Fezolinetant for Moderate to Severe Vasomotor Symptoms Associated with Menopause: Effectiveness and Value

Draft Evidence Report

October 11, 2022

Prepared for

ICER Staff and Consultants	The University of Colorado Modeling Group
Francesca L. Beaudoin, MD, PhD, MS Senior Medical Advisor Institute for Clinical and Economic Review **Abigail Wright, PhD, MSc** Senior Research Lead, Evidence Synthesis Institute for Clinical and Economic Review **Kai Yeung, PharmD, PhD** Assistant Scientific Investigator Kaiser Permanente Washington Health Research Institute **Ashton Moradi, PharmD, MS** Health Economist Institute for Clinical and Economic Review **Serina Herron-Smith, BA** Associate Research Manager Institute for Clinical and Economic Review **Steven D. Pearson, MD, MSc** President Institute for Clinical and Economic Review **David M. Rind, MD, MSc** Chief Medical Officer Institute for Clinical and Economic Review **Grace A. Lin, MD** Medical Director for Health Technology Assessment Institute for Clinical and Economic Review	**R. Brett McQueen, PhD** Assistant Professor School of Pharmacy, University of Colorado **Eric Gutierrez, MPH** Professional Research Assistant School of Pharmacy, University of Colorado *The role of the University of Colorado is limited to the development of the cost-effectiveness model, and the resulting ICER report does not necessarily represent the views of the University of Colorado.*

DATE OF PUBLICATION: October 11, 2022

How to cite this document: Beaudoin FL, McQueen RB, Wright A, Yeung K, Moradi A, Herron-Smith S, Gutierrez E, Rind DM, Pearson SD, Lin GA. Fezolinetant for Moderate to Severe Vasomotor Symptoms Associated with Menopause: Effectiveness and Value; Draft Evidence Report. Institute for Clinical and Economic Review, October 11, 2022. https://icer.org/assessment/vasomotor-symptoms-menopause-2022/#overview

Acknowledgements: Francesca L. Beaudoin served as the lead author for the report, with Kai Yeung making substantial contributions to the first draft of the report. Abigail Wright led the systematic review in collaboration with Serina Herron-Smith. R. Brett McQueen and Eric Gutierrez developed the economic model. Ashton Moradi provided oversight of the cost-effectiveness analyses and developed the budget impact model. Grace A. Lin, David M. Rind, and Steven D. Pearson provided methodologic guidance on the clinical and economic evaluations. We would also like to thank Kelsey Gosselin, Yasmine Kayali, and Shahariar Mohammed Fahim for their contributions to this report. Additionally, we would like to thank the National Coalition for LGBTQ Health for their input on gender inclusivity.

About ICER

The Institute for Clinical and Economic Review (ICER) is an independent non-profit research organization that evaluates medical evidence and convenes public deliberative bodies to help stakeholders interpret and apply evidence to improve patient outcomes and control costs. Through all its work, ICER seeks to help create a future in which collaborative efforts to move evidence into action provide the foundation for a more effective, efficient, and just health care system. More information about ICER is available at https://icer.org/.

The funding for this report comes from government grants and non-profit foundations, with the largest single funder being the Arnold Ventures. No funding for this work comes from health insurers, pharmacy benefit managers, or life science companies. ICER receives approximately 24% of its overall revenue from these health industry organizations to run a separate Policy Summit program, with funding approximately equally split between insurers/PBMs and life science companies. There are no life science companies relevant to this review who participate in this program. For a complete list of funders and for more information on ICER's support, please visit https://icer.org/who-we-are/independent-funding/.

For drug topics, in addition to receiving recommendations from the public, ICER scans publicly available information and also benefits from a collaboration with IPD Analytics, an independent organization that performs analyses of the emerging drug pipeline for a diverse group of industry stakeholders, including payers, pharmaceutical manufacturers, providers, and wholesalers. IPD provides a tailored report on the drug pipeline on a courtesy basis to ICER but does not prioritize topics for specific ICER assessments.

About the Midwest CEPAC

The Midwest Comparative Effectiveness Public Advisory Council (Midwest CEPAC) – a core program of ICER – provides a public venue in which the evidence on the effectiveness and value of health care services can be discussed with the input of all stakeholders. The Midwest CEPAC seeks to help patients, clinicians, insurers, and policymakers interpret and use evidence to improve the quality and value of health care.

The Midwest CEPAC Panel is an independent committee of medical evidence experts from across the Midwest, with a mix of practicing clinicians, methodologists, and leaders in patient engagement and advocacy. All Panel members meet strict conflict of interest guidelines and are convened to discuss the evidence summarized in ICER reports and vote on the comparative clinical effectiveness and value of medical interventions. More information about the Midwest CEPAC is available at https://icer.org/who-we-are/people/independent-appraisalcommittees/midwest-comparative-effectiveness-public-advisory-council-m-cepac/ .

The findings contained within this report are current as of the date of publication. Readers should be aware that new evidence may emerge following the publication of this report that could potentially influence the results. ICER may revisit its analyses in a formal update to this report in the future.

The economic models used in ICER reports are intended to compare the clinical outcomes, expected costs, and cost-effectiveness of different care pathways for broad groups of patients. Model results therefore represent average findings across patients and should not be presumed to represent the clinical or cost outcomes for any specific patient. In addition, data inputs to ICER models often come from clinical trials; patients in these trials may differ in real-world practice settings.

In the development of this report, ICER's researchers consulted with several clinical experts, patients, manufacturers, and other stakeholders. The following experts provided input that helped guide the ICER team as we shaped our scope and report. It is possible that expert reviewers may not have had the opportunity to review all portions of this draft report. None of these individuals is responsible for the final contents of this report, nor should it be assumed that they support any part of it. The report should be viewed as attributable solely to the ICER team and its affiliated researchers.

