base case, by including (i) use of Year 1 clinical validation estimates for baseline distribution, and (ii) use of Year 3 and 6 clinical validation estimates for natural history. The ICER changes significantly to £30,434 per QALY gained for fezolinetant compared to no treatment, which highlights the sensitivity of the cost-effectiveness results to the assumptions used in the model.

7 SEVERITY MODIFIER

The CS does not apply the severity modifier to the cost effectiveness results. The EAG considers this to be appropriate as the severity weights are not applicable for people with VMS associated with menopause.

8 REFERENCES

1. Hunter M, Gentry-Maharaj A, Ryan A, Burnell M, Lanceley A, Fraser L, et al. Prevalence, frequency and problem rating of hot flushes persist in older postmenopausal women: impact of age, body mass index, hysterectomy, hormone therapy use, lifestyle and mood in a cross-sectional cohort study of 10 418 British women aged 54–65. *BJOG* 2012;**119**:40-50.
2. Thurston RC, Bromberger JT, Joffe H, Avis NE, Hess R, Crandall CJ, et al. Beyond frequency: who is most bothered by vasomotor symptoms? *Menopause* 2008;**15**:841-7.
3. NHS. *Menopause - overview*. NHS; Available from: https://www.nhs.uk/conditions/menopause/ [accessed 2nd December 2024].
4. Politi MC, Schleinitz MD, Col NF. Revisiting the duration of vasomotor symptoms of menopause: a meta-analysis. *J Gen Intern Med* 2008;**23**:1507-13.
5. Avis NE, Crawford SL, Greendale G, Bromberger JT, Everson-Rose SA, Gold EB, et al. Duration of menopausal vasomotor symptoms over the menopause transition. *JAMA Intern Med* 2015;**175**:531-9.
6. Kiran A, Schultz NM, Siddiqui E, Todorova L, Van der Poel B, Stoelzel M, et al. Epidemiology and treatment patterns of UK women diagnosed with vasomotor symptoms: findings from the Clinical Practice Research Datalink GOLD database. *Maturitas* 2022;**164**:1-8.
7. Hickey M, Szabo RA, Hunter MS. Non-hormonal treatments for menopausal symptoms. *BMJ* 2017;**359**:j5101.
8. Kumari M, Stafford M, Marmot M. The menopausal transition was associated in a prospective study with decreased health functioning in women who report menopausal symptoms. *J Clin Epidemiol* 2005;**58**:719-27.
9. Nappi RE, Siddiqui E, Todorova L, Rea C, Gemmen E, Schultz NM. Prevalence and quality-of-life burden of vasomotor symptoms associated with menopause: a European cross-sectional survey. *Maturitas* 2023;**167**.
10. DePree B, Shiozawa A, King D, Schild A, Zhou M, Yang H, et al. Association of menopausal vasomotor symptom severity with sleep and work impairments: a US survey. *Menopause* 2023;**30**:887-97.
11. Freeman EW, Sammel MD, Gross SA, Pien GW. Poor sleep in relation to natural menopause: a population-based 14-year follow-up of midlife women. *Menopause* 2015;**22**:719-26.
12. Williams RE, Levine KB, Kalilani L, Lewis J, Clark RV. Menopause-specific questionnaire assessment in US population-based study shows negative impact on health-related quality of life. *Maturitas* 2009;**62**:153-9.
13. Natari RB, Clavarino AM, McGuire TM, Dingle KD, Hollingworth SA. The bidirectional relationship between vasomotor symptoms and depression across the menopausal transition: a systematic review of longitudinal studies. *Menopause* 2018;**25**:109-20.
14. DiBonaventura MD, Chandran A, Hsu MA, Bushmakin A. Burden of vasomotor symptoms in France, Germany, Italy, Spain, and the United Kingdom. *Int J Womens Health* 2013:261-9.
15. English M, Stoykova B, Slota C, Doward L, Siddiqui E, Crawford R, et al. Qualitative study: burden of menopause-associated vasomotor symptoms (VMS) and validation of PROMIS Sleep Disturbance and Sleep-Related Impairment measures for assessment of VMS impact on sleep. *J Patient Rep Outcomes* 2021;**5**:37.
16. Whiteley J, Wagner JS, Bushmakin A, Kopenhafer L, Dibonaventura M, Racketa J. Impact of the severity of vasomotor symptoms on health status, resource use, and productivity. *Menopause* 2013;**20**:518-24.

17. Woods NF, Mitchell ES. Symptoms during the perimenopause: prevalence, severity, trajectory, and significance in women's lives. *Am J Med* 2005;**118**:14-24.
18. Fawcett. *Menopause and the workplace*. 2022. Available from: https://www.fawcettsociety.org.uk/menopauseandtheworkplace [accessed 2nd December 2024].
19. Hardy C, Thorne E, Griffiths A, Hunter MS. Work outcomes in midlife women: the impact of menopause, work stress and working environment. *Womens Midlife Health* 2018;**4**:3.
20. National Institute for Health and Care Excellence. *Menopause: identification and management. NICE guideline [NG23]. Last updated: 07 November 2024*. NICE; 2015. Available from: https://www.nice.org.uk/guidance/ng23 [accessed 13th November 2024].
21. European Medicines Agency. *Guideline on clinical investigation of medicinal products for hormone replacement therapy of oestrogen deficiency symptoms in postmenopausal women*. Amsterdam: EMA; 2005. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-hormone-replacement-therapy-oestrogen-deficiency_en.pdf
22. National Institute for Health and Care Excellence (NICE). Menopause: diagnosis and management [NG23]. Available at: https://www.nice.org.uk/guidance/NG23 [Accessed September 2024]. 2024.
23. Constantine GD, Graham S, Clerinx C, Bernick BA, Krassan M, Mirkin S, et al. Behaviours and attitudes influencing treatment decisions for menopausal symptoms in five European countries. *Post Reprod Health* 2016;**22**:112-22.
24. British Menopause Society. *BMS & WHC's 2020 recommendations on hormone replacement therapy in menopausal women*. Marlow: British Menopause Society; 2023.
25. Morga A, Ajmera M, Gao E, Patterson-Lomba O, Zhao A, Mancuso S, et al. Systematic review and network meta-analysis comparing the efficacy of fezolinetant with hormone and nonhormone therapies for treatment of vasomotor symptoms due to menopause. *Menopause* 2023;**31**:68-76.
26. Shams T, Firwana B, Habib F, Alshahrani A, Alnouh B, Murad MH, et al. SSRIs for hot flashes: a systematic review and meta-analysis of randomized trials. *J Gen Intern Med* 2014;**29**:204-13.
27. Toulis KA, Tzellos T, Kouvelas D, Goulis DG. Gabapentin for the treatment of hot flashes in women with natural or tamoxifen-induced menopause: a systematic review and meta-analysis. *Clin Ther* 2009;**31**:221-35.
28. Umland EM, Falconieri L. Treatment options for vasomotor symptoms in menopause: focus on desvenlafaxine. *Int J Womens Health* 2012;**4**:305-19.
29. National Institute for Health and Care Excellence. *Fezolinetant for treating vasomotor symptoms associated with the menopause (ID5071). Response to stakeholder organisation comments on the draft remit and draft scope*. London: NICE; 2024. Available from: https://www.nice.org.uk/guidance/gid-ta11058/documents/draft-scope-and-provisional-matrix-comments-table-post-referral-2 [accessed 5th December 2024].
30. Astellas. *Study report: Network Meta-Analysis (NMA) comparing fezolinetant and comparators for the treatment of vaso-motor symptoms (VMS) associated with menopause in hormone therapy unsuitable women*: Astellas; 2024.
31. Astellas Pharma Global Development Inc. *A study of fezolinetant to treat hot flashes in women going through menopause (Daylight)*. ClinicalTrials.gov. Bethesda (MD): National Library of Medicine (US); 2024. Available from: https://classic.clinicaltrials.gov/ct2/show/NCT05033886 [accessed 5th December 2024].
32. Lederman S, Shapiro CM, Stute P, Lee M, Wang X, Neal-Perry G. Phase 3 study of fezolinetant for treatment of moderate-to-severe vasomotor symptoms associated with menopause. *Obstet Gynecol* 2022;**139**:39S.
33. Nappi RE, Johnson K, Stute P, Wolfman W, Blogg M, English M, et al. Treatment of moderate-to-severe vasomotor symptoms associated with menopause with fezolinetant: pooled responder analysis from SKYLIGHT 1 and 2. *Maturitas* 2023;**173**:75.
34. Simon JA, Portman DJ, Kaunitz AM, Mekonnen H, Kazempour K, Bhaskar S, et al. Low-dose paroxetine 7.5 mg for menopausal vasomotor symptoms: two randomized controlled trials. *Menopause* 2013;**20**:1027-35.
35. National Institute for Health and Care Excellence. *Menopause: identification and management. NICE Guideline [NG23] Last updated: 07 November 2024.* NICE; 2015. Available from: https://www.nice.org.uk/guidance/NG23 [accessed 5th December 2024].

36. Astellas. UK HRT-unsuitable Consensus Statement.
37. Goldštajn M, Mikuš M, Ferrari FA, Bosco M, Uccella S, Noventa M, et al. Effects of transdermal versus oral hormone replacement therapy in postmenopause: a systematic review. *Arch Gynecol Obstet* 2023;**307**:1727-45.
38. Mohammed K, Abu Dabrh AM, Benkhadra K, Al Nofal A, Carranza Leon BG, Prokop LJ, et al. Oral vs transdermal estrogen therapy and vascular events: a systematic review and meta-analysis. *J Clin Endocrinol Metab* 2015;**100**:4012-20.
39. Golder S, Farrah K, Mierzwinski-Urban M, Barker B, Rama A. Updated generic search filters for finding studies of adverse drug effects in Ovid MEDLINE and Embase may retrieve up to 90% of relevant studies. *Health Info Libr J* 2023;**40**:190-200.
40. Fraser GL, Lederman S, Waldbaum A, Kroll R, Santoro N, Lee M, et al. A phase 2b, randomized, placebo-controlled, double-blind, dose-ranging study of the neurokinin 3 receptor antagonist fezolinetant for vasomotor symptoms associated with menopause. *Menopause* 2020;**27**:382-92.
41. Fraser GL, Obermayer-Pietsch B, Laven J, Griesinger G, Pintiaux A, Timmerman D, et al. Randomized controlled trial of neurokinin 3 receptor antagonist fezolinetant for treatment of polycystic ovary syndrome. *J Clin Endocrinol Metab* 2021;**106**:e3519-e32.
42. Depypere H, Lademacher C, Siddiqui E, Fraser GL. Fezolinetant in the treatment of vasomotor symptoms associated with menopause. *Expert Opin Investig Drugs* 2021;**30**:681-94.
43. Schultz NM, Morga A, Siddiqui E, Rhoten SE. Psychometric evaluation of the MENQOL instrument in women experiencing vasomotor symptoms associated with menopause. *Adv Ther* 2024;**41**:2233-52.
44. European Medicines Agency. *Assessment report veoza (fezolinetant)*. Amsterdam: EMA; 2023.
45. Douxfils J, Beaudart C, Dogné JM. Risk of neoplasm with the neurokinin 3 receptor antagonist fezolinetant. *Lancet* 2023;**402**:1623-5.
46. Dias S, Ades AE, Welton NJ, Jansen JP, Sutton AJ. *Network meta-analysis for decision making*. Hoboken, NJ: John Wiley & Sons Ltd; 2018.
47. Center for Drug Evaluation and Research, Food and Drug Administration. *Clinical review: fezolinetant (216578Orig1s000)*. Silver Spring, MD: FDA; 2023.
48. Douxfils J, Beaudart C, Dogné JM. Totality of evidence refutes neoplasm risk with fezolinetant - Authors' reply. *Lancet* 2024;**403**:1988.
49. Yu Q, Ming F, Ma J, Cai Y, Wang L, Ren M, et al. Long-term safety of fezolinetant in Chinese women with vasomotor symptoms associated with menopause: the phase 3 open-label MOONLIGHT 3 clinical trial. *J Int Med Res* 2024;**52**:3000605241246624.
50. US Food & Drug Administration. *FDA adds warning about rare occurrence of serious liver injury with use of Veozah (fezolinetant) for hot flashes due to menopause*. FDA; 2024. Available from: https://www.fda.gov/drugs/drug-safety-and-availability/fda-adds-warning-about-rare-occurrence-serious-liver-injury-use-veozah-fezolinetant-hot-flashes-due#:~:text=What%20safety%20concern%20is%20FDA,return%20liver%20function%20to%20normal [accessed 23rd October 2024].
51. Skorupskaite K, George JT, Veldhuis JD, Millar RP, Anderson RA. Neurokinin 3 receptor antagonism reveals roles for neurokinin B in the regulation of gonadotropin secretion and hot flashes in postmenopausal women. *Neuroendocrinology* 2018;**106**:148-57.
52. Prague JK, Roberts RE, Comninos AN, Clarke S, Jayasena CN, Nash Z, et al. Neurokinin 3 receptor antagonism as a novel treatment for menopausal hot flushes: a phase 2, randomised, double-blind, placebo-controlled trial. *Lancet* 2017;**389**:1809-20.
53. Archer DF, Dupont CM, Constantine GD, Pickar JH, Olivier S. Desvenlafaxine for the treatment of vasomotor symptoms associated with menopause: a double-blind, randomized, placebo-controlled trial of efficacy and safety. *Am J Obstet Gynecol* 2009;**200**:238.e1-10.
54. Archer DF, Seidman L, Constantine GD, Pickar JH, Olivier S. A double-blind, randomly assigned, placebo-controlled study of desvenlafaxine efficacy and safety for the treatment of vasomotor symptoms associated with menopause. *Am J Obstet Gynecol* 2009;**200**:172.e1-10.
55. Bouchard P, Panay N, de Villiers TJ, Vincendon P, Bao W, Cheng RJ, et al. Randomized placebo- and active-controlled study of desvenlafaxine for menopausal vasomotor symptoms. *Climacteric* 2012;**15**:12-20.

56. Depomed. *Study of Gabapentin Extended Release (G-ER) in the treatment of vasomotor (hot flashes/hotflushes) symptoms in postmenopausal women*. ClinicalTrials.gov. Bethesda (MD): National Library of Medicine (US); 2009. Available from: https://clinicaltrials.gov/study/NCT00755417 [accessed 15th November 2024].
57. Depomed. *Study of Gabapentin Extended Release (G-ER) in the treatment of vasomotor (hot flashes/hot flushes) symptoms in postmenopausal women*. ClinicalTrials.gov. Bethesda (MD): National Library of Medicine (US); 2009. Available from: https://clinicaltrials.gov/study/NCT00777023 [accessed 15th November 2024].
58. Pinkerton JV, Kagan, R., Portman, D., Sathyanarayana, R., Sweeney, M. and Breeze 3 Investigators. Phase 3 randomized controlled study of gastroretentive gabapentin for the treatment of moderate-to-severe hot flashes in menopause. *Menopause* 2014;**21**:567–73.
59. Pinkerton JV, Constantine G, Hwang E, Cheng RF. Desvenlafaxine compared with placebo for treatment of menopausal vasomotor symptoms: a 12-week, multicenter, parallel-group, randomized, double-blind, placebo-controlled efficacy trial. *Menopause* 2012;**20**:28-37.
60. Pinkerton JV, Archer DF, Guico-Pabia CJ, Hwang E, Cheng RF. Maintenance of the efficacy of desvenlafaxine in menopausal vasomotor symptoms: a 1-year randomized controlled trial. *Menopause* 2012;**20**:38-46.
61. Reddy SY, Warner H, Guttuso T, Messing S, DiGrazio W, Thornburg L, et al. Gabapentin, estrogen, and placebo for treating hot flushes: a randomized controlled trial. *Obstet Gynecol* 2006;**108**:41-8.
62. Sassarini J, Fox H, Ferrell W, Sattar N, Lumsden MA. Hot flushes, vascular reactivity and the role of the α-adrenergic system. *Climacteric* 2012;**15**:332-8.
63. Astellas Data on File. DAYLIGHT Clinical Study Report.
64. Speroff L, Gass M, Constantine G, Olivier S, Study 315 Investigators. Efficacy and tolerability of desvenlafaxine succinate treatment for menopausal vasomotor symptoms: a randomized controlled trial. *Obstet Gynecol* 2008;**111**:77-87.
65. General Medical Council. *Prescribing unlicensed medicines*. General Medical Council; 2021. Available from: https://www.gmc-uk.org/professional-standards/the-professional-standards/good-practice-in-prescribing-and-managing-medicines-and-devices/prescribing-unlicensed-medicines#:~:text=You%20should%20usually%20prescribe%20licensed,specific%20needs%20of%20the%20patient [accessed 4th December 2024].
66. Dias S, Welton NJ, Sutton AJ, Ades AE. *NICE DSU technical support document 2: a generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials (updated September 2016)*. Sheffield: Decision Support Unit, ScHARR, University of Sheffield; 2011.
67. Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G, Ades AE. *NICE DSU technical support document 4: inconsistency in networks of evidence based on randomised controlled trials (updated April 2014)*. Sheffield: Decision Support Unit, ScHARR, University of Sheffield; 2011.
68. Beaudoin F, McQueen R, Wright A. Fezolinetant for Moderate to Severe Vasomotor Symptoms Associated with Menopause: Effectiveness and Value; Final Evidence Report. *Institute for Clinical and Economic Review* 2023.
69. Botteman MF, Shah NP, Lian J, Pashos CL, Simon JA. A cost-effectiveness evaluation of two continuous-combined hormone therapies for the management of moderate-to-severe vasomotor symptoms. *Menopause* 2004;**11**:343-55.
70. 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**:199-207.
71. *SIGN Filters*. Available from: https://www.sign.ac.uk/what-we-do/methodology/search-filters/ [accessed 03.10.2023].
72. Coyle D, Cranney A, Tugwell P. Economic evaluation of norethisterone acetate/ethinylestradiol (FemHRT) for women with menopausal symptoms. *Pharmacoeconomics* 2003;**21**:661-9.
73. Brown A, Coyle D, Chen S, Cumming D MS. *Transdermal Hormone Replacement Therapy Patches for Women with Postmenopausal Symptoms: Economic Analysis of Short Term Use*. Ottawa: Canadian Coordinating Office for Health Technology Assessment; 2006.
74. Astellas. *(Data on File). Structured Expert Elicitation on Vasomotor Symptoms in Menopause to Inform Health Technology Assessment in the UK.* In.

75. Food and Drug Administration. *Guidance for industry. Estrogen and estrogen/progestin drug products to treat vasomotor symptoms and vulvar and vaginal atrophy symptoms - recommendations for clinical evaluation*. FDA; 2003. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/estrogen-and-estrogenprogestin-drug-products-treat-vasomotor-symptoms-and-vulvar-and-vaginal-atrophy [accessed 5th December 2024].
76. *Summary of Product Characteristics (SPC) veoza 45 mg film-coated tablets (fezolinetant)*. London: Medicines & Healthcare products Regulatory Agency (MHRA); 2023.
77. Iliodromiti S, Wang W, Lumsden MA, Hunter MS, Bell R, Mishra G, et al. Variation in menopausal vasomotor symptoms outcomes in clinical trials: a systematic review. *Bjog* 2020;**127**:320-33.
78. National Institute for Health and Care Excellence (NICE). *NICE health technology evaluations: the manual. NICE process and methods [PMG36]*. Available from: https://www.nice.org.uk/process/pmg36/chapter/introduction-to-health-technology-evaluation [accessed 6 June 2024
79. Centre for Health Economics. *Structured expert elicitation resources (STEER)*. University of York; Available from: https://www.york.ac.uk/che/economic-evaluation/steer/ [accessed 5th December 2024].
80. Office for National Statistics. *National life tables: UK*. ONS; 2024. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables [accessed 5th December 2024].
81. Welton AJ, Vickers MR, Kim J, Ford D, Lawton BA, MacLennan AH, et al. Health related quality of life after combined hormone replacement therapy: randomised controlled trial. *BMJ* 2008;**337**:a1190.
82. Rowen D, Brazier J, Wong R, Wailoo A. *Measuring and valuing health-related quality of life when sufficient EQ-5D data is not available. Report by the Decision Support Unit*. Sheffield: Decision Support Unit, ScHARR, University of Sheffield; 2020.
83. British National Formulary (BNF). *Fezolinetant. Medicinal Forms*. Available from: https://bnf.nice.org.uk/drugs/fezolinetant/medicinal-forms/ [accessed June 2024].
84. Astellas Data on File. UK Individual Expert Validation Meetings Report. 2024.
85. Astellas Data on File. PAG Engagement Meetings: Fezolinetant (Veoza™) HTA appraisals.
86. Personal Social Services Research Unit. *New unit costs of health and social care programme*. PSSRU; Available from: https://www.pssru.ac.uk/unitcostsreport/ [accessed 5th December 2024].
87. National Health Service (NHS). *2021/22 National Cost Collection Data Publication.* Available from: https://www.england.nhs.uk/publication/2021-22-national-cost-collection-data-publication/ [accessed 6 June 2024].
88. Douxfils J, Beaudart C, Dogne JM. Letter to the Editor: efficacy and safety of fezolinetant for vasomotor symptoms in postmenopausal women: a systematic review and meta-analysis of randomized controlled trials. *Int J Gynaecol Obstet* 2024;**166**:905-6.
89. Astellas Data on File. Model QC Checklist.
90. Büyükkaramikli NC, Rutten-van Mölken MPMH, Severens JL, Al M. TECH-VER: A Verification Checklist to Reduce Errors in Models and Improve Their Credibility. *Pharmacoeconomics* 2019;**37**:1391-408.

APPENDIX 1. BAYESIAN META-ANALYSIS OF THE DOUXFILS DATA

A Bayesian pairwise fixed (common) effect (FE) meta-analysis on the Douxfils data[45] was carried out in WinBUGS 1.4.3 using a complementary log-log link[46] to allow for the different follow up periods. A random effects model was not considered, as the number of studies was low and fit for the fixed effect model was good.

The fixed effect code WinBUGS used was as follows (adapted from Dias et al[46, 66])

```
model{
for(i in 1:ns){
mu[i] ~ dnorm(0,.0001)
for (k in 1:2) {
r[i,k] ~ dbin(p[i,k],n[i,k])
cloglog(p[i,k]) <- log(time[i]) + mu[i] + d[k]
rhat[i,k] <- p[i,k] * n[i,k]
dev[i,k] <- 2 * (r[i,k] * (log(r[i,k])-log(rhat[i,k]))  +  (n[i,k]-r[i,k]) *
(log(n[i,k]-r[i,k]) - log(n[i,k]-rhat[i,k])))}
resdev[i] <- sum(dev[i,1:2])}
totresdev <- sum(resdev[])
d[1]<-0
d[2] ~ dnorm(0,.0001)
HR<-exp(d[k])
prob.more<-step(d[k])
}
```

Data were as given below:

ns= number of studies

list(ns=3)

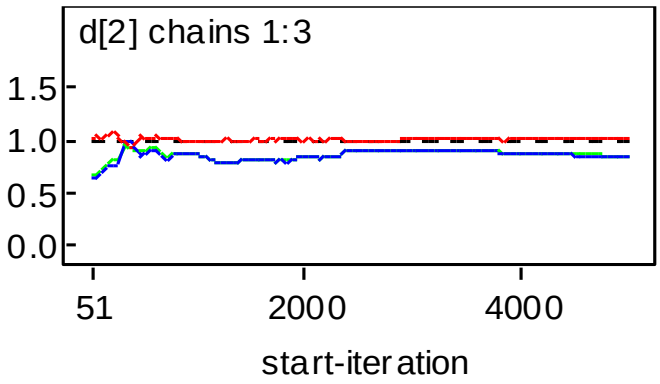
time[]	r[,1]	n[,1]	r[,2]	n[,2]
12	0	175	3	173
12	0	167	2	167
52	2	610	9	609

END

Initial values for each of the 3 chains were as follows:

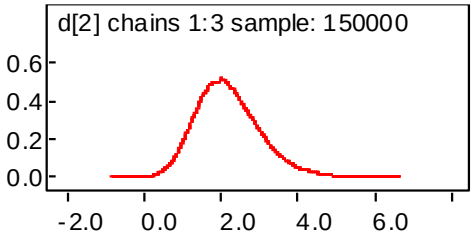
#chain 1
list(d=c(NA,0), mu=c(0,0,0))
#chain 2
list(d=c(NA,0.4), mu=c(.1,.1,.1))
#chain 3
list(d=c(NA,-0.3), mu=c(0.5,0.5,.1))

A burn in of 10,000 was carried out, which led to satisfactory convergence, as shown in the Gelman-Rubin plot below. 50,000 simulations were then carried out for each chain.



Results

A density plot for the posterior distribution of the ln HR (d[2]) is given below:



The node statistics were as follows:

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
HR	14.13	24.6	0.3493	2.177	8.218	61.36	10001	150000
d[2]	2.195	0.8508	0.01013	0.778	2.106	4.117	10001	150000
dev[1,1]	0.7562	0.7096	0.005382	0.05529	0.5463	2.657	10001	150000
dev[1,2]	1.13	1.574	0.004068	0.001131	0.521	5.586	10001	150000
dev[2,1]	0.4973	0.5236	0.003558	0.02481	0.3324	1.933	10001	150000
dev[2,2]	1.132	1.593	0.004161	0.001101	0.5222	5.633	10001	150000
dev[3,1]	1.398	1.823	0.02211	0.001493	0.7014	6.552	10001	150000
dev[3,2]	0.9571	1.357	0.003552	9.456E-4	0.4359	4.827	10001	150000
prob.more	0.9995	0.02129	6.447E-5	1.0	1.0	1.0	10001	150000
totresdev	5.871	2.993	0.01782	2.152	5.191	13.44	10001	150000

The median total residual deviance was 5.19, which shows that the DE model was a good fit to the data, given that this was less than the number of unconstrained datapoints (3 studies and 2 treatments = 6 datapoints).

APPENDIX 2. SYSTEMATIC LITERATURE REVIEWS – COST EFFECTIVENESS, HRQOL, HEALTHCARE COSTS AND RESOURCE USE

Appendix G of the CS included the company searches to identify: i) economic evaluations of fezolinetant or non-hormonal comparator therapies for VMS in menopause, ii) HRQoL studies

reporting utility values for VMS in menopause and iii) healthcare cost and resource use studies of VMS in menopause.

The original database searches were carried out in June 2023 with further searching of supplementary sources in August and September 2023. An update search was undertaken in April 2024.

The EAG found that the searches reported in Appendix G were well reported, good quality and, for the most part, appropriate. Some minor weaknesses were noted: brand names for fezolinetant were not included in the strategies, search terms for postmenopause and perimenopause were missing from the strategies in some of the supplementary sources, and the database EconLit was not searched.

The EAG appraisal of the searches can be found in Table 39

Table 39 ERG appraisal of evidence identification for SLR of cost-effectiveness, SLR of HRQoL and SLR of cost and healthcare resource use

Topic	ERG response	Note
Is the report of the search clear and comprehensive?	YES	There was an error at line 2 of the MEDLINE strategy, table 50, pg. 140, Appendix G. The company confirmed that this was a copying error and provided a corrected version of this line in their response to the PfCs.
Were appropriate sources searched?	YES	A good range of relevant databases, conference proceedings, UK HTA websites and reference checking of relevant studies was undertaken. EconLit might have been another source that could have been searched for completeness, along with other non-UK HTA websites.
Was the timespan of the searches appropriate?	YES	Databases: inception to 9th April 2024 Supplementary databases: inception to 26th April 2024 Conference proceedings: January 2021 to April 2024 HTA websites: to 26th April 2024
Were appropriate parts of the PICOS included in the search strategies?	YES	**MEDLINE and Embase** SLR cost-effectiveness: [P] VMS in menopause AND ([I] fezolinetant OR [C] non-hormonal therapies) AND [S] economic evaluations SLR HRQoL: [P] VMS in menopause AND HSUV SLR healthcare cost and resource use: [P] VMS in menopause AND direct healthcare costs and resource use **NHS EED, INAHTA** [P] VMS in menopause **CEA Registry, ScHaRRHUD, EQ5D database, conference proceedings, HTA websites** [P] VMS OR menopause. (n.b.Menopause conference proceedings were searched appropriately using broad terms for costs, economics and HSUVs.)