For a complete list of stakeholders from whom we requested input, please visit: https://icer.org/wp-content/uploads/2022/06/Menopause-Revised-Key-Stakeholders-List.pdf

Expert Reviewers

Yoko Allen, MPH
Senior Program Manager
Black Women's Health Imperative
No relevant conflicts of interest to disclose, defined as more than $10,000 in healthcare company stock or more than $5,000 in honoraria or consultancies during the previous year from health care manufacturers or insurers.

Louise Crathorne, MSc
Senior Research Fellow in Health Technology Assessment
University of Exeter
Ms. Crathorne has received salary or other payments for services such as consulting fees or honoraria in excess of $5,000 from Aesara Europe.

Deborah Grady, MD, MPH
Professor of Medicine
University of California, San Francisco
No relevant conflicts of interest to disclose, defined as more than $10,000 in healthcare company stock or more than $5,000 in honoraria or consultancies during the previous year from health care manufacturers or insurers.

Kathryn Rexrode, MD, MPH
Chief, Division of Women's Health
Brigham and Women's Hospital
No relevant conflicts of interest to disclose, defined as more than $10,000 in healthcare company stock or more than $5,000 in honoraria or consultancies during the previous year from health care manufacturers or insurers.

Table of Contents

List of Acronyms and Abbreviations Used in this Report

AE	Adverse events
AHRQ	Agency for Healthcare Research and Quality
ALT	Alanine transaminase
AST	Aspartate aminotransferase
BMI	Body mass index
CI	Confidence interval
CR	Controlled release
DRSP	Drospirenone
DVT	Deep vein thrombosis
DYD	Dydrogesterone
evLY	Equal value life year
FDA	Food and Drug Administration
FSH	Follicle stimulating hormone
GCS	Greene Climacteric Scale
KNDy	Kisspeptin, neurokinin B, and dynorphin
MCID	Minimum clinically important difference
MENQoL	Menopause-specific quality of life
MG	Milligram
MHT	Menopausal hormone therapy
MI	Myocardial infarction
N	Number
NETA	Norethisterone acetate
NMA	Network meta-analysis
NR	Not reported
QALY	Quality-adjusted life years
RCT	Randomized control trial
REF	Reference group
SD	Standard deviation
SE	Standard error
SERM	Selective estrogen receptor modulators
SNRI	Serotonin-norepinephrine reuptake inhibitor
SSRI	Selective serotonin reuptake inhibitor
US	United States
VMS	Vasomotor symptoms
WHI	Women's Health Initiative

ICER recognizes that gender language is evolving and that individuals experiencing menopause may have diverse and dynamic gender identities. In this report, when we use the word "woman" (and the pronouns "she" and "her"), we are describing adult individuals whose biologic sex is female, whether they identify as female, male, or non-binary, among others. When referencing study populations used in specific research studies, we will use the gender language used by the study investigators. As gender language continues to evolve in efforts to build inclusivity and equity into the health system, ICER will periodically reassess this language, consult with subject matter experts, and make appropriate adjustments as necessary in future versions of this and other reports to ensure that this language is fully inclusive and affirming.

Executive Summary

Vasomotor symptoms (VMS), namely hot flashes and night sweats, are the hallmark symptoms of menopause. The vast majority of women undergoing menopause experience some degree of VMS (>75%), but a substantial proportion (32 – 46%) experience moderate to severe VMS such that there is impairment in quality of life or interference with normal activities.[1] The median total duration of moderate to severe VMS is 9.4 years.[1] Frequent moderate to severe VMS episodes (7 or more per day) are associated with interference with sleep, concentration, mood, energy, and sexual activity.[2] VMS duration and severity are also known to differ by race and ethnicity.[3] In addition to impacts on quality of life, VMS also has both direct healthcare costs and indirect economic costs due to missed work.[4]

A number of therapeutics (e.g., anti-depressants, gabapentinoids) have been investigated to treat VMS, with Menopausal Hormone Therapy (MHT) generally considered the mainstay of treatment.[5] However, MHT may be medically contraindicated in some patients and not desired by others. Fezolinetant (Astellas Pharma Inc.) is a once daily oral nonhormonal therapy being investigated for the treatment of moderate to severe VMS associated with menopause. It acts by regulating neuronal activity in the hypothalamus thereby affecting temperature regulation. If approved, it would be the first selective neurokinin-3 (NK3) receptor antagonist available in the United States (US). On June 23rd, 2022, Astellas submitted a New Drug Application for fezolinetant 45 mg to the FDA.[6]

We compared the clinical and cost effectiveness of fezolinetant to no pharmacologic treatment (prescription nor non-prescription), to MHT, and where evidence was available, to selective serotonin reuptake inhibitors (SSRIs)/serotonin–norepinephrine reuptake inhibitors (SNRIs), gabapentin and pregabalin. Two doses of fezolinetant (30 mg and 45 mg) were studied in two Phase 3 randomized controlled trials (RCTs) conducted primarily in the US (Skylight 1 and 2) and an additional Phase 3 RCT conducted in Asia which studied the 30 mg dose only (Moonlight 1). While

both doses in the US-based studies showed statistically significant improvements in VMS frequency and severity compared with no pharmacologic treatment, the Asia-based study did not find statistically significant improvements in either outcome. In terms of safety, fezolinetant was generally well tolerated and liver injury only occurred in higher doses (≥60 mg).

When compared to no pharmacologic treatment and averaged across trials, only MHT achieved clinically significant differences for both VMS frequency and severity. When averaged across trials, fezolinetant achieved clinically significant reductions in VMS frequency; although reductions in VMS severity were statistically significant, they did not achieve clinical significance. For SSRIs/SNRIs and gabapentin, reductions in both outcomes were not clinically significant. For pregabalin, VMS frequency reductions were clinically significant. There was considerable heterogeneity across included studies.