Were appropriate search terms used?	PARTLY	Missing brand names veoza (UK and Europe) and also veozah (US). Search terms for postmenopause and perimenopause were missing in some of the manual conference proceeding searches, HTA website searches and in the supplementary database searches (CEA Registry, ScHARRHUD and EQ-5D Publications Database).
Were any search restrictions applied appropriate?	YES	Date limit of 2013 onwards applied to the searches for HRQoL and healthcare cost and resource use sections of the strategies.
Were any search filters used validated and referenced?	PARTLY	A bespoke search filter rather than a single previously validated search filter was used to limit retrieval to economic evaluations in MEDLINE and Embase. The bespoke filter was based upon terms from published SIGN, CADTH and NHS EED economic filters. It is unclear how well this bespoke filter would have performed for the retrieval of all relevant economic evaluations, but terms appeared to be comprehensive. Similarly, a bespoke filter was created to limit retrieval to studies reporting on HSUV based upon terms from HSUV filters produced by ScHARR and YHEC. The terms included in the bespoke filter appeared to be comprehensive and appropriate for retrieval of HSUV studies.

ERG response = YES/NO/PARTLY/UNCLEAR/NOT APPLICABLE

Single Technology Appraisal

Fezolinetant for treating vasomotor symptoms associated with the menopause [ID5071]

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 Monday 16th December 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 [redacted] should be highlighted in turquoise and all information submitted as '[redacted]' in pink.

Issue 1 SKYLIGHT 1 and SKYLIGHT 2 primary endpoints

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 15, Key Issue 5. The "Description of issue and why the EAG has identified it as important" row states: "the primary endpoint of DAYLIGHT and SKYLIGHT 1 and 2 was frequency of moderate to severe VMS, a key secondary endpoint of the trials was severity of VMS." This is inaccurate. The DAYLIGHT primary endpoint was frequency of moderate to severe VMS and the key secondary endpoint was severity of moderate to severe VMS. While SKYLIGHT 1 and SKYLIGHT 2 used co-primary endpoints of frequency of moderate to severe VMS and severity of moderate to severe VMS, and the key secondary endpoint was PROMIS SD SF 8b total score.	Please amend to: "the primary endpoint of DAYLIGHT **was frequency of moderate to severe VMS and the key secondary endpoint was severity of moderate to severe VMS. The co-primary endpoints of** SKYLIGHT 1 and SKYLIGHT 2 were frequency of moderate to severe VMS **and severity of moderate to severe VMS, and the key secondary endpoint was PROMIS SD SF 8b total score**".	Important amendment to improve accuracy of the report.	The description of Issue 5 has been revised to note that severity of moderate to severe VMS was a key secondary endpoint of DAYLIGHT and a co-primary endpoint of SKYLIGHT 1 and SKYLIGHT 2.

Page 68, Section 4.2.2.1 “Points for critique” states: “Although the primary endpoint of DAYLIGHT and SKYLIGHT 1 and 2 was frequency of moderate to severe VMS, the EAG notes that a key secondary endpoint of the trials was severity of moderate to severe VMS” This is inaccurate. The DAYLIGHT primary endpoint was frequency of moderate to severe VMS and the key secondary endpoint was severity of moderate to severe VMS. While SKYLIGHT 1 and SKYLIGHT 2 used co-primary endpoints of frequency of moderate to severe VMS and severity of moderate to severe VMS, and the key secondary endpoint was PROMIS SD SF 8b total score.	Please amend to this statement as above to clarify that the primary endpoint of DAYLIGHT was frequency of moderate to severe VMS and the key secondary endpoint was severity of moderate to severe VMS. The co-primary endpoints of SKYLIGHT 1 and SKYLIGHT 2 were frequency of moderate to severe VMS and severity of moderate to severe VMS, and the key secondary endpoint was PROMIS SD SF 8b total score	Important amendment to improve accuracy of the report.	The text on page 68, Section 4.2.2.1 “Points for critique” has also been revised to state “Although the primary endpoint of DAYLIGHT was frequency of moderate to severe VMS, the EAG notes that a key secondary endpoint of DAYLIGHT was severity of moderate to severe VMS, while the co-primary endpoints of SKYLIGHT 1 and SKYLIGHT 2 were frequency of moderate to severe VMS and severity of moderate to severe VMS.”

Issue 2 Definition of severity of VMS and severity of moderate to severe VMS

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 15, Key Issue 5. The "Description of issue and why the EAG has identified it as important" row states: "In the model, VMS frequency represents an important part of the definition of VMS severity, which is not part of the definition used by the US Food and Drug Administration (FDA) and European Medicines Agency (EMA)." This is inaccurate, as in the model VMS frequency reflects the frequency of *moderate to severe VMS* in line with the trial outcomes used in DAYLIGHT and SKYLIGHT 1 and SKYLIGHT 2. Additionally, the definition of moderate to severe VMS were in line with the regulatory guidance from the FDA on the conduct of clinical trials for VMS, which defines mild, moderate and severe VMS as follows:	Please amend to: "…**VMS frequency represents the frequency of moderate to severe VMS, which is defined in line with the definition for moderate and severe VMS outlined by the US Food and Drug Administration (FDA) and European Medicines Agency (EMA).**" Please delete this: "~~**The company stated that seven moderate to severe VMS per day was used to define VMS severity to be moderate to severe**~~"	There is a misunderstanding about how the clinical trials for fezolinetant defined moderate and severe VMS. As such, this is an important amendment to improve accuracy of the report.	In the CS, it was stated that the frequency of a minimum average of 7 moderate to severe VMS was used to define the health states in the economic model. Please see for example: • "Given that all people enter DAYLIGHT with a minimum average of 7 moderate to severe VMS per day, it was necessary to use this as **a threshold for moderate to severe VMS frequency health state definitions** to ensure that most people in the trial who improve from baseline, are modelled to experience a

• Mild: sensation of heat without sweating • Moderate: sensation of heat with sweating, able to continue activity • Severe: sensation of heat with sweating, causing cessation of activity Further, the EAG stated that "The company stated that seven moderate to severe VMS per day was used to define VMS severity to be moderate to severe." This is inaccurate, as the Company never made the claim that seven moderate to severe VMS per day was used to define VMS severity to be moderate to severe. Instead, the inclusion criteria of seven moderate to severe VMS was driven by regulatory guidance from the FDA on the conduct of clinical trials for VMS.			change of health state in the model" • "To select plausible **cut-offs to define health states**, first, clinically **meaningful values were assessed along with pre-existing trial definitions**. **7 is a cut-off** used in clinical trials to define moderate to severe VMS […] These statements and the fact that the frequency of a minimum average of 7 moderate to severe VMS was used to define health states indicated that this frequency was part of the definition of moderate to severe VMS used to define the model structure. While these statements included in the CS and the fact that the fixed cut-off frequency of 7 moderate to severe VMS was used to define the health states

			used in the model makes it very unclear. We thank the company for now providing clarification at FAC. We have amended the EAR accordingly. The amended text is as follow: “In the model, VMS frequency was used to define the structure of the model. A frequency of seven moderate to severe VMS per day represents the baseline cut-off to start receiving fezolinetant in the trials. However, in the model it is assumed that postmenopausal people who experience a decrease in the frequency of moderate to severe VMS after treatment keep receiving the treatment, even if they experience only 0 to <2 moderate to severe VMS per day. The EAG is concerned that the model structure is not adequately capturing the impact of fezolinetant on

			the severity of moderate to severe VMS.”
Page 15, Key Issue 6. The “Description of issue and why the EAG has identified it as important” row states: “The company assumed that the frequency of seven moderate to severe VMS per day used in the trial eligibility criteria was required to define VMS as moderate to severe and used this as a fixed cut-off threshold. However, the definition of VMS severity used in the trials did not include any requirement of frequency of VMS.” This is inaccurate. As noted in Issue 2 above, the Company has not claimed that seven moderate to severe moderate to severe VMS per day was used to define VMS as moderate to severe VMS. Consistent with the FDA guidance, participants needed to have at least seven moderate to severe VMS at baseline for the indication of	Please amend to: “**In line with the definition for moderate and severe VMS outlined by the US FDA and EMA guidance on clinical trials for VMS, VMS frequency health states in the model represents the frequency of moderate to severe VMS. The seven moderate to severe VMS per day threshold reflects the inclusion criterion of DAYLIGHT, which was mandated by guidance from FDA on clinical trials for VMS**.”	There is a misunderstanding about how the clinical trials for fezolinetant defined moderate and severe VMS. As such, this is an important amendment to improve accuracy of the report.	Please see our response to the comment above regarding this issue. We have amended the text to the following: “The company assumed a frequency of seven moderate to severe VMS per day, which was used in the trial eligibility criteria, as a fixed cut-off threshold to define model health states. The EAG notes that the utility values for frequency between 7 to < 8 and 8 to < 9 are higher than the utility values between 6 to < 7. […]”

treatment of moderate to severe VMS.			
Page 64, Section 4.2.2.1 states: “In the analysis, VMS frequency of seven per day and above represents the severity of moderate to severe VMS, which was based on the DAYLIGHT eligibility criteria (i.e., a minimum average of seven moderate to severe VMS per day, within the last 10 days prior to randomization).” This is inaccurate. The Company has not claimed that seven moderate to severe moderate to severe VMS per day was used to define VMS as moderate to severe VMS. Consistent with the FDA guidance, participants needed to have at least seven moderate to severe VMS at baseline for the indication of treatment of moderate to severe VMS	Please amend to: “In the analysis, VMS frequency of seven per day and above ~~**represents the severity of moderate to severe VMS, which**~~ was based on the DAYLIGHT eligibility criteria (i.e., a minimum average of seven moderate to severe VMS per day, within the last 10 days prior to randomization).”	There is a misunderstanding about how the clinical trials for fezolinetant defined moderate and severe VMS. As such, this is an important amendment to improve accuracy of the report.	Please see our response to the comment above regarding this issue. We have amended the text as suggested by the company.
Page 67, Section 4.2.2.1 “Points for critique” states: “In NHS clinical practice, VMS frequency is not commonly	Please clarify in the report that the company does not use frequency of	There is a misunderstanding about how the clinical trials for fezolinetant defined moderate and severe VMS. As such, this is	Please see our response to the comment above regarding this issue. We have amended the text to

used to define moderate to severe VMS” In the company submission, frequency of moderate to severe VMS is not used to define moderate to severe VMS.	moderate to severe VMS to define moderate to severe VMS.	an important amendment to improve accuracy of the report.	remove the statement that the frequency was used to define the severity of VMS.
Page 67, Section 4.2.2.1 “Points for critique” states: “In the model, VMS frequency represents an important part of the definition of VMS severity. More specifically, the company stated that seven moderate to severe VMS per day was used to define VMS severity to be moderate to severe.” In the company submission, frequency of moderate to severe VMS is not used to define moderate to severe VMS.	Please clarify in the report that the company does not use frequency of moderate to severe VMS to define moderate to severe VMS.	There is a misunderstanding about how the clinical trials for fezolinetant defined moderate and severe VMS. As such, this is an important amendment to improve accuracy of the report.	Please see our response to the comment above regarding this issue. We have amended the text to remove the statement that the frequency was used to define the severity of VMS.
Page 67, Section 4.2.2.1 “Points for critique” states: “It remains unclear how the company defines moderate to severe VMS and how the use of moderate to severe VMS	Please clarify in the report that the company does not use frequency of moderate to severe VMS to define moderate to severe VMS.	There is a misunderstanding about how the clinical trials for fezolinetant defined moderate and severe VMS. As such, this is an important amendment to improve accuracy of the report.	Please see our response to the comment above regarding this issue. We have amended the text to remove the statement that the frequency was used to

frequency reflects the changes in VMS severity” This is inaccurate, as in the model VMS frequency reflects the frequency of *moderate to severe VMS* in line with the trial outcomes used in DAYLIGHT and SKYLIGHT 1 and SKYLIGHT 2. Additionally, the definition of moderate to severe VMS were in line with the regulatory guidance from the FDA on the conduct of clinical trials for VMS.			define the severity of VMS.
Page 68, Section 4.2.2.1 “The company assumed that the frequency of seven moderate to severe VMS per day used in the trial eligibility criteria was required to define VMS as moderate to severe VMS and used this as a fixed cut-off threshold” This is inaccurate, as in the model VMS frequency reflects the frequency of moderate to severe VMS in line with the trial outcomes used in DAYLIGHT and SKYLIGHT 1 and SKYLIGHT 2. Additionally, the	Please clarify in the report that the company does not use frequency of moderate to severe VMS to define moderate to severe VMS.	There is a misunderstanding about how the clinical trials for fezolinetant defined moderate and severe VMS. As such, this is an important amendment to improve accuracy of the report.	Please see our response to the comment above regarding this issue. We have amended the text to remove the statement that the frequency was used to define the severity of VMS.

definition of moderate to severe VMS were in line with the regulatory guidance from the FDA on the conduct of clinical trials for VMS.			
Page 69, Section 4.2.2.1, states: "The definition of moderate to severe VMS in DAYLIGHT is consistent with EMA and FDA guidance (i.e., moderate: sensation of heat with sweating, able to continue activity; severe: sensation of heat with sweating, causing cessation of activity) but eligibility criteria for DAYLIGHT required participants to have a minimum average of seven moderate to severe VMS per day, within the last 10 days prior to randomization." This sentence does not differentiate between the definition of moderate to severe VMS events in DAYLIGHT trial and the definition of moderate to severe VMS as a condition as a whole.	Please add an additional statement which provides this justification: "The definition of **mild**, moderate ~~**to**~~ **and** severe VMS **episodes** in DAYLIGHT is consistent with EMA and FDA guidance (i.e., moderate: sensation of heat with sweating, able to continue activity; severe: sensation of heat with sweating, causing cessation of activity) but eligibility criteria for DAYLIGHT required participants to have a minimum average of seven moderate to severe VMS per day, within the last 10 days prior to randomization."	To clearly distinguish between how the severity of VMS events are defined and how the severity of VMS as a condition is defined.	This was in Section 4.2.3.1, and was amended as suggested by the company.