The discrepancy in outcomes between the Skylight and Moonlight trials raises uncertainty regarding the efficacy of fezolinetant, particularly at lower doses. The short duration of the trials (12 weeks) also creates uncertainty about long-term efficacy. In terms of safety, fezolinetant was well tolerated and there were no serious adverse events noted in the Phase 3 RCTs. However, only limited data from the three Phase 3 RCTs were available for review in this report. In addition, fezolinetant possesses a unique mechanism of action without other in-class data available and liver injury has been documented at higher doses. Given the modest benefit observed in RCTs and uncertainty about long-term benefit and overall safety, we have assigned an ICER Evidence Rating for the overall net health benefits of fezolinetant versus no pharmacologic treatment for VMS of "Promising but Inconclusive" (P/I).

In comparing fezolinetant to other interventions (e.g., MHT), there have not been any head-to-head trials with active comparators. Due to low quality or insufficient evidence, we did not compare fezolinetant to SNRI/SSRI, gabapentin, or pregabalin. In indirect comparison of the treatment effects of fezolinetant vs. MHT, MHT resulted in greater reductions in both VMS frequency and severity when compared to fezolinetant, but heterogeneity across the trials creates uncertainty about this conclusion. Over the short-term, the safety and tolerability of fezolinetant and MHT appear comparable. However, longer term use of MHT carries serious increased risks including coronary heart disease, stroke, venous thromboembolism, breast cancer and mortality; particularly among women older than 60 years old. There is uncertainty over longer term use of fezolinetant. In sum, there is considerable uncertainty and insufficient evidence to draw conclusions about the overall net health benefit of fezolinetant vs. MHT, leading to an ICER Evidence Rating of "Insufficient" (I).

Table ES1. Evidence Ratings

Treatment	Comparator	Evidence Rating
Fezolinetant	No pharmacologic treatment	P/I
Fezolinetant	MHT	I

MHT: menopausal hormone therapy

We developed a *de novo* decision analytic model to evaluate fezolinetant for the treatment of VMS compared with no pharmacologic treatment. We also conducted an economic evaluation of MHT compared with no pharmacologic treatment. At a placeholder price of $6,000 annually, fezolinetant exceeds commonly accepted cost-effectiveness benchmarks. However, draft results suggest that fezolinetant would meet these benchmarks and be considered cost-effective if priced at $1,900 annually. MHT is widely available as generic medication and is cost-effective. Table ES2 presents the incremental results from the cost-effectiveness analyses for fezolinetant and MHT.

Table ES2. Incremental Cost-effectiveness Ratios for Fezolinetant and MHT

Treatment	Comparator	Cost per QALY Gained	Cost per evLY Gained	Cost per Symptom-free Day
Fezolinetant*	No pharmacologic treatment	$390,000	$390,000	$500
MHT	No pharmacologic treatment	$13,000	$13,000	$12

evLYs: equal value life year, MHT: menopausal hormone therapy, QALYs: quality-adjusted life year

*Based on annual placeholder price of $6,000. Interpret findings with caution

In sum, fezolinetant appears promising in the treatment of VMS, but considerable uncertainty about efficacy and long-term safety still exists. Currently available evidence suggests that it is likely either comparable or may be inferior to MHT in terms of benefits, but this must also be balanced against the known safety profile of MHT. The cost-effectiveness of fezolinetant will depend upon its price and whether it is considered an alternative to treatment to MHT for all women or whether it will primarily be used by women who cannot or will not take MHT.

1. Background

Eighty percent of women undergoing menopause experience vasomotor symptoms (VMS). The pathophysiology of VMS, characterized by hot flashes and night sweats, has not been fully elucidated. Purported mechanisms include changes in estrogen levels and increased neurokinin B (NKB) activity acting on the hypothalamus, a region of the brain which regulates body temperature.[7-9] Changes in thermoregulation may increase blood flow to the skin, resulting in the VMS. Hot flashes are the sudden onset of heat in the upper chest and face which spreads throughout the body, typically lasting two to four minutes. Hot flashes are often accompanied by profuse sweating and, when this occurs at night (night sweats), can cause sleep disruption, and negatively affect mood. Additionally, the combination of vasodilation and sweating can rapidly lower body temperature and cause shivering.

The frequency of VMS varies from one per day to one per hour. Most women who have them experience several per day. The intensity of VMS can be classified as mild (sensation of heat without sweating), moderate (sensation of heat with sweating but able to continue activity), or severe (sensation of heat with sweating, causing cessation of activity). Moderate to severe VMS affects 32% to 46% of women undergoing menopause.[1] Women with frequent moderate to severe VMS (i.e., 7 or more episodes per day) often report interference with sleep (94%), concentration (84%), mood (85%), energy (77%), and sexual activity (61%).[2] Risk factors for developing VMS include obesity, smoking, reduced physical activity, high follicle-stimulating hormone (FSH) levels, and mutations in the tachykinin receptor 3 gene, which encodes the NKB receptor.[10,11]

Available data suggest that the median total duration of moderate to severe VMS is 9.4 years[1]; bothersome VMS may last around 4 years.[12] However, severity and duration of VMS symptoms appears to be heterogeneous, with racial and ethnic differences. For example, a higher proportion of Black women experience increased severity and duration of VMS symptoms compared to White women. [3] More limited evidence suggests that Chinese women typically have the shortest duration of symptoms and Native American women may have the highest prevalence of VMS. [13,14] [15] VMS are estimated to increase direct healthcare costs by $1,300 per person per year compared to women without these symptoms, and increase indirect economic costs due to missed work by another $770 per person per year.[4]

Treatment options vary based on symptom severity. For women with mild VMS symptoms, behavioral approaches (e.g., lowering ambient temperature, dressing in layers of clothing) can be effective. For women with moderate to severe VMS and no contraindications, menopausal hormone therapy (MHT), consisting of estrogen and progesterone (for women with an intact uterus) or estrogen alone (for women who have undergone a hysterectomy) is recommended as first-line therapy. Contraindications to MHT include a history of breast cancer, coronary heart

disease, venous thromboembolic (VTE) event or stroke, active liver disease, unexplained vaginal bleeding, high-risk for endometrial cancer, or transient ischemic attack[16], since MHT can potentially increase the risk of these adverse outcomes.

In women who have contraindications to MHT, nonhormonal treatments may be considered for treatment of VMS. Complementary and alternative therapies (e.g., yoga, supplements) have been studied, but evidence for the effectiveness of such treatments is, at best, inconclusive due to heterogeneity between trials and low-quality evidence.[17,18] Antidepressants such as selective serotonin reuptake inhibitors (SSRIs) and serotonin-norepinephrine reuptake inhibitors (SNRIs) have been studied for the treatment of VMS. While paroxetine is the only antidepressant with an FDA-approved indication for VMS[19], other SSRIs and SNRIs have also been shown to be effective in reducing symptoms in some trials.[20,21] Gabapentin and pregabalin have also been studied for the treatment of VMS, particularly for women who have sleep disturbances as they can be sedating.

Fezolinetant (Astellas Pharma Inc.) is a once daily oral nonhormonal therapy being investigated for the treatment of moderate to severe VMS associated with menopause. It acts by regulating neuronal activity in the hypothalamus thereby affecting temperature regulation. If approved, it would be the first selective neurokinin-3 (NK3) receptor antagonist available in the US. On June 23rd, 2022, Astellas submitted a New Drug Application for fezolinetant 45 mg to the FDA.[6]

Table 1.1. Interventions of Interest

Intervention Brand Name (Generic Name)	Mechanism of Action	Delivery Route	Prescribing Information
Fezolinetant	Neurokinin-3 receptor antagonist	Oral	45 mg once daily

mg: milligrams

2. Patient and Caregiver Perspectives

We spoke with representatives from three patient advocacy groups, as well as five individual patients who are experiencing VMS from menopause. We supplemented information obtained from our interviews with patients and advocacy groups with an online survey conducted by the National Menopause Foundation (see supplement for survey methods details).

Every patient we interviewed described the large effect of VMS on their lives. During the day, patients mentioned that VMS episodes would cause them to be "completely drenched", feel "on fire", and that skin flushing would cause embarrassment and negatively affect their ability to function in the workplace and interact with others. Patients also discussed the anxiety associated with the unpredictable timing and rapid onset of hot flashes and having no way to plan for an episode. During the night, patients mentioned that night sweats severely reduced sleep quality thereby limiting their function during the day. Patients also indicated that they are less likely to embrace others or be sexually intimate given the uncomfortable temperature increase.

Some patients discussed the lack of recognition of their symptoms by their healthcare providers, causing patients to worry that their symptoms were indicative of the onset of other health conditions. Other patients mentioned that their healthcare providers considered their symptoms as unavoidable parts of menopause and did not offer further information about treatment. The lack of information and recognition of the burden of VMS for menopausal women caused some women to feel disempowered and prevented them from engaging with their healthcare providers on this topic. Instead, in the absence of discussion from their healthcare providers, they sought alternative information sources, such as family members, friends and church members. Our interviews mirrored findings from a survey conducted by the National Menopause Foundation, where respondents were more likely to speak to friends about menopause than their primary care physician or gynecologist, even though healthcare providers were viewed as a more reliable source of information.

To manage VMS, women discussed changing their wardrobe, bedding, diet, and behaviors to stay cooler, and avoiding physical and emotional triggers. Some patients were offered MHT and antidepressants by their providers. Patients also tried holistic and over the counter treatments. In terms of future treatments, patients were concerned about potential adverse effects, such as increased risk of cancer, and that health plans would require high cost-sharing, prior authorization, or may not cover the medication It was also highlighted that clinical trials are often not demographically diverse and may therefore not adequately represent the burden of VMS symptoms and treatment effects in all racial and ethnic groups.

In discussions with clinical experts, we heard that VMS is undertreated, in part due to the lack of clinicians with expertise in treating menopause symptoms. Furthermore, clinicians highlighted that the heightened risk of cardiovascular disease with estrogen and progestin seen in the Women's Health Initiative (WHI) study may not be generalizable to the entire population of women with VMS, given that treatment may start at younger ages, use different formulations, and have different durations of treatment than in the WHI.[22] However, depending on individual patient characteristics as well as the type (e.g., route of administration, dose, combination hormones) and duration of MHT, for some women, the risks of MHT may outweigh the benefits.[23] Thus, safe and effective nonhormonal treatment options are an important need. Finally, we heard that while the FDA guidance for industry on clinical evaluation is to include women with 7 to 8 moderate to severe hot flashes per day,[24] this is a only a subset of women who experience VMS, so clinical trial results may not necessarily be applicable to women whose symptoms are not as severe.

3. Comparative Clinical Effectiveness

3.1. Methods Overview

Procedures for the systematic literature review assessing the evidence on fezolinetant, menopausal hormone therapy (MHT), SSRIs/SNRIs, gabapentin, and pregabalin for the treatment of VMS associated with menopause are detailed in Section D of the Supplement.

Scope of Review

We systematically reviewed the clinical effectiveness of the medications for treatment of moderate to severe VMS compared to no pharmacologic therapy, as represented by the placebo arm of clinical trials. For studies evaluating multiple doses, we reviewed only the doses that met the level considered to be the minimally effective dose. Our primary outcomes were changes in frequency and severity of VMS from baseline. Other outcomes included menopause-specific quality of life (MENQoL), sleep disturbances, urogenital symptoms, and mood. To ensure comparability of results, we present trials in the main report that are most similar to the Phase 3 fezolinetant clinical trials in terms of: study design (i.e., Randomized Control Trial), relevant population (i.e., general healthy women experiencing frequent [7 per day] VMS associated with menopause), assessment of VMS (i.e., self-reported frequency and severity of VMS), and length of follow-up (i.e., between 8 to 16 weeks). We also sought evidence on longitudinal outcomes associated with the risks of MHT, focusing on trials that reported outcomes a beyond 1-year duration. Other included trials not meeting the comparability criteria above are reported in the supplement. We found there was heterogeneity in the assessment of VMS frequency and VMS severity (such as measuring moderate to severe or all [mild, moderate, and severe] VMS, measuring weekly vs. daily VMS, measuring in real time vs. retrospective). Throughout the results, we report any differences in assessment. While we had proposed conducting a network meta-analysis (NMA) for VMS frequency and severity and MENQoL outcomes, after performing a literature search and applying inclusion criteria, there was an insufficient number of studies that measured VMS in a similar manner (e.g., daily frequency of moderate to severe VMS) or provided enough data on changes in MENQoL and thus we were unable to conduct an NMA for this outcome. The full scope of the review is available in Section D2 of the Supplement.