Page 69, Section 4.2.2.1 states: "The company assumed that the frequency of seven moderate to severe VMS per day used in the trial eligibility criteria was required to define VMS as moderate to severe VMS and used this as a fixed cut-off threshold… The EAG notes that the company did not analyse other cut-off thresholds for defining moderate to severe VMS other than seven per day" This is inaccurate, as in the model VMS frequency reflects the frequency of *moderate to severe VMS* in line with the trial outcomes used in DAYLIGHT and SKYLIGHT 1 and SKYLIGHT 2. Additionally, the definition of moderate to severe VMS were in line with the regulatory guidance from the FDA on the conduct of clinical trials for VMS.	Please clarify this in the report	There is a misunderstanding about how the clinical trials for fezolinetant defined moderate and severe VMS. As such, this is an important amendment to improve accuracy of the report.	Please see our response to the comment above regarding this issue. We have amended the text to remove the statement that the frequency was used to define the severity of VMS.
Page 70, Section 4.2.3.1 states: "Importantly, the evidence used in the model to assess the cost-effectiveness	Please amend to: "Importantly, the evidence used in the model to assess the cost-effectiveness of	This is an important amendment to improve the accuracy of the report.	The text is amended as suggested by the company.

of fezolinetant is for a DAYLIGHT eligible population and therefore a moderate to severe VMS frequency limit of greater than seven per day is considered in the target population of the decision problem" The company has not defined the target population based on the frequency of moderate to severe VMS. The use of the 7 or greater moderate to severe VMS in the inclusion criteria was driven by FDA guidance.	fezolinetant is for a DAYLIGHT eligible population and therefore **based on a population with a frequency of seven or more** moderate to severe VMS ~~***frequency limit of greater than or seven***~~ per day ~~**is considered in the target population of the decision problem**~~"		
Page 82, Section 4.2.6.2 states: "…but that the SEE exercise did not appear to define moderate to severe VMS in terms of the baseline VMS frequency distribution used in the model with a minimum average of seven moderate to severe VMS per day." The company has not used frequency of moderate to severe VMS to define moderate to severe VMS	Please delete this statement: "…~~**but that the SEE exercise did not appear to define moderate to severe VMS in terms of the baseline VMS frequency distribution used in the model with a minimum average of seven moderate to severe VMS per day.**~~"	This is an important amendment to improve the accuracy of the report.	The text is amended to state that the SEE exercise was not anchored on the baseline VMS frequency distribution used in the model, which was confirmed by the company in response to EAG clarifications.

neither in the company submission or the SEE study.			
Page 105, Section 6.1.1.1 states: "…or the company's baseline VMS frequency distribution derived from DAYLIGHT is representative of a very high frequency of moderate to severe VMS and where participants are likely to have had…" Describing 7 or more moderate to severe VMS episodes per day as "very high" appears to be an overstatement and may not align with clinical understanding	Please amend to "…or the company's baseline VMS frequency distribution derived from DAYLIGHT is representative of ~~**very high**~~ a population with ~~**frequency of**~~ 7 or greater moderate to severe VMS and where participants are likely to have had…"	This is an important amendment to improve the accuracy of the report.	The text is amended as suggested by the company.
Page 117, Section 6.4 states: "The company states that seven moderate-to-severe VMS events per day was used to define VMS severity to be moderate to severe in the model because…" The company has not used frequency of moderate to severe VMS to define moderate to severe VMS in this submission.	Please delete this statement.	This is an important amendment to improve the accuracy of the report.	Please see our response to the comment above regarding this issue. We have amended the text to remove the statement that the frequency was used to define the severity of VMS. The amended text is as follows: "Another key concern is in relation to the use of moderate to severe VMS

			frequency as a proxy for VMS severity in the model. The company states that this assumption was made in the model because: 1) it is aligned with the DAYLIGHT and SKYLIGHT trials; 2) calculations incorporating both frequency and severity of moderate to severe VMS would not be possible for a probabilistic analysis; and 3) correlation between frequency and severity from the trials is demonstrated.”

Issue 3 Company’s stance on the current stand of care treatment

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 23, Section 2.2.4.2 states: “…the company believes that people for whom HRT is unsuitable, or people who have stopped HRT for any reason, should receive no active treatment.”	Please amend to: “…the company believes that people for whom HRT is unsuitable, or people who have stopped HRT for any reason, ~~**should**~~ **would** receive no active treatment.”	The company have not claimed that menopausal people who are HRT-unsuitable *should* receive no active treatment. Rather, based on extensive feedback from UK clinical experts, it is understood that most menopausal people who are	The amendment has been made.

		deemed unsuitable for HRT are currently receiving no active treatment. This reflects the limitations of existing non-hormonal pharmacological options, including limited or conflicting efficacy evidence and high discontinuation rates due to adverse events. Therefore, the wording (i.e., 'should') must be revised to avoid suggesting that the company is being prescriptive in its stance.	

Issue 4 Evidence base for non-hormonal treatments and relevance to the UK setting

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 23, Section 2.2.5 states: "The company also notes that there is a lack of robust evidence supporting these treatments. However, the EAG disagrees with this assertion. All the non-hormonal pharmacological treatments have been tested in randomised controlled trials (RCTs).[25] Reviews of such RCTs[26-29] have generally	The EAG should reflect the limitations in the applicability of the evidence to UK clinical practice. Specifically, the commentary should acknowledge that the dosages and formulations investigated in RCTs often do not correspond to those typically prescribed in the UK for VMS.	While these treatments have been evaluated in randomised controlled trials (RCTs), the dosages and formulations studied often differ from those typically used in UK clinical practice for VMS. For example, paroxetine 7.5 mg, as tested in trials, is not available in the UK and it is uncertain whether it can be assumed to be equivalent to the 10 mg dose prescribed in	P23 and p29: not a factual inaccuracy. The EAG report accurately describes the evidence for these trials. The EAG description of these trials has framed positive comments cautiously. The NICE scope did not specify dosages or formulations of the comparators.

demonstrated some efficacy over placebo." Page 29, Table 4 "Comparator(s) row states: "All the non-hormonal pharmacological treatments have been tested in RCTs which have generally demonstrated a benefit over placebo." The EAG's assertion that all non-hormonal pharmacological treatments have robust evidence supporting their efficacy is misleading.		the UK, and gabapentin studies frequently assessed extended-release formulations at 1200 or 1800 mg, which exceed the commonly prescribed 900 mg dose. These discrepancies undermine the direct applicability of the trial findings to NHS clinical practice, and the lack of alignment between trial conditions and real-world prescribing practices should have been critically assessed by the EAG.	

Issue 5 Applicability of fezolinetant RCT evidence to perimenopausal population

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 28, Section 2.3 "Population" subheading states: "The company also confirmed that both these groups are covered by the marketing authorisation. However, the company also stated that the fezolinetant RCT evidence did not cover	Please amend to: "The company also confirmed that both these groups are covered by the marketing authorisation**, as the EMA's Committee for Medicinal Products for Human Use deemed the indication wording to be approvable, finding no major clinical efficacy or safety concerns in the**	It is important to include an amendment that clearly directs readers to the section where the rationale for the EMA's decision not to restrict the licensed wording for fezolinetant to only postmenopause is explained.	Not a factual inaccuracy. It remains true that the RCT evidence did not cover the perimenopausal population, and it is reasonable for the EAG to state that this may affect the applicability of trial

the perimenopausal population. This may affect the applicability of the trials' results to the NHS population, and is a potential source of uncertainty"	**postmenopausal population studied in the pivotal Phase 3 trials for fezolinetant (Section 3.2.1.2).** However, the company also stated that the fezolinetant RCT evidence..."		results to people who are perimenopausal.

Issue 6 Decision problem population

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 29, Table 4, "Population" row states: "Menopausal people who are HRT-averse have been excluded from the targeted population, as it is expected that HRT remains the relevant comparator for these people." As per discussion under the 'Population' subheading in Section 2.3, please clarify that the decision problem population includes HRT-averse	Please amend to: "Menopausal people who are HRT-averse have been excluded from the targeted population in the company submission, as it is expected that HRT remains the relevant comparator for these people. **However, following misunderstanding at scoping which was resolved at clarification, the decision problem population was updated to include HRT-averse"**	Important amendment to ensure Table 4 aligns with the discussion above the table.	This has been amended as suggested.

Issue 7 Databases searched in the clinical SLR

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 32, Table 5 "Were appropriate sources searched?" states: "The International Health Technology Assessment database (INAHTA) was not searched in the original or update searches. Epistemonikos or KSR Evidence (databases of systematic reviews) were not searched in either the original or update searches." INAHTA was not searched in the clinical SLR as this is an economic database. Additionally, it is inaccurate that no databases of systematic reviews were searched as the Cochrane database of systematic reviews (CDSR) was searched in the clinical SLR.	Please amend the "EAG response" column to "YES" and the "EAG Notes" column to: "The International Health Technology Assessment database (INAHTA) was not searched in the original or update searches, **as this is an economic database**. ~~**Epistemonikos or KSR Evidence (databases of systematic reviews) were not searched in either the original or update searches.**~~ **The Cochrane database of systematic reviews (CDSR) was searched in the clinical SLR**".	Important amendment to improve accuracy of report	The proposed amended text in bold is not entirely correct. The International Health Technology Assessment database (INAHTA) contains Health Technology Assessments that cover both clinical effectiveness as well as cost-effectiveness assessments of interventions so is a valid source to search for clinical effectiveness evidence. The proposed deletion of text is not applicable. A search of Epistemonikos or KSR Evidence is necessary to identify previous non-Cochrane systematic reviews. The proposed text "**The Cochrane database of**

			systematic reviews (CDSR) was searched in the clinical SLR" is acceptable and has been added. For the reasons outlined above the "EAG response" column to the question Were appropriate sources searched? should remain as "PARTLY"

Issue 8 Population included in the clinical SLR

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 32, "Population" subheading states: "3. The decision problem refers to 'post-menopausal people', which was clarified by the company to mean both perimenopausal and postmenopausal people. The SLR restricts the population to 'post-menopausal individuals'." Search terms for peri-menopausal were included in the search strategies although	Please amend to: "The SLR ~~**restricts the population**~~ **to includes both 'peri- and** post-menopausal individuals'.	Important amendment to improve accuracy of report	Not a factual inaccuracy. The company SLR protocol clearly states that the population for inclusion were post-menopausal (p39, Table 12, CS appendices). Regardless of search terms used, it is possible that a systematic reviewer would respond to the wording of the protocol by only including studies

Astellas acknowledge that it is not clear from the eligibility criteria that peri-menopausal women were also eligible for inclusion			involving post-menopausal people.

Issue 9 HRQoL outcome was captured in the economic SLR

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 36, Section 3.1.2 "Outcomes" states: "…However, the SLR does not contain the important decision problem outcomes of sleep disturbance or health-related quality of life (HRQoL)." For completeness, the EAG should acknowledge or signpost the reader to the economic SLR where HRQoL outcomes were captured.	Please amend to: "…However, the SLR does not contain the important decision problem outcomes of sleep disturbance or health-related quality of life (HRQoL). **That said, the economic SLR identified HRQoL studies reporting utility values for VMS in menopause (Appendix 2)**"	Important amendment to improve accuracy of report	Not a factual inaccuracy. The clinical SLR did not contain these outcomes, and it is the clinical SLR that was being described.

Issue 10 Estimand method and approaches to handling of missing data

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Pages 37, Section 3.2.1.1 "Risk of bias" states: "…The EAG	The following amendments must be made to ensure the EAG report is accurate,	The inclusion of this additional evidence would provide a	Not a factual inaccuracy.

identified two issues with missing data, described below, which it believes place the results of all three trials at high risk of bias (due to missing data) with respect to continuous outcomes (this includes the VMS frequency and VMS severity outcomes)" The EAG's assessment of the SKYLIGHT 1 and SKYLIGHT 2 trials as being at high risk of bias due to missing data does not fully account for additional sensitivity analyses that have been conducted. Specifically, the company performed a tipping-point analysis to evaluate sensitivity to potential violations in the missing at random assumption used in the MMRM analysis for the primary endpoints, which the company acknowledge the EAG may have missed, that was used to support regulatory approvals. This analysis demonstrated that the assumption of data being missing at random was justifiable and that the missing data did not impact the efficacy outcome assessments.	balanced, and reflective of all available evidence: • Include acknowledgment of the tipping-point analysis conducted by the company, which demonstrated that the assumption of data being missing at random (MAR) was justifiable and that the missing data did not impact the efficacy outcome assessments for SKYLIGHT 1 and SKYLIGHT 2. • State explicitly that the tipping point analysis was performed on the full population, not just on the subgroup of patients who discontinued due to treatment-emergent adverse events (TEAEs). • Change the classification of DAYLIGHT, SKYLIGHT 1 and SKYLIGHT 2 from "high risk of bias due to missing data" to "low risk of bias"	more complete and accurate assessment of the trials' risk of bias due to missing data.	The EAG does not think that the tipping point analysis "demonstrated that the assumption of data being missing at random was justifiable and that the missing data did not impact the efficacy outcome assessments." Tipping point analyses relate to the amount of shift needed to make results non-significant. While this may be useful for regulatory purposes it is not so useful when evaluating cost-effectiveness. Moreover, the methods details provided on this analysis is limited. For example, it is stated (in the SKYLIGHT 1 and SKYLIGHT 2 Integrated Summary of Efficacy_2 document) that "*subjects in the Fezolinetant treated groups who fell into a missing not at random pattern had k*

Additionally, addressing the concerns of the EAG, the tipping-point analysis was conducted in the full population, not solely in the subgroup of patients who discontinued due to TEAEs. The results of this analysis provide robust evidence supporting the reliability of the continuous outcomes (e.g., VMS frequency and severity) and should be considered in the EAG's evaluation. Furthermore, the point around the estimand strategy is likely moot, as it appears the EAG's real concerns centre around the use of a hypothetical strategy leading to higher rates of missing data. However, this concern relies on the assumption that the MAR approach is biased—a claim that the tipping point analysis clearly refutes. The methodology and results of the tipping-point analysis were provided as part of the reference pack for original submission. Please refer to:			*times the treatment differences obtained from the primary MMRM analysis subtracted from the imputed values after the dropout time*." But it is unclear both how patients were categorised as having a missing not at random pattern, and how many patients were classified as having data missing not at random.