Evidence Base

Fezolinetant

Evidence informing our review of fezolinetant for treatment of VMS was derived from three Phase 3 trials: Skylight 1, Skylight 2, and Moonlight 1.[25-28] At the time of this report, published data from both Skylight trials was available for review and a summary of key findings from Moonlight 1 was issued as a press release only. Additional data on long-term harms was provided in a press release describing results from the single-arm Moonlight 3 trial.[29] Full results of additional Phase 2 trials of fezolinetant are described in Section D2 and Tables D3.1-9 of the Supplement.

Skylight 1 and 2 evaluated oral fezolinetant 30 mg and 45 mg daily versus placebo for 12 weeks, with an unblinded non-controlled extension phase lasting an additional 40 weeks.[25-27] Moonlight 1 evaluated oral fezolinetant 30 mg daily versus placebo for 12 weeks, with an unblinded non-controlled extension phase lasting an additional 12 weeks.[30] Moonlight 3 was a single-arm trial that evaluated the safety and tolerability of oral fezolinetant 30 mg daily versus placebo for 52 weeks, with a follow-up visit three weeks after the last dose.[31] In this report, we report results from the 30 mg dose as well as the 45 mg dose as both doses were evaluated in the Phase 3 trials.[32] Skylight 1 and 2 were conducted across 93 international locations, with the majority of sites located in the US, whereas Moonlight 1 and 3 were conducted across 48 locations exclusively in Asia.

Participants in all fezolinetant trials were women aged 40-65 years of age, BMI ≥18 kg/m^2 and ≤38 kg/m^2, seeking relief for VMS associated with menopause with a minimum average of 7-8 moderate to severe hot flashes per day or 50-60 per week, and in good general health. They could not be taking any other pharmacologic (prescription nor non-prescription) treatment for VMS. Detailed inclusion/exclusion criteria are described in Table 3.1. At the time of posting this report, demographic information was unavailable for Skylight 1 and Moonlight 1 and 3. Baseline characteristics for Skylight 2 are described in Table 3.1. All fezolinetant trials measured daily moderate-severe VMS frequency and severity. Severity was rated by the participant as 1=mild (sensation of heat without sweating), 2=moderate (sensation of heat with sweating and able to continue activity), and 3=severe (sensation of heat with sweating causing cessation of activity).

Menopausal Hormone Therapy

Evidence informing our review of MHT for treatment of VMS was derived from ten RCTs that examined the use of a combined estrogen and progesterone dose or estrogen-only in postmenopausal women. There were five RCTs, including the REPLENISH trial, that evaluated a standard dose of oral estrogen (1 mg estradiol daily) with progesterone versus placebo for 12[33,34], 13[35], and 16 weeks.[36,37] Five RCTs, including CHOICE and MsFLASH 03 trials, evaluated a low dose of oral estrogen (0.5 mg estradiol daily) with progesterone versus placebo for 8[38], 12[33,39,40], and 13 weeks.[35] Two estrogen-only RCTs, published in one manuscript, evaluated transdermal estradiol

(standard dose: 0.04 mg or low dose: 0.02 mg per day) versus placebo for 12 weeks in women with a prior hysterectomy.[41] Baseline characteristics and inclusion/exclusion criteria are outlined in Table 3.1 and Tables D3.1-3. Baseline characteristics for the eight RCTs appeared to be similar, except participants in the two estrogen-only trials were younger (mean age: 49 years) and were mostly White (81%).[41] Five MHT trials included only changes in moderate-severe VMS frequency and/or severity (similar to the fezolinetant trials), while the other MHT trials additionally included mild VMS in their measurement. Table A.1 in the Supplement describes how each included study defines VMS severity.

SSRIs/SNRIs

Evidence informing our review of SSRIs/SNRIs for the treatment of VMS was derived from 10 RCTs in postmenopausal women: three RCTs examined the efficacy of SSRIs (paroxetine and escitalopram)[42,43] and seven RCTs examined the efficacy of SNRIs (venlafaxine or desvenlafaxine).[38,44-49] For SSRIs, two RCTs evaluated oral paroxetine 7.5 mg daily versus placebo for 12 weeks[42] and one multisite RCT, MsFLASH 01 trial, evaluated oral escitalopram 10 mg daily versus placebo for 8 weeks.[43] For SNRIs, two RCTs, including the multisite RCT MsFLASH 03, evaluated oral venlafaxine (37.5 mg daily for the first week and then increased to 75 mg daily) versus placebo for 8 weeks[38] and 12 weeks.[44] Five RCTs evaluated oral desvenlafaxine 100 mg daily versus placebo for 12 weeks[45-49] Some trials also included 150 mg[45,46,49] and/or 200 mg doses.[46,50] Baseline characteristics and inclusion/exclusion criteria are outlined in Table 3.1 and Tables D3.1-3. Majority of trials recruited predominately White participants (75%-93%), except three trials that included at least 25% Black/African American participants.[38,42,43] Similar to the fezolinetant trials, Simon et al. measured only moderate-severe VMS, whereas all other SSRIs trials measured all VMS (mild, moderate, and severe). Trials differed in their measurement of VMS severity (see Table A.1 in Supplement).

Gabapentin

Evidence informing our review of gabapentin for the treatment of VMS was derived from three RCTs including postmenopausal women [51-53]. The trials evaluated: oral gabapentin 900 mg daily[53], gastroretentive gabapentin 1800 mg daily[51], or oral gabapentin titrated to 2,400 mg (400mg per capsule) daily [52], versus placebo for 12 weeks. A larger dose was chosen in the Reddy et al. (2006) trial as an open-label trial reported a plateau of efficacy at 1,800-2,400 mg. (Table 3.1. and Tables D3.1.-3.). Baseline characteristics for the three RCTs were similar (see Tables 3.1.), except Guttuso et al.[53] included a majority White sample (93%). All trials measured daily VMS (mild, moderate, and severe) frequency, except Reddy et al.[52] which measured weekly VMS frequency. Severity was determined differently across all trials (see Table A.1 in Supplement).

Pregabalin

There were no studies of pregabalin that met our criteria for inclusion in the systematic review therefore it is not considered in the remainder of the report.

Table. 3.1. Baseline Characteristics and Inclusion and Exclusion Criteria.

Title/Author	Inclusion	Exclusion	Baseline characteristics
Fezolinetant			
SKYLIGHT 1 [54,55] N=522 Fezolinetant 30 mg (N= 173) Fezolinetant 45 mg (n=174) Placebo (n= 175)	Healthy postmenopausal women aged 40-65 years, BMI ≥18 kg/m^2 and ≤38 kg/m^2, with an average of 7-8 moderate-severe hot flashes per day or 50-60 per week.	Current use of a prohibited therapy (any pharmacologic treatment for VMS), known substance or alcohol use disorder, history of suicide attempt, previous/current history of malignant tumor, high systolic (≥ 130 mmHg) or diastolic (≥80 mmHg) blood pressure, severe allergy or intolerance to drugs, presence of disordered proliferative endometrium, endometrial hyperplasia, or endometrial cancer, or has any other medical disorder that could confound study outcome.	NR
SKYLIGHT 2 [55-57] N=500 Fezolinetant 30 mg (n=166) Fezolinetant 45 mg (n=167) Placebo (n=167)	See Skylight 1	See Skylight 1	Age: 54.3 (SD=5.0) Race/Ethnicity: 79.1% White, 20% Black, 0.8% Other BMI: 28.04 (range: 18-38) Hysterectomy: 32.4%
Moonlight 1 N=302 Fezolinetant 30 mg (n=NR) Placebo (n=NR)	See Skylight 1	See Skylight 1	NR
Moonlight 3 N= 150 Fezolinetant 30 mg (n=150)	See Skylight 1	See Skylight 1	NR

MHT – standard dose (Estradiol 1 mg)			
Schürmann et al. (2004)[36] N=225 Placebo (n=61) Estradiol 1 mg/DRSP 1 mg (n=55) Estradiol 1 mg/DRSP 2 mg (n=52) Estradiol 1 mg/DRSP 3 mg (n=57)	Healthy post-menopausal women aged 45–65 years, who complained of at least 5 moderate to severe hot flushes per day during the screening period.	Contraindications for MHT, treatment with anticoagulant medications, recent use of oral, transdermal, or transvaginal hormonal preparations. Past medical history for cardiovascular disease, depression, diabetes, hypertension, or other diseases that could affect the study results.	Age: 53.7 (SD=4.75) Race/Ethnicity: NR BMI: 26.2 (4.13) Natural menopause: NR
Endrikat et al. (2007)[34] N=324 Estradiol valerate 1 mg/dienogest 2 mg (n=162) Placebo (n=162)	Women aged 52–65 years in general or aged 40–51 years in case of previous bilateral oophorectomy, and had an intact uterus.	Contraindications to HRT; any disease/conditions that compromised the function of the body systems; abnormal cervical smear; abnormal baseline lab values considered clinically significant; history of alcohol or drug abuse; current significant liver dysfunction; insulin-dependent diabetes; hypertension; concomitant medication with drugs known to influence the study medication; any severe systemic disease that could interfere with the study.	Age: 56.3 (SD=4.9) Race/Ethnicity: NR BMI: 26.6 (SD=3.8) Hysterectomy: 33%
Lin et al. (2011)[37] N=244 Estradiol 1 mg/DRSP 2 mg (n= 183) Placebo (n=61)	Women who had 24 or more moderate to severe hot flushes over 7 consecutive days during the screening period, had a negative pregnancy test and negative bilateral mammography results.	History of cardiovascular disease, uncontrolled thyroid disorders, clinical depression, malignant or premalignant disease, abnormal gynecologic findings, hepatic disease, adrenal insufficiency or renal failure, abnormal glucose tolerance and severe or congenital hyper-triglyceridemic; abnormal baseline laboratory findings; a history of alcohol/drug abuse or current smoking; recent hormonal therapy; use of herbal/other medicines for climacteric disorders.	Age: 51.9 (3.75) Race/Ethnicity: 100% Asian BMI: 23.4 (SD=2.84) Natural menopause: NR

REPLENISH: Lobo et al. (2018)† [33,39,58-61] N=1,411 Estradiol 1 mg and progesterone 100 mg (n=415) Estradiol 0.5 mg and progesterone 100 mg (n=424) Estradiol 0.5 mg and progesterone 50 mg (n=421) Placebo (n=151)	Healthy menopausal women aged 40–65 years with BMI 34 kg/m^2 or less, had an intact uterus and at least 12 months of spontaneous amenorrhea.	Contraindications or allergy to MHT; a history of endometrial hyperplasia or undiagnosed vaginal bleeding; uterine fibroids diagnosed at screening; heavy smoking, or a history of drug or alcohol abuse; recent use of another therapy for VMS.	Age: 54.7 (SD=4.4) Race/Ethnicity: 65.9% White, 31.6% Black, 2.5% Other BMI: 20.7 (SD=4.1) Natural menopause: NR
Simon et al. (2019)¤ [62] VMS substudy of Lobo et al. (2018) N=572 Estradiol 1 mg and progesterone 100 mg (n= 141) Estradiol 0.5 mg and progesterone 100 mg (n= 149) Estradiol 0.5 mg and progesterone 50 mg (n= 147) Placebo (n= 135)	See Lobo et al. (2018). Additional requirement: Women who had a minimum of 7 moderate to severe VMS daily or 50 per week before enrollment.	See Lobo et al. (2018)	Age: 54.68 (SD=4.6) Race/Ethnicity: 67.1% White, 31% Black, 1.9% Other BMI: 26.7 (SD=3.98) Natural menopause: NR
MsFLASH 03: Joffe et al.* (2014) [38] N=339 Estradiol 0.5 mg (n=97) Venlafaxine 75 mg (n=96) Placebo (n=146)	Healthy women aged 40 to 62 years in the menopause transition, were postmenopausal, had FSH level exceeding 20 mIU/mL, and an estradiol level not exceeding 50 pg/mL in the absence of a reliable menstrual marker and were required to have at least 14 VMS per week	Pregnancy or breastfeeding; suicide attempt in the past 3 years; diagnosis of bipolar disorder or psychosis; psychotropic medications for VMS in the past month; major depressive episode or drug or alcohol abuse in the past year, recent or current use of MHT; hormonal contraceptives, SERM or aromatase inhibitors, and some comorbidities.	Age: 54.6 (SD=3.8) Race/Ethnicity: 59.9% White, 34.2% Black, 20% Other BMI: 28.3 (SD=6.8) Natural menopause: NR