• SKYLIGHT 1 and SKYLIGHT 2 Integrated Summary of Efficacy_1
  ○ Section 1.3.4.2 Pooled Efficacy Analyses Methods
• SKYLIGHT 1 and SKYLIGHT 2 Integrated Summary of Efficacy_2
  ○ Tables 8.4.3.1.1, 8.4.3.1.2, 8.4.3.2.1 and 8.4.3.2.2.
  ○ Table 8.4.3.3 Sensitivity Analysis Overview Table: Tipping Point
 	
Lastly, given that DAYLIGHT used a treatment policy estimand which includes patients who discontinued treatment, missing data is less of a concern here.			

Issue 11 Clinically important thresholds for outcomes in DAYLIGHT, SKYLIGHT 1 and 2

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 45, section 3.2.2.5 states: “The differences between groups in either the	Please amend to: “The differences between groups in either the total score or the VMS domain did not reach this	Important amendment to improve accuracy of report	Amended to: “The company suggested an MCID value of 2.0 for

total score and VMS domain did not reach this threshold, so clinical significance cannot be assumed." This is inaccurate. The minimal clinically important difference provided by the company was for within-group comparison; as such, it should not be used to assess between-group differences.	threshold~~, so clinical significance cannot be assumed.~~ **However, it is important to note that the MCID analysis provided by the company was specifically designed for within-group comparisons. This analysis is not intended to assess between-group differences, and conclusions about clinical significance at the group comparison level cannot be drawn from it**"		the vasomotor domain, based on **a within-group** anchor-based study of data from SKYLIGHT 1 and 2.[44] The differences between groups in either the total score and VMS domain did not reach this **(within-group)** threshold, so clinical significance cannot be assumed."

Issue 12 Efficacy data used in the model

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 43, section 3.2.2.2 states: "Also, the efficacy data used in the modelling were observed data which did not take account of any biases arising from missing trial data." This is statement is misleading, as DAYLIGTH efficacy data (i.e., transition probabilities) in the model did include patients who discontinued treatment.	Please update this statement to acknowledge that DAYLIGHT efficacy data included patients who discontinued treatment.	Important amendment to improve accuracy of report	Not a factual inaccuracy. The statement relates to missing data, not to patients who discontinued treatment. 22% of fezolinetant 45mg patients in DAYLIGHT had missing data at the primary endpoint (Table 7 in EAR).

Issue 13 DAYLIGHT trial period

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 47 Section 3.2.3 states: "A major limitation of this meta-analysis was the pooling of studies with different follow up periods (the double blind period of the SKYLIGHT 4 trial was 52 weeks, but only 12 weeks for the DAYLIGHT trials)." This is inaccurate, as the double-blind placebo period for DAYLIGHT was 24 weeks	Please amend to: "…(the double blind period of the SKYLIGHT 4 trial was 52 weeks, but only **24** ~~**12**~~ weeks for the DAYLIGHT trials)."	Minor amendment to improve accuracy of report	We had incorrectly written 'DAYLIGHT' rather than the 'SKYLIGHT 1 and 2 trials'. This has been amended.

Issue 14 Liver test elevations and liver monitoring in the model

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 14, Section 1.5 Issue 4 "What is the expected effect on the cost-effectiveness estimates?" row states: "Expected to have minimal impact on cost-effectiveness, but additional costs associated with liver function tests and	Please amend to: "Expected to have minimal impact on cost-effectiveness, **as supported by a scenario provided by the company in the updated model following clarification questions which included costs associated with liver function tests for patients receiving fezolinetant** ~~**but additional costs**~~	Minor amendment to improve accuracy of report.	The text is amended.

monitoring may have been excluded." It is important to clarify that a scenario that includes liver monitoring was provided at clarification.	**~~associated with liver function tests and monitoring may have been excluded~~**."		
Page 50, Section 3.2.3 "Liver test elevations" states that: "The FDA warning, as well as the data provided by the CS on liver test elevations and special liver tests, do not match the company's implication that liver adverse events are unimportant". This is inaccurate. Given that the risk of liver test elevations was very similar between the placebo and fezolinetant 45 mg arms (Table 13 of EAG report), it was considered reasonable to not model additional liver monitoring for fezolinetant in the company submission	Please delete this statement: "**~~The FDA warning, as well as the data provided by the CS on liver test elevations and special liver tests, do not match the company's implication that liver adverse events are unimportant~~**"	The current phrasing misrepresents the company's position. The company does not consider liver adverse events to be unimportant. Rather, the decision not to model additional liver monitoring in the submission was based on RCT evidence showing that the risk of liver test elevations was very similar between the placebo and fezolinetant 45 mg arms (as outlined in Table 13 of the EAG report).	We have amended our statement to: *"The FDA warning, as well as the data provided by the CS on liver test elevations and special liver tests, indicate that this may be important.*"

Issue 15 Desvenlafaxine as a relevant comparator to represent SNRI as a class

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 53 Section 3.3.1 states: "… Nevertheless, the EAG thinks that it should have been included in the NMA, to represent SNRIs as a class."	Please remove this statement.	Including desvenlafaxine in the NMA is unreasonable, as it is not available as a treatment option in the UK. While the EAG argues that licence status should not preclude its inclusion and suggests it could represent SNRIs as a class, it is important to recognize that desvenlafaxine has not been licensed for any indication in the UK. This lack of licensing makes it effectively unavailable for clinical use within the NHS, unlike other treatments that, while unlicensed for VMS, are at least accessible for off-label prescribing. Additionally, the EAG have provided no evidence to suggest that a class effect exists. Lastly, were a submitting pharmaceutical company to make a similar claim to that made by the EAG without any evidence to substantiate it, it would not be accepted in the NICE process.	Not a factual inaccuracy. This represents the EAG's opinion. However, this has been amended for clarity: "… Nevertheless, the EAG considers that it **could** have been included in the NMA, to represent SNRIs as a class, **along with evidence supporting a class effect**".

Issue 16 NMA results of Morga et al 2023

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 57, section 3.5 “Results of Morga et al 2023” states: “The EAG concludes that the results from this NMA indicate that there is little evidence to suggest that fezolinetant 45mg has a clinically-meaningful benefit over SSRIs, SNRIs and gabapentin for reducing the frequency of moderate to severe vasomotor symptoms, and that there is no evidence to suggest any benefit for reducing the severity of vasomotor symptoms. Although the EAG also noted some methodological and reporting limitations with the NMA, its results are somewhat uncertain, primarily due to the high risk of bias due to missing outcome data seen across trials.” The conclusion drawn by the EAG appears to be unreasonable, given the limited and biased evidence	Please amend to: “The EAG acknowledges that the results from this NMA provide limited evidence to suggest that fezolinetant 45 mg has a clinically meaningful benefit over SSRIs, SNRIs, and gabapentin for reducing the frequency of moderate to severe vasomotor symptoms. However, this conclusion is heavily influenced by the limited and biased evidence available for many comparators included in the analysis, as well as the complete lack of evidence for other treatments sometimes used in clinical practice. Additionally, the high risk of bias due to missing outcome data across trials and the methodological and reporting limitations of the NMA significantly increase the uncertainty of these results. Acknowledging these limitations, the conclusions should be interpreted with caution rather than definitively implying a lack of efficacy for fezolinetant”	To improve clarity and avoid overgeneralisation, the EAG must specify the exact drugs and doses used in their conclusions rather than making broad generalisations about entire drug classes. Additionally, it would be helpful for the EAG to explicitly acknowledge in their conclusions that the comparator data were both limited and of questionable quality, which significantly impacts the reliability of the NMA results. This would ensure a more balanced and accurate interpretation of the findings.	Not a factual inaccuracy. The NMA limitations have been adequately covered in the EAR.

for the comparators included in the NMA and the complete lack of evidence for many comparators likely to be used in clinical practice. It would be more appropriate for the EAG to provide a fairer comment on the lack of robust data, rather than implying that fezolinetant may not be effective based on such limited and biased comparator evidence.			

Issue 17 Reporting of Institute for Clinical and Economic Review (ICER) model results

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 59, Section 4.1, "Table 55 of the CS summarises the cost-effectiveness model, patient population, quality-adjusted life years (QALYs), costs, and comparative cost-effectiveness of fezolinetant compared to no pharmacologic treatment. The resulting incremental cost per QALY gained was $390,000." It is important to note that this analysis was done using a	Please amend to: "Table 55 of the CS summarises the cost-effectiveness model, patient population, quality-adjusted life years (QALYs), costs, and comparative cost-effectiveness of fezolinetant compared to no pharmacologic treatment. The resulting incremental cost per QALY gained was $390,000, **based on a placeholder price for fezolinetant of $6,000 per annum**. **The UK price per annum would be significantly lower (approximately £584).**"	It is important to provide context regarding the high ICER result from The Institute for Clinical and Economic Review's model.	Not a factual inaccuracy. The EAG summarized the information provided in the report published by the Institute for Clinical and Economic Review (ICER). The UK price, which is included in the company's suggestions for revisions, is not included in that report.

placeholder cost for fezolinetant of $6,000, which is significantly higher than the cost of fezolinetant used in the CS.			

Issue 18 EAG comment on the source of treatment effect used in the company submission

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 60, Section 4.1 "Points for critique" states that: "In the CS, the treatment effect only included a reduction in moderate to severe VMS frequency per day from DAYLIGHT up to week 24, pooled SKYLIGHT 1 and 2 from week 24-52 and a constant average based on pooled SKYLIGHT 1 and 2 from week 52 onwards for fezolinetant, while the placebo effect from DAYLIGHT was applied to week 12 only." The phrasing "only included" understates the complexity and granularity of the	Given the granularity of the model, it is more accurate to highlight the comprehensive nature of the data inputs and their role in driving these granular predictions, rather than implying a simplistic or limited approach used in the company's model.	It is important to provide context of the important differences between the simplistic ICER model and the more granular model developed by the company.	Not a factual inaccuracy. The EAG is not implying that one approach is more appropriate than another in this text. Instead, this text is highlighting the differences between the ICER model and the company's model. The specifics of the company's model are explored in the sections that follow. We removed the word 'only' in the text to avoid any misinterpretation.

company model. This model incorporates probabilities of transitioning between health states, including four moderate to severe VMS frequency health states; thus, offering a more detailed and realistic projection of long-term costs and health outcomes. By capturing these dynamic changes, a Markov cohort model enhances the robustness and reliability of the cost-effectiveness analysis for fezolinetant. Moreover, the model incorporates an extensive dataset, including RCT data from DAYLIGHT and pooled SKYLIGHT 1 and 2 trials.			

Issue 19 Treatment discontinuation for fezolinetant and no active treatment used in the model

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 60, Section 4.1 “Points for critique” states that: “The company’s model applied treatment-specific discontinuation rates to all	Please amend to: “The company’s model applied treatment-specific discontinuation rates to all model cycles for fezolinetant **over the full model**	Minor amendment to improve accuracy of report.	Text amended as suggested by the company to improve clarity.

model cycles for fezolinetant and for no active treatment until week 12.”	**duration** and for no active treatment until week 12”		

Issue 20 Fezolinetant impact on proportion of women in lower VMS frequency health states

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 61, Section 4.2, states “Fezolinetant is modelled to affect QALYs by increasing the proportion of people who transition from a high moderate to severe VMS frequency of ≥ 7 events per day to low moderate to severe VMS frequencies of between 0 and 2 and between 2 and 7 events per day…” Fezolinetant also increases the proportion of people who transition from moderate to severe VMS frequency of ≥9 to moderate to severe VMS frequency of between 7 and less than 9	Please amend to: “Fezolinetant is modelled to affect QALYs by increasing the proportion of people who transition from a high moderate to severe VMS frequency of ≥ 7 events per day to low moderate to severe VMS frequencies of between 0 and 2**, ~~and~~** between 2 and 7 **and between 7 and 9** events per day…”	Minor amendment to reflect the results of the model.	We have amended the text to the following: “Fezolinetant is modelled to affect QALYs by increasing the proportion of people who transition from a high moderate to severe VMS frequency of ≥ 7 events per day to low moderate to severe VMS frequencies of between 0 and 2, between 2 and 7 and between 7 and 9 (from frequency of ≥ 9) events per day […]”
Page 61, Section 4.2, states “Fezolinetant is modelled to affect costs by having higher	Please amend to: “Fezolinetant is modelled to affect costs by having higher drug acquisition costs	Minor amendment to reflect the results of the model.	Text has been amended as suggested by the company.

drug acquisition costs compared to no active treatment but a lower percentage of menopausal people in the low moderate to severe VMS frequency health states, …" Fezolinetant is more effective than no treatment in reducing VMS and therefore results in a higher proportion of menopausal people in the low moderate to severe VMS frequency health states.	compared to no active treatment but a ~~lower~~ **higher** percentage of menopausal people in the low moderate to severe VMS frequency health states, …"		

Issue 21 Rationale for moderate to severe VMS frequency cut-off thresholds used in the model

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 63, Section 4.2.2.1 states: "These cut-off thresholds were validated by UK clinical experts for clinical plausibility. The detailed rationale for choice of each cut-off is provided in CS Section B.3.2.2.1 and in an appendix to the company's supplementary material,	Please amend to: "These cut-off thresholds were validated by UK clinical experts for clinical plausibility. The detailed rationale for choice of each cut-off is provided in CS Section B.3.2.2.1 and in **~~an appendix to the company's supplementary material,~~ the appendix of the** Structured Expert Elicitation (SEE) **report**[75]"	Minor amendment to improve clarity of the report.	Text has been amended as suggested by the company.

Structured Expert Elicitation (SEE)[75]" For clarity, please refer to this document as the SEE report.			
Page 68, Section 4.2.2.1 states: "The EAG assessed the information about GEE model 7, which was provided in the appendix to the supplementary material, SEE." For clarity, please refer to this document as the SEE report.	Please amend to: "The EAG assessed the information about GEE model 7, which was provided in the appendix **~~to the supplementary material,~~ of the** SEE **report**."	Minor amendment to improve clarity of the report.	Text has been amended as suggested by the company.
Page 69, Section 4.2.2.1 states:" The EAG assessed the information about GEE model 7, which was provided in the appendix to the supplementary material, SEE." For clarity, please refer to this document as the SEE report.	Please amend to: "The EAG assessed the information about GEE model 7, which was provided in the appendix **~~to the supplementary material,~~ of the** SEE **report**."	Minor amendment to improve accuracy of report.	This is the same as the previous comment.
Page 95, Section 4.2.8.1 states: "As described in Section 4.2.2.1, the appendix to the supplementary material, SEE provides some information on the utility values and on how	Please amend to: "As described in Section 4.2.2.1, the appendix **~~to~~ of** the **~~supplementary material,~~** SEE **report** provides some information on the utility values and on how they were used to define the health states in the model"	Minor amendment to improve clarity of the report.	Text has been amended as suggested by the company.

they were used to define the health states in the model” For clarity, please refer to this document as the SEE report.			