MHT – low dose (Estradiol 0.5 mg or lower)			
CHOICE: Panay et al. (2007) [40] N=575 Estradiol 0.5 mg/0.1 mg NETA (n=194) Estradiol 0.5 mg/0.25 mg NETA (n= 181) Placebo (n= 200)	Women who had at least 50 moderate to severe hot flushes per week, no menses during the past year or 6 months spontaneous amenorrhea with FSH levels 440 mIU/ml and estradiol levels 525 pg/ml.	Recent exposure to MHT. Suspected or previous history of breast cancer or estrogen-dependent neoplasia, untreated endometrial hyperplasia and abnormal genital bleeding. History of diabetes mellitus, hypertension, any thrombo-embolic conditions and hepatic or renal impairment.	Age: 55.5 (SD=4.6) Race/Ethnicity: 95% White, 0% Black, 1% Asian BMI: 67.8 (10.5) Natural menopause: NR
Stevenson et al. (2010)‡ [35,63] N=313 Estradiol 1 mg/DYD 5 mg (n=62) Estradiol 0.5 mg/DYD 2.5 mg (n=124) Placebo (n=127)	Non-hysterectomized, postmenopausal women aged 45–65 years who had been amenorrhoeic for ≥ 12 months, had serum estradiol and FSH levels within the post-menopausal range, had ≥ 50 moderate to severe hot flushes during the screening period.	Endometrial biopsy showing clinically relevant abnormalities and/or bilayer endometrial thickness of ≥5 mm, recent abnormal vaginal bleeding, a history of or current estrogen.	Age: 53.8 (SD=4.2) Race/Ethnicity: NR BMI: 26.36 (SD=6.42) Natural menopause: NR
Archer et al. (2013) [39] N=675 (full analysis set) Estradiol 0.5 mg/DRSP 0.25 mg (n=177) Estradiol 0.5 mg /DRSP 0.5 mg (n=178) Placebo (n=176)	Women aged 40 years or older, experienced spontaneous amenorrhea for 12 months or more, had a minimum of 7 to 8 moderate to severe VMS per day, or 50 to 60 moderate to severe VMS per week during the screening period.	Recent use of oral hormonal products.	Age: 53.5 (6.0) Race/Ethnicity: 67.6% White, 24.2% Black, 0.6% Asian, 7% Hispanic BMI: 28.5 (5.84) Hysterectomy: 54.4%

Speroff et al. (1996)[41] N=324 Study 1: Placebo (n=54) Estradiol transdermal system: 0.02 mg (n=54) Estradiol transdermal system: 0.04 mg (n=53) Study 2: Placebo (n=37) Estradiol transdermal system: 0.02 mg (n=37) Estradiol transdermal system: 0.04 mg (n=37)	Women at least 50 years of age, undergone hysterectomy, had natural menopause or at least 35 years of age, had surgical menopause, and screened for baseline VMS (at least 56 per week).	Contraindications to MHT; those with a skin condition that may be exacerbated by use of transdermal system.	Age: 49 (SD=NR) Race/Ethnicity: 81% White (other categories NR) BMI: NR Natural menopause: 28%
SNRIs			
Evans et al. (2005)[44] N=80 Venlafaxine 75 mg (n=40) placebo (n=40)	Women with natural or surgical menopause and had more than 14 hot flushes per week.	Receiving estrogens, progestins, androgens, antidepressants, or chemotherapy.	Age: 52.2 (5.5) Race/Ethnicity: 76.5% White, 8.5% Black, 8.5% Asian, 6.5% Hispanic BMI: NR Natural menopause: 79.3
Speroff et al. (2008)[46,50] N=563 Desvenlafaxine 100 mg (n=157) Desvenlafaxine 150 mg (n=163) Desvenlafaxine 200 mg (n=155) Placebo (n=78)	Healthy postmenopausal women with BMI 40 kg/m^2 or less who experienced at least 7 moderate-to-severe hot flushes per day (or 50 or more per week).	Recent use of MHT or therapies for VMS; history of seizure disorder; myocardial infarction; malignancy other than basal or squamous cell carcinoma; glaucoma or raised intraocular pressure; or hepatic, renal medical disease; current major depressive, bipolar, psychotic, or generalized anxiety disorder; other clinically important abnormalities at screening.	Age: 53.6 (SD=4.97) Race/Ethnicity: 83.9% White, 9.95% Black, 6.1% Other BMI: 26.9 (SD=4.6) Natural menopause: 77.9%
Archer et al. (2009) (12-week)[64] N=451 Desvenlafaxine 150 mg (n= 151) Desvenlafaxine 100 mg	Healthy postmenopausal women with BMI 40 kg/m^2 or less who experienced at least 7 moderate to severe hot flushes per day or 50 or more per week for 2 consecutive weeks at baseline.	Recent use of any hormone-containing drug or VMS therapy; history of seizure disorder, myocardial infarction, or malignancy or treatment for malignancy other than basal or squamous cell carcinoma; hepatic, renal, or	Age: 53.36 (SD=4.8) Race/Ethnicity: 82.7% White, 15.7% Black, 1.7% Other BMI: 27.86 (4.96) Natural menopause: 80%