Issue 22 Frequency of moderate to severe VMS as a proxy for severity of moderate to severe VMS

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 67, Section 4.2.2.1 “Points for critique” states: “The EAG has several concerns about the model structure. The primary concern relates to the use of VMS frequency as a proxy for VMS severity” And later states: “More specifically, the company stated that seven moderate to severe VMS per day was used to define VMS severity to be moderate to severe.” These statements are inaccurate. Frequency of moderate to severe VMS is not used as a proxy for	Please delete these statements.	To provide clarity regarding the modelling of moderate to severe VMS frequency.	The first statement is correct as VMS frequency was used in the model as a proxy for VMS severity. This is described in the CS section B.3.2.2.1. The second statement was addressed in response to the comments regarding Issue 2 above.

severity of moderate to severe VMS, it is considered to implicitly capture elements of severity based on the correlation between frequency of moderate to severe VMS and severity of moderate to severe VMS in the clinical trials.			
Page 67, Section 4.2.2.1 "Points for critique" states: "Clinical advice to the EAG expressed concern about using VMS frequency as a proxy for VMS severity because a menopausal person with two severe night sweats in a 24-hour period, requiring a change of bed sheets and clothing, is not comparable to someone with a frequency of greater than seven per day where the severity of symptoms of each event are mild." This is inaccurate. Frequency of moderate to severe VMS is not used as a proxy for severity of moderate to severe VMS, it is considered to implicitly capture elements of	Please delete this statement: "**~~Clinical advice to the EAG expressed concern about using VMS frequency as a proxy for VMS severity because a menopausal person with two severe night sweats in a 24-hour period, requiring a change of bed sheets and clothing, is not comparable to someone with a frequency of greater than seven per day where the severity of symptoms of each event are mild.~~**"	This statement is not relevant as the company do not consider mild events in the model at all.	Not a factual inaccuracy. VMS frequency was used as a proxy for VMS severity in the model structure. This is described in the CS section B.3.2.2.1

severity based on the correlation between frequency of moderate to severe VMS and severity of moderate to severe VMS in the clinical trials.			
Page 117, Section 6.4 states: "Another key concern is in relation to the use of moderate to severe VMS frequency as a proxy for VMS severity in the model." This is inaccurate. Frequency of moderate to severe VMS is not used as a proxy for severity of moderate to severe VMS, it is considered to implicitly capture elements of severity based on the correlation between frequency of moderate to severe VMS and severity of moderate to severe VMS in the clinical trials.	Please delete this statement: "~~**Another key concern is in relation to the use of moderate to severe VMS frequency as a proxy for VMS severity in the model**~~ ."	To provide clarity regarding the modelling of moderate to severe VMS frequency.	Not a factual inaccuracy. VMS frequency was used in the model to define health states. Thus, it represents a proxy for VMS severity. This is described in the CS section B.3.2.2.1. The company stated that some clinical experts suggested capturing VMS severity alongside VMS frequency in the model and the Company considered that approach but decided to use only VMS frequency for model structure.

Issue 23 Inclusion criteria of DAYLIGHT and SKYLIGHT 1 and 2

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 67, Section 4.2.2.1 "Points for critique" states: "However, this requirement was to increase the chance of observing a change in symptoms from baseline between the arms of the trials." This is inaccurate. The inclusion criterion was mandated by the FDA and was not used for the purpose quoted above	Please delete this statement: "~~**However, this requirement was to increase the chance of observing a change in symptoms from baseline between the arms of the trials.**~~"	This is an important amendment to improve the accuracy of the report.	Text is amended as suggested by the company.

Issue 24 Modelling of treatment discontinuation

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 67, Section 4.2.2.1, "…, it is assumed in the model that women who experience a decrease in the frequency of moderate to severe VMS after treatment keep receiving the	Please amend this statement to make clear the biological plausibility of this assumption in light of the mechanism of action of fezolinetant and UK clinical opinion	This is an important amendment to improve the accuracy of the report.	The text is amended to make it clear that treatment discontinuation is considered independent of VMS frequency health states.

treatment, even if they experience only 0 to <2 moderate to severe VMS per day." Treatment discontinuation is considered independently of VMS frequency health state occupation in the model. Treatment discontinuation rates are based on discontinuation data from the DAYLIGHT and SKYLIGHT trials and are applied to people 'on treatment' across all VMS frequency health states in the model. As such women in the 0 to <2 may discontinue treatment in line with the discontinuation rates from the DAYLIGHT/SKYLIGHT trials. Additionally, the assumption that women who experience a decrease in the frequency of moderate to severe VMS continue treatment aligns with the mechanism of action of fezolinetant. This treatment addresses the underlying imbalance in oestrogen signalling in the brain region responsible for VMS. As a result, if treatment is stopped before natural cessation of VMS, it is expected that symptom frequency will return			

immediately to pre-treatment levels. This assumption was validated by UK clinicians.			

Issue 25 Correlation between frequency and severity of moderate to severe VMS

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 68, Section 4.2.2.1, "The company further justified their approach with a correlation analysis of frequency and severity from the trials, but no details were provided on the type of correlation analysis and only results for SKYLIGHT 1 and SKYLIGHT 2 were provided despite the fact that DAYLIGHT was used to define the health states in the model and estimate the health state utility values." Further justification for the use of VMS frequency to define health states was provided in the original	Please amend to: "The company further justified their approach **1) by noting that a combined frequency and severity 'scoring scale' would lack credibility due to lack of validation and absence from common clinical practice 2)** with a correlation analysis of frequency and severity from the trials, but no details were provided on the type of correlation analysis and only results for SKYLIGHT	Important amendment to provide additional clarity regarding the justification for using VMS frequency to define health states.	Text has been amended as suggested by the company.

company submission, namely limitations regarding post hoc construction of a composite endpoint based on the available data from phase III trials. Combining data for these two endpoints into a combined frequency and severity 'scoring scale' would lack credibility due to lack of validation and absence from common clinical practice.			

Issue 26 EQ-5D utility values

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 69, Section 4.2.2.1 states: "In general, the EAG is also concerned about the use of utility values to define VMS frequency health states because the EQ-5D utility values are likely to reflect changes in quality of life due to other menopause symptoms over and above VMS."	Please delete this statement: "~~**In general, the EAG is also concerned about the use of utility values to define VMS frequency health states because the EQ-5D utility values are likely to reflect changes in quality of life due to other menopause symptoms over and above VMS.**~~"	This is an important amendment to improve the accuracy of the report.	Not a factual inaccuracy. This is the EAG's opinion.

This is inaccurate. The analysis of utility values based on moderate to severe VMS frequency directly addresses this concern. Due to fezolinetant specifically targeting VMS, no improvements in other menopause symptoms are expected. Therefore, any observed utility improvements would be directly attributable to changes in the frequency of moderate to severe VMS. This relationship refutes the claim that utility values reflect quality-of-life changes from other menopause symptoms beyond VMS.			
Page 96, Section 4.2.8.1 states “Consequently, the company decided to adjust these values based on the opinion of one clinical expert.” This is inaccurate. The adjustments were based on two clinical expert opinions, two changes were made each justified by one clinical expert.	Please amend to: “Consequently, the company decided to adjust these values based on the opinion**s of two clinical experts, which adjustment validated by** one clinical expert **and further validated by 4 patient experts.**”	This is an important amendment to improve the accuracy of the report.	Text has been amended to the following: “Consequently, the company decided to adjust these values based on the opinion of two clinical experts. This adjustment was validated by one clinical expert and 4 patient experts.”

Page 109, Section 6.1.1.5 states: “Thus, the company decided to adjust these values based on the opinion of one clinical expert.” This is inaccurate. The adjustments were based on two clinical expert opinions, two changes were made each justified by one clinical expert.	Please amend to: “Thus, the company decided to adjust these values based on the **opinions of two clinical experts, which adjustment validated by** one clinical expert **and further validated by 4 patient experts.**”	This is an important amendment to improve the accuracy of the report.	Text has been amended to the following: “Thus, the company decided to adjust these values based on the opinion of two clinical experts. This adjustment was validated by one clinical expert and 4 patient experts.”
Page 119, Section 6.4: “More importantly, the EAG considers that the adjustment of the utility values with the opinion of one clinical expert is not an appropriate approach.” This is inaccurate. The adjustments were based on two clinical expert opinions, two changes were made each justified by one clinical expert.	Please amend to acknowledge that the company decided to adjust these values based on the opinions of two clinical experts, which adjustment validated by one clinical expert.	This is an important amendment to improve the accuracy of the report	Text has been amended to the following: “More importantly, the EAG considers that the adjustment of the utility values with the opinion of two clinical experts is not an appropriate approach.”

Issue 27 Treatment effect from exploratory NMA conducted by the company

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 74, Section 4.2.4.1: "… (ii) the singular treatment effect estimate from the exploratory NMA had to be mapped to the four-by-four transition probability matrices (VMS frequency categories) at different 4-weekly cycles creating a significant source of uncertainty, since the model structure did not allow adequate use of a relative treatment effect from the NMA" The current phrasing could be reworded to reflect the inherent limitations of using an NMA to inform a granular model structure. Specifically, the NMA inherently assumes that a treatment effect can be summarised in a singular value of relative treatment effect, which is inadequate for capturing the nuanced	Please amend to: "An NMA cannot provide adequate outputs to inform the company's granular model structure, as the NMA generates a singular treatment effect estimate, which must then be mapped to the four-by-four transition probability matrices. This creates a significant source of uncertainty"	This is an important amendment to improve the accuracy of the report	This is not a factual inaccuracy. The company chose a model structure that did not allow incorporation of a singular treatment effect from a NMA.

transitions required in the company's model			

Issue 28 Time horizon

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 75, Section 4.2.5.1 states: "In addition, although the median treatment duration for fezolinetant is modelled to be approximately 2.4 years, a proportion of the cohort is still receiving fezolinetant by year 10." This statement requires further clarification to enhance transparency and allow for a more accurate interpretation of treatment persistence and its impact on model outcomes.	Please amend to: "In addition, although the median treatment duration for fezolinetant is modelled to be approximately 2.4 years, a **very small** proportion of the cohort is still receiving fezolinetant by year 10"	This is an important amendment to improve the accuracy of the report	Text has been amended as suggested by the company.

Issue 29 Treatment effectiveness

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 77, Section 4.2.6.1: "Both outcomes showed a statistically significant LS mean difference for fezolinetant versus placebo, but the relative difference between treatments was smaller for the outcome of mean weighted severity compared to frequency… The EAG notes that a much smaller relative mean difference of 0.26 (95% CrI, 0.15 to 0.37) was estimated for the comparison of fezolinetant 45 mg versus placebo for reduction in VMS severity…" The statement comparing the relative difference in mean weighted severity to frequency is factually inaccurate because the two outcomes are not on the same scale, making direct comparison invalid.	The phrasing must be revised to avoid misleading implications and instead focus on the significance of the outcomes within their respective scales.	This is an important amendment to improve the accuracy of the report	Text amended to state: "Both outcomes showed a statistically significant LS mean difference for fezolinetant versus placebo…. ~~The EAG notes that a much smaller relative~~ and a mean difference of 0.26 (95% CrI, 0.15 to 0.37) ~~was estimated~~ for the comparison of fezolinetant 45 mg versus placebo for reduction in VMS severity…"

Issue 30 Pooled DAYLIGHT and SKYLIGHT 1 and 2 Scenario

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 79, Section 4.2.6.1: "…The company have presented a scenario analysis that uses pooled efficacy data from DAYLIGHT and SKYLIGHT 1 and 2 (HRT-unsuitable) [based on naïve pooling – see Section 3.2.2], which is demonstrated to have a material impact on the incremental cost-effectiveness ratio (ICER) for fezolinetant relative to no active treatment (see Table 76, p181 of CS)… It is important to note that the 'Pooled DAYLIGHT and SKYLIGHT 1 and 2' scenario referenced here also changes a number of other settings to use data from DAYLIGHT and SKYLIGHT 1 and 2 (HRT-unsuitable) beyond efficacy being discussed in this section. If only the efficacy is changed there is <u>not</u> a material impact on the	The phrasing must be revised to avoid misleading suggestions of what the 'Pooled DAYLIGHT and SKYLIGHT 1 and 2' scenario presented by the company represents. Please include the other model settings that are changed in this scenario: • Utility values - pooled DAYLIGHT & SKYLIGHT 1 & 2 (HRT-unsuitable) analysis (<u>un</u>adjusted by clinicians) • Baseline patient distribution pooled DAYLIGHT & SKYLIGHT (HRT-unsuitable) • Baseline age pooled DAYLIGHT & SKYLIGHT (HRT-unsuitable) • Fezolinetant efficacy week 0–12 (pooled DAYLIGHT & SKYLIGHT (HRT-unsuitable) 0–12 per cycle) • 'No active treatment' efficacy week 0–12 (pooled DAYLIGHT & SKYLIGHT (HRT-unsuitable) 0–12 per cycle) • Fezolinetant treatment discontinuation pooled DAYLIGHT & SKYLIGHT (HRT-unsuitable) week 0–24	This is an important amendment to improve the accuracy of the report	The text referring to this scenario analysis is removed to avoid any misunderstanding.

incremental cost-effectiveness ratio (ICER).	• 'No active treatment' discontinuation pooled DAYLIGHT & SKYLIGHT (HRT-unsuitable) week 0–12 •		

Issue 31 UK clinical experts' estimates of natural history of moderate to severe VMS for women receiving no active treatment

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 82, Section 4.2.6.2 states: "As noted previously, the EAG believes that the baseline distribution should be closer to the elicited values at one year for the natural history of postmenopausal women in the UK because such a rapid reduction in the frequency of moderate to severe VMS between baseline and one year in the absence of treatment would suggest a lack of need for treatments for moderate to severe VMS because the symptoms would resolve rapidly by themselves over time."	Please amend to: "As noted previously, the EAG believes that the baseline distribution should be closer to the elicited values at one year for the natural history of postmenopausal women **receiving no active treatment for moderate to severe VMS** in the UK…"	This is an important amendment to improve the accuracy of the report	The EAG has amended the text to the following: ". As noted previously, the EAG believes that the baseline distribution should be closer to the elicited values at one year for the natural history of postmenopausal women in the UK who do not receive any active treatment for moderate to severe VMS because such a rapid reduction in the frequency of moderate to severe VMS between baseline and one year in the absence of treatment would suggest a lack of

It is important to note that the derived values were specifically for women receiving no active treatment for moderate to severe VMS.			need for treatments for moderate to severe VMS because the symptoms would resolve rapidly by themselves over time."
Page 82, Section 4.2.6.2 states: "This was also confirmed by the clinical experts informing the company who supported the view that such a significant reduction in symptoms is unlikely to occur within the first year." It is important to clarify what "This" is here as this sentence could easily be misinterpreted.	Please amend to: "~~**This**~~ **The implausibility of the rapid resolution of moderate to severe VMS at Year 1** was also confirmed by the clinical experts informing the company who supported the view that such a significant reduction in symptoms is unlikely to occur within the first year."	Minor amendment to clarify the accuracy of the report	The EAG has amended the text to the following: "The implausibility of the significant reduction in symptoms occurring within the first year was also confirmed by the clinical experts informing the company."
Page 82, Section 4.2.6.2 states: "…or that the baseline VMS frequency distribution is not representative of moderate to severe VMS observed in UK clinical practice in the absence of treatment. " It is unclear how the use of Year 6 SEE estimates only suggests that the baseline VMS frequency distribution is not representative of moderate to severe VMS observed in UK	Please amend this statement to a provide clear justification for this stance.	This is an important amendment to improve the accuracy of the report	Not a factual inaccuracy. The sentence is referring to the company discarding the year 1 and year 3 SEE estimates and only using the year 6 estimates, which implies that either the SEE estimates are not reliable (because the company does not trust the estimates derived from year 1 and year 3) or the baseline VMS frequency

clinical practice in the absence of treatment.			distribution is not representative of UK practice (because the baseline distribution does not align with the elicited year 1 estimates).

Issue 32 Principles followed for the SEE study

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 83 Section 4.2.6.2 states: "The bisection method was used to determine uncertain judgements. The bisection approach uses a variable interval method; 1%, 99%, 50%, 25% and 75% percentiles. This is an alternative to the fixed interval methods. The STEER guidance concludes that either approach is appropriate for SEE to inform health care decision making." The placement of the quoted text in the section entitled "Several observations can be made regarding the methods employed in the SEE and the	Please incorporate this methodology as part of the following text: "Several principles, suggested by the STEER framework, were adhered to in the company's elicitation exercise. Specifically: only substantive experts were recruited using pre-specified recruitment and selection criteria, the experts were trained prior to the task, including awareness on forms of bias, a SEE protocol was developed, and experts were frequently asked to reflect on and revise their answers as needed following each quantity elicited. **The bisection method was used to determine uncertain judgements. The bisection approach uses a variable interval method; 1%, 99%, 50%, 25% and 75% percentiles. This is an alternative to the fixed**	This is an important amendment to improve the accuracy of the report	This is not a factual inaccuracy. We state clearly that "The STEER guidance concludes that either approach is appropriate for SEE to inform health care decision making."

likely credibility of the resulting elicited estimates" is inappropriate, given that placement of the text here implies that the bisection method is considered inappropriate, which is not the case as this method is line with the STEER guidance.	**interval methods. The STEER guidance concludes that either approach is appropriate for SEE to inform health care decision making**"		

Issue 33 Uncertainty of original SEE estimates

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 84, Section 4.2.6.2 states: "The company states that discussion with experts suggested that the high uncertainty captured in the estimates is representative of the highly variable experience of menopausal people with VMS and reflects the counterfactual nature of trying to estimate the VMS frequency in a cohort that would rarely exist in reality because clinicians would not leave people experiencing moderate to severe VMS untreated. The EAG notes that this suggests	Please amend the text to provide further clarification: "The company states that discussions with experts suggested that the high uncertainty captured in the estimates is representative of the highly variable experience of menopausal people with VMS and reflects the counterfactual challenge nature of trying to estimate the VMS frequency in a cohort that would rarely exist in reality because clinicians would not leave people experiencing moderate to severe VMS untreated. **However, clinicians also acknowledged that many menopausal people deemed unsuitable for HRT are not currently**	The interpretation that "no active treatment is not the relevant comparator" misrepresents the feedback from clinicians. The feedback indicates that while patients with moderate to severe VMS who are unsuitable for HRT may not be receiving any active treatment, this does not imply that clinicians would intentionally leave them untreated. Instead, it reflects the reality of limited effective non-hormonal treatment options available in practice. The comparator of no active treatment remains relevant, as it aligns with the	Text amended to include the additional wording suggested by the company. The EAG's opinion is not deleted.

that no active treatment is not the relevant comparator in this target population who are unsuitable for HRT because such a population rarely exists in NHS practice" This is a misinterpretation of the feedback from clinicians.	**receiving active treatment due to the limitations of available options. ~~The EAG notes that this suggests that no active treatment is not the relevant comparator in this target population who are unsuitable for HRT because such a population rarely exists in NHS practice.~~** This highlights the relevance of no active treatment as the sole relevant comparator, as it reflects the current treatment landscape and unmet needs within this population."	current treatment landscape for this target population. This distinction should be clarified to avoid mischaracterising clinical practice and undermining the relevance of the chosen comparator.	

Issue 34 Company model results

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 85, Section 4.2.6.2 states "The cost-effectiveness results for the scenario analyses demonstrate that the alternative options have a material impact on the ICER for fezolinetant relative to no active treatment (see Table 76, p182 of CS), which highlights how sensitive the cost-effectiveness results are to the natural history assumed in the model."	The EAG should provide a quantitative description of the impact to clarify the extent of sensitivity in the cost-effectiveness results and improve transparency.	This is an important amendment to improve the transparency of the report	This is not a factual inaccuracy. The EAG refers to Table 76, p182 of CS and the reader can make their own judgement on the impact. The use of the term 'material impact' implies that the ICER is changed by a large amount compared to the company's base case. This section of

The use of the term “material impact” is unnecessarily vague.			the EAR is not presenting cost-effectiveness results.

Issue 35 Median duration of VMS (SWAN study)

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 90, Section 4.2.6.4 states: “The study reported that younger age was a risk factor for duration of VMS, including showing a difference in the risk of longer VMS duration for ages 50 to 54.9 compared to age ≥ 55 years, which suggests that the assumption of a constant rate over time as menopausal people age in the model may not be appropriate.” It is important to specify that the study identified younger age at first report of VMS as a risk factor for longer VMS duration, rather than just younger age in general.	Please amend to: “The study reported that younger age **at first report of VMS** was a risk factor for duration of VMS, including showing a difference in the risk of longer VMS duration for ages 50 to 54.9 compared to age ≥ 55 years, which suggests that the assumption of a constant rate over time as menopausal people age in the model may not be appropriate.”	Minor amendment to improve the accuracy of the report	Text amended as suggested by the company.

Issue 36 Search filters used in the economic SLR

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 128, Appendix 2 “Were any search filters used validated and referenced?” states that “”			No description or amendment of problem given.

References

Description of problem	Description of proposed amendment	Company response	EAG Response
Page 21, Section 2.2.2 Burden of disease states: “For example, the company states that VMS causes a decrease in self-care in 81% of patients, citing Makara-Studzinska (2014),[17] but perusal of the article does not reveal this finding.”	N/A	The company acknowledges the EAG’s comment. Makara-Studzinska 2014 was incorrectly cited in this instance, instead the statement should have been referenced to English et al. Qualitative study: burden of menopause-associated vasomotor symptoms (VMS) and validation of PROMIS Sleep Disturbance and Sleep-Related Impairment measures for assessment of VMS impact on sleep. J Patient Rep Outcomes 2021 Apr 26;5(1):37.	The EAG comment on Studinska has been deleted to acknowledge this correction.

Page 21, Section 2.2.2 Burden of disease states: “Similarly, the company states how VMS causes reduced quality of life or depressed mood, but the cited studies[1, 2] were designed to demonstrate how mood and other factors affect how bothersome VMS are, and were not designed to show the effects of VMS on such factors.”	N/A	The company acknowledges the EAG’s comment. However, both papers (Hunter et al., 2012 and Thurston et al., 2008) do note that there is a relation between VMS and a negative impact on QoL. These studies, along with additional supporting references, were used to substantiate these statements.	No change required.
Page 21, Section 2.2.2 Burden of disease states: “Likewise, the review by Woods (2005)[18] is referenced as evidence that VMS causes an impact on relationships, involving less intimacy, a loss of closeness, arguments, and tension. However, whilst an effect on intimacy is supported, the other effects do not appear to be demonstrated in the article.”	N/A	The company acknowledges the EAG’s comment.	No change required.
Page 21, Section 2.2.2 Burden of disease states: “In terms of social and work-related effects, a reference to back up a statement that VMS has a negative socioeconomic	N/A	While it’s true that the referenced webpage discusses general menopause symptoms rather than exclusively focusing on VMS, it still provides relevant context that informs the broader	No change required.

impact is a webpage that focussed on general menopause symptoms, and does not specifically refer to the effects of VMS on socioeconomic variables.[19]"		socioeconomic impacts of menopause. Since VMS is one of the most common and disruptive symptoms of menopause, it is reasonable to infer that the broader socioeconomic challenges discussed in the report are, at least in part, influenced by VMS. Therefore, while the reference may not directly isolate VMS-specific socioeconomic effects, it remains a valuable resource for understanding the wider implications of menopause symptoms on social and work-related outcomes.	
Page 21, Section 2.2.2 Burden of disease states: "In the same way, Hardy, 2018[20] is cited in order to support the claim that VMS leads to reduced work performance, greater workplace stress and intention to leave the labour force. However, the survey carried out by the authors only shows a link between VMS at work and intention to stop working, whilst specifically stating that	N/A	The company acknowledges the EAG's comment.	No change required.

VMS does not affect work outcomes."			

Typographical errors

Description of problem	Description of proposed amendment	Justification for amendment	EAG Response
Page 12, Section 1.3 The decision problem: summary of the EAG's key issues, Issue 1 states: "In addition to no active treatment, other non-hormonal pharmacological treatments, including anti-depressants (such as selective serotonin reuptake inhibitors [SSRIs] and norepinephrine reuptake inhibitors [SNRIs]),…"	Please amend to "In addition to no active treatment, other non-hormonal pharmacological treatments, including anti-depressants (such as selective serotonin reuptake inhibitors [SSRIs] and **serotonin-**norepinephrine reuptake inhibitors [SNRIs]),…"	Word missing from abbreviation.	This has been amended
Page 20, Section 2.1 states: "…the clinical and cost-effectiveness of fezolinetant (Veozah®)…"	Please amend to "…**the clinical and cost-effectiveness of fezolinetant (Veoza®)**…"	Minor amendment to reflect the correct brand name for fezolinetant.	This has been amended
Page 24, Section 2.2.5 Intended positioning of fezolinetant states: "Although antidepressants and gabapentinoids are not	Please amend to "Although antidepressants and gabapentinoids are not licensed for VMS, clinical advice to the EAG confirms that the pharmacological non-hormonal treatments in the NICE	The reference supplied, Constantine et al. 2016 includes survey results from UK patients, and does not appear to specify whether all of these patients	This has been amended

licensed for VMS, clinical advice to the EAG confirms that the pharmacological non-hormonal treatments in the NICE scope (clonidine, selective serotonin reuptake inhibitors (SSRIs)/ serotonin and norepinephrine reuptake inhibitors (SNRIs) and gabapentinoids) are commonly prescribed in the NHS to about 1 in 5 people with VMS,[24] and will therefore be a more appropriate comparator than 'no active treatment' for estimations of efficacy and cost-effectiveness."	scope (clonidine, selective serotonin reuptake inhibitors (SSRIs)/ serotonin and norepinephrine reuptake inhibitors (SNRIs) and gabapentinoids) are commonly prescribed in the **UK** to about 1 in 5 people with VMS,[24] and will therefore be a more appropriate comparator than 'no active treatment' for estimations of efficacy and cost-effectiveness."	received their treatment on the NHS.	
Page 25, Section 2.3 Critique of company's definition of decision problem states: "The HRT contraindications described in the fezolinetant DAYLIGHT trial comprise undiagnosed vaginal bleeding, a history of breast cancer or oestrogen dependent tumours, arterial thromboembolic disease, venous thrombophilic disorder, hypersensitivity to oestrogen and progesterone	Please amend to "The HRT contraindications described in the fezolinetant DAYLIGHT trial comprise undiagnosed vaginal bleeding, a history of breast cancer or oestrogen dependent tumours, **liver disease, pregnancy**, arterial thromboembolic disease, venous thrombophilic disorder, hypersensitivity to oestrogen and progesterone therapy or any of the excipients or porphyria."	Liver disease and pregnancy were missing from the list of contraindications.	This has been amended

therapy or any of the excipients or porphyria."			
Page 26, Section 2.3 "Population" subheading states: "The NICE scope population includes all people with moderate or severe VMS, which equate to the licensed population, but the company restricts the decision problem population to people for whom HRT is unsuitable for medical reasons."	Please amend to "The NICE scope population includes all people with moderate or severe VMS, which equate to the licensed population, but the company restricts the decision problem population to people for whom HRT is unsuitable ~~**for medical reasons**~~."	To align with the updated decision problem population at clarification stage. Astellas does not consider people who are HRT averse to be HRT unsuitable for medical reasons, only those who are HRT-contraindicated, HRT-caution and HRT-stoppers.	This has been amended
Page 27, Section 2.3 "Population" subheading states: "The company define people unsuitable for HRT for medical reasons as those with HRT-contraindications, those with HRT cautions, those that have stopped HRT for any reason, or those who are HRT-averse..."	Please amend to "The company define people unsuitable for HRT ~~**for medical reasons**~~ as those with HRT-contraindications, those with HRT cautions, those that have stopped HRT for any reason, or those who are HRT-averse…"	To align with the updated decision problem population at clarification stage. Astellas does not consider people who are HRT averse to be HRT unsuitable for medical reasons, only those who are HRT-contraindicated, HRT-caution and HRT-stoppers.	This has been amended
Page 29, Section 2.3, Table 4, row 2, column 3 states:	Please amend to "Menopausal people with moderate to severe vasomotor-	To align with the updated decision problem population at clarification stage. Astellas does	This has been amended

"Menopausal people with moderate to severe vasomotor-predominant symptoms for whom HRT is deemed unsuitable for medical reasons:"	predominant symptoms for whom HRT is deemed ~~**for medical reasons**~~:"	not consider people who are HRT averse to be HRT unsuitable for medical reasons, only those who are HRT-contraindicated, HRT-caution and HRT-stoppers.	
Page 29, Section 2.3, Table 4, row 2, column 4 states: "Consequently, UK clinical experts expect that, in NHS primary care, fezolinetant will mainly be used in menopausal people with moderate to severe VMS for whom HRT is deemed unsuitable for medical reasons,…"	Please amend to "Consequently, UK clinical experts expect that, in NHS primary care, fezolinetant will mainly be used in menopausal people with moderate to severe VMS for whom HRT is deemed unsuitable ~~**for medical reasons**~~,…"	To align with the updated decision problem population at clarification stage. Astellas does not consider people who are HRT averse to be HRT unsuitable for medical reasons, only those who are HRT-contraindicated, HRT-caution and HRT-stoppers.	This has been amended
Page 29, Section 2.3, Table 4, row 4, column 3 states: "Menopausal people for whom HRT is not deemed suitable for medical reasons:"	Please amend to "Menopausal people for whom HRT is not deemed suitable ~~**for medical reasons**~~:"	To align with the updated decision problem population at clarification stage. Astellas does not consider people who are HRT averse to be HRT unsuitable for medical reasons, only those who are HRT-contraindicated, HRT-caution and HRT-stoppers.	This has been amended
Page 36, Section 3.1.5 Evidence synthesis states:	Please amend to "Placebo comparisons were made with desvenlafaxine (**5** studies), gabapentin (1 study), gabapentin extended release (3 studies), paroxetine	Table 19, CS Appendix D lists five studies using desvenlafaxine versus placebo: Archer 2009a, Archer 2009b,	This has been amended (with removal of repeated gabapentin study as well)

“Placebo comparisons were made with desvenlafaxine (4 studies), gabapentin (1 study), gabapentin extended release (3 studies), paroxetine (2 studies), gabapentin (1 study), clonidine (1 study) and fezolinetant (2 studies).”	(2 studies), gabapentin (1 study), clonidine (1 study) and fezolinetant (2 studies).”	Bouchard 2012, Pinkerton 2012, Speroff 2008.	
Page 42, Section X, Table 10 Responder Analyses of Change from Baseline in Frequency of Moderate to Severe VMS	Please amend the DAYLIGHT Week 24 OR (95% CI) to “[redacted]”	Rounding error.	This has been amended
Page 45, Section 3.2.2.5 Health-related quality of life states: “Higher scores represent a worse impairment (CS, p 42).”	Please amend to “Higher scores represent a worse impairment (CS, p **5**2).”	The section of the CS is reported incorrectly and should be updated.	This has been amended
Page 47, Section 3.2.3 Adverse event and safety data, Table 11 lists the following sources: “Source: Tables 43, 44, 53 of the CS. AEs were coded using MedDRA.”	Please amend to “Source: Table 43 **of the CS appendices, and Tables** 44, 53 of the CS. AEs were coded using MedDRA.”	The section of the CS is reported incorrectly and should be updated.	This has been amended

Page 49, Section 3.2.3 Adverse event and safety data states: "Although the company claims that there were no malignant neoplasms in MOONLIGHT 3, 1/150 developed endometrial adenocarcinoma, and 1/150 patients developed endometrial carcinoma."	Please amend to "Although the company claims that there were no malignant neoplasms in MOONLIGHT 3, 1/150 developed endometrial adenocarcinoma~~,~~ **~~and 1/150 patients developed endometrial carcinoma~~**."	The MOONLIGHT 3 trial only reported 1 case of endometrial adenocarcinoma, this appears to be a repetition.	This has been amended
Page 55, Section 3.5 Additional work on clinical effectiveness undertaken by the EAG states: "Studies reporting outcome definitions comparable to those used in the SKYLIGHT trials (see CS p52) were deemed appropriate for inclusion in the NMA."	Please amend to "Studies reporting outcome definitions comparable to those used in the SKYLIGHT trials (see CS p5**1**) were deemed appropriate for inclusion in the NMA."	The section of the CS is reported incorrectly and should be updated.	This has been amended
Page 60, Section 4.1 EAG comment on company's review of cost-effectiveness evidence states: "However, MENQOL includes more health domains related to VMS and symptoms associated with VMS, which	Please amend to "However, MENQOL includes more health domains related to **menopause** ~~VMS~~ and symptoms associated with **menopause** ~~VMS~~, which means the utility values were associated not only with VMS, but also other symptoms correlated with VMS."	Minor amendment to improve accuracy of the report.	This has been amended.

means the utility values were associated not only with VMS, but also other symptoms correlated with VMS."			
Page 62, Section 4.2.2 Model structure states: "The company justified the use of VMS frequency for the health states based on discussions with UK clinical experts and four patient experts who considered VMS frequency to be a more objective measure of the health effects of treatment compared to VMS severity; however, it is also stated in the CS that some clinical experts suggested that VMS severity should be incorporated and captured in the model."	Please amend to "The company justified the use of VMS frequency for the health states based on discussions with UK clinical experts and four patient experts who considered VMS frequency to be a more objective measure of the health effects of treatment compared to VMS severity; however, it is also stated in the CS that some clinical experts suggested that VMS severity should be ~~**incorporated and**~~ captured **alongside VMS frequency** in the model."	This wording should be updated to include "alongside VMS frequency" as per the CS statement in the CS.	This has been amended.
Page 64, Section 4.2.2.1, Table 16 "Mean Estimate (95%)" column	Please amend to: "Mean Estimate (95% **CI**)"	Typographical error.	This has been amended.
Page 67, Section 4.2.2.1 "Points for critique" states: "This frequency also represents the baseline cut-off	Please amend to: "This frequency ~~**also**~~ represents the baseline cut-off to start receiving fezolinetant (see Section 3.2.2.2)."	Minor amendment to improve accuracy of the report.	This has been amended.

to start receiving fezolinetant (see Section 3.2.2.2)."			
Page 68, Section 4.2.2 Model structure states: "The EAG notes that the precision estimates for both the mean frequency of moderate to severe VMS (Figure 7 in the CS, p. 74) and the mean weighted severity of VMS experienced per 24 hours (Figure 8 in the CS, p. 75) are provided."	Please amend to "The EAG notes that the precision estimates for both the mean frequency of moderate to severe VMS (Figure 7 in the CS, p. 74) and the mean weighted severity of VMS experienced per 24 hours (Figure 8 in the CS, p. 7**6**) are provided."	The section of the CS is reported incorrectly and should be updated.	This has been amended.
Page 68, Section 4.2.2.1 "item 1. Frequency of moderate to severe VMS is not commonly used to define severity of VMS in NHS clinical practice." heading	This heading should be positioned right below the "Point for critique" heading.	Minor formatting error.	Not a factual inaccuracy.
Page 69, Section 4.2.2.1 "item 2. Arbitrary cut-off thresholds are used to define moderate to severe VMS frequency health states in the model"	This heading should be positioned right above the following paragraph: "The EAG's second concern relates to the definition of the health states. The assumptions underlying the cut-off thresholds were based on statistical analysis of the distribution…"	Minor formatting error.	Not a factual inaccuracy.

Page 73, Section 4.2.3.1, “item 3. The baseline distribution with a moderate to severe VMS frequency limit of greater than seven per day from the DAYLIGHT trial is unlikely to be appropriate for assessing the need for treatment in the NHS and is unlikely to match that seen in UK clinical practice for postmenopausal people.”	This heading should be positioned right below the “Point for critique” heading.	Minor formatting error.	Not a factual inaccuracy.
Page 74, Section 4.2.5.1, “item 4. No active treatment is unlikely to be the sole relevant comparator to fezolinetant because menopausal people presenting with moderate to severe VMS and deemed HRT-unsuitable are more likely to be offered non-hormonal pharmacological treatments as opposed to ‘nothing’ in NHS clinical practice.”	This heading should be positioned right below the “Point for critique” heading.	Minor formatting error.	Not a factual inaccuracy.
Page 75, Section 4.2.6.1 Short-term efficacy from clinical trials up to week 52 states: “At the end of week 12, people in the no active treatment revert to the natural history of	Please amend to “At the end of week 12, people in the no active treatment **arm** revert to the natural history of VMS in the absence of treatment and follow the natural history transitions (Section 4.2.6.2 below).”	Minor grammatical error.	This has been amended.

VMS in the absence of treatment and follow the natural history transitions (Section 4.2.6.2 below)."			
Page 75, Section 4.2.6.1 Short-term efficacy from clinical trials up to week 52 states: "…, which is unlikely to be observed in clinical practice as people will not have additional follow-up or GP interactions associated with partaking in a clinical trial.. "	Please amend to: "…, which is unlikely to be observed in clinical practice as people will not have additional follow-up or GP interactions associated with partaking in a clinical trial."	Minor typographical error (double full stop).	This has been amended.
Page 79, Section 4.2.6.1 states: "[based on naïve pooling – see Section 3.2.2, which is demonstrated to have a material impact on the incremental cost-effectiveness ratio (ICER) for fezolinetant relative to no active treatment (see Table 76, p181 of CS)"	Please amend to: "[based on naïve pooling – see Section 3.2.2, which is demonstrated to have a material impact on the incremental cost-effectiveness ratio (ICER) for fezolinetant relative to no active treatment (see Table 76, p181 of CS)"	Formatting error.	This has been amended.
Page 80, Section 4.2.6.1 states: "item 5. The absolute changes from baseline in the frequency of moderate to severe VMS observed in the fezolinetant arm of the trials is not adjusted for the placebo	This heading should be positioned right below the "Point for critique" heading.	Formatting error.	Not a factual inaccuracy.

effect from week 12 onwards in the model."			
Page 85, Section 4.2.6.2 states: "item 6. The credibility of the SEE estimates for informing the natural history of moderate to severe VMS over time in NHS clinical practice for postmenopausal people who are deemed HRT-unsuitable is unclear."	This heading should be positioned right below the "Point for critique" heading.	Formatting error.	Not a factual inaccuracy.
Page 92, Section 4.2.7.1: "item 7. The median total VMS duration of 7.4 years that includes women pre-menopause or early perimenopause through to postmenopause does not align with the modelled population of postmenopausal people with a mean starting age of 51 years at the onset of menopause and the modelled time horizon of 10 years."	This heading should be positioned right below the "Point for critique" heading.	Formatting error.	Not a factual inaccuracy.
Page 94, Section 4.2.8.1 Summary of company submission, Table 23: Health state utility values used in the base case and scenario	Please amend to "Table 23: Health state utility values used in the base case and scenario analysis (reproduced from Table 64 of CS, p162 **and the company cost-**	This table is not reproduced from Table 64 of CS, p162. Only the first two sets of columns (UK clinician adjusted DAYLIGHT data and DAYLIGHT) were	This has been amended.

analysis (reproduced from Table 64 of CS, p162)	**effectiveness model updated in response to the EAG CQs**)"	taken from the CS. The third set of columns (Pooled DAYLIGHT & SKYLIGHT 1 & 2) were taken from the company's cost-effectiveness model. The fourth set of columns (Pooled SKYLIGHT 1 & 2) were taken from the company's updated cost-effectiveness model that was updated in response to the EAG's CQs.	
Page 95, Section 4.2.8.1 Summary of company submission states: "Utility values for all health states, which are provided in Tables 47 and 48 in the company's response to EAG clarifications, are notably higher from the SKYLIGHT 1 and 2 trials compared to DAYLIGHT ([redacted]-[redacted] vs. [redacted]-[redacted] for VMS frequency 0 to ≥9 in SKYLIGHT 1 and 2 and DAYLIGHT, respectively),…"	Please amend to "Utility values for all health states, which are provided in Tables 47 and 48 in the company's response to EAG clarifications, are notably higher from the SKYLIGHT 1 and 2 trials compared to DAYLIGHT ([redacted]-[redacted] vs. [redacted]-[redacted] for VMS frequency 0 to ≥9 in SKYLIGHT 1 and 2 and DAYLIGHT, respectively),…"	Rounding error.	This has been amended.
Page 96, Section 4.2.8.1 states: "item 8. The modelled health state utility values are highly uncertain."	This heading should be positioned right below the "Point for critique" heading.	Formatting error.	Not a factual inaccuracy.

Page 99, Section 5.1.1 Summary of company's submission, Table 26: Deterministic and probabilistic base-case results (reproduced from Table 74 and Table 71, CS).	Please amend the Probabilistic QALYs to: Fezolinetant: 6.659 QALYs No active treatment: 6.543 QALYs	Data input error.	This has been amended.
Page 100, Section 5.1.1 "Point for critique"	This heading is incorrectly positioned.	Formatting error.	Not a factual inaccuracy.
Page 119, Section 6.4 Conclusions of the cost effectiveness section states: "The EAG notes that the study reports separately a median VMS duration of 3.4 years in the subset of women who were postmenopause at the onset of VMS (20.2% of the total cohort of 3302 women), which is more reflective of the company's modelled base case population of postmenopausal people."	Please amend to "The EAG notes that the **Avis et al., (2015)** study reports separately a median VMS duration of 3.4 years in the subset of women who were postmenopause at the onset of VMS (20.2% of the total cohort of 3302 women), which is more reflective of the company's modelled base case population of postmenopausal people."	The study referred to here was not specified.	This has been amended. Although it was stated in Page 109, Section 6.1.1.4 Scenario 4: VMS duration.