(n=150) Placebo (n=150)		other medical disease; presence of psychiatric disease requiring therapy.	
Archer et al. (2009) (26-week study)[65] N=541 Desvenlafaxine 100 mg (n=182) Desvenlafaxine 150 mg (n=179) Placebo (n=180)	Healthy, postmenopausal women with BMI ≤ 40 kg/m^2 who experienced at least 7 moderate to severe hot flashes per day (or 50/ week) recorded by participants for 7 consecutive days during screening.	Recent use of any hormone-containing drug or VMS therapy; history of seizure disorder, myocardial infarction, or malignancy or treatment for malignancy other than basal or squamous cell carcinoma; hepatic, renal, or other medical disease; presence of psychiatric disease requiring therapy.	Age: 53.7 (SD=5.03) Race/Ethnicity: 87.3% White, 10.9% Black, 1.8% Other BMI: 27.1 (4.59) Natural menopause: 76.2%
Bouchard et al. (2012)[47] N=287 Placebo (n=150) Desvenlafaxine 100 mg (n=137)	Healthy women who had completed their last natural menstrual period ≥12 months prior to screening, had an intact uterus, a BMI of ≤34.0 kg/m2, and a minimum of 7 moderate and severe VMS per day, or ≥50 moderate and severe VMS per week recorded for 7 consecutive days during screening.	Recent use of any hormone-containing drug or VMS therapy; estrogen-dependent neoplasia; seizure disorder; active or recent arterial thrombo-embolic disease; cerebrovascular accident or stroke; venous thromboembolism; malignancy or treatment for malignancy within 2 years; hepatic, renal medical disease; major depressive, bipolar, psychotic, or generalized anxiety disorder requiring therapy; narrow-angle glaucoma or current raised intraocular pressure; undiagnosed vaginal bleeding.	Age: 54 (SD=4.5) Race/Ethnicity: 92.5% White, 0.5% Black, 7% Other BMI: 26 (SD=4) Natural menopause: NR
Pinkerton et al. (2013)§[48,66] N=365 Desvenlafaxine 100 mg (n=181) Placebo (n=184)	Women aged 45 years or older, had a BMI of 34.0 kg/m^2 or lower, and had confirmed menopause status. Efficacy sub-study: Approximately 20% of the enrolled participants met the additional criterion of 7 or more moderate and severe VMS per day (or ≥50 VMS per week) recorded for 2 weeks before randomization.	Recent use of any VMS therapy; history of seizure disorder, myocardial infarction, narrow-angle glaucoma, or malignancy or treatment of malignancy other than basal or squamous cell carcinoma; important medical disease; major depressive, bipolar, psychotic, or generalized anxiety disorder requiring therapy; other clinically important abnormalities at screening.	Age: 54 (SD=5) Race/Ethnicity: 86.5% White, 12% Black, 1.5% Other BMI: 26.5 (SD=4) Natural menopause: 79%
SSRIs			

MSFLASH 01 (Freeman et al. 2011) [43,67] N=205 Escitalopram 10 mg (n=104) Placebo (n=101)	Women who had at least 28 hot flashes or night sweats per week for 3 weeks where hot flashes or night sweats rated as bothersome or severe on 4 or more days per week and the frequency in week 3 did not decrease by more than 50% from the mean weekly levels in weeks 1 and 2.	Use of therapies for hot flashes in the past 30 days; current severe medical illness, major depressive episode, drug or alcohol abuse in the past year, suicide attempt in the past 3 years, lifetime diagnosis of bipolar disorder, or psychosis; or uncontrolled hypertension, history of endometrial or ovarian cancer, myocardial infarction, angina or cerebrovascular events, or other preexisting medical conditions.	Age: 53.9 (SD=4.03) Race/Ethnicity: 49.8% White, 46.4% Black, 3.9% Other BMI: 29.1 (SD=6.51) Hysterectomy: 13%
Simon et al. (2013) (Study 1 & 2) [68-70] N=606 Paroxetine 7.5 mg(n= 301) Placebo (n=305)	Postmenopausal women 40 years or older who had an average of more than 7-8 moderate to severe hot flashes per day, or 50-60 moderate to severe hot flashes per week.	Hypersensitivity to paroxetine; known nonresponse to previous SSRI or SNRI treatment of VMS; untreated hypertension; impaired liver or kidney function; unstable cardiac disease; pregnancy; a history of psychiatric disorder; and any other medical condition.	Age (median): 54 (range: 40-79) Race/Ethnicity: 69.9% White, 27.4% Black, 0.95% Asian BMI (median): 28.05 (range: 16.8-60.7) Natural menopause: 81.1%

Gabapentin			
Guttuso et al. (2003) [53,71] N=59 Gabapentin 900 mg (n=30) Placebo (n=29)	Postmenopausal women with an average of 7 or more hot flashes per day accompanied by sweating.	Estrogen, progestin, leuprolide, or tamoxifen therapy within the past 2 months. More than 50% of a patient's hot flashes were associated with occurrence of migraine headaches or ingestion of particular foods or beverages.	Age: 52.9 (SD=3.4) Race/Ethnicity: 93.2% White, 6.8% Black BMI: NR Natural menopause: NR
Reddy et al. (2006) [52] N=40# Gabapentin 900 mg (n=20) Placebo (n= 20)	Menopausal women, aged 35 - 60 years, experiencing at least 50 moderate to severe hot flashes per week > 2 months.	History of DVT, history of MI, stroke, and/or functional decline, history of malignancies or undiagnosed vaginal bleeding, history of chronic liver, gallbladder, chronic renal, cardiac, or endocrine diseases.	Age: 51.75 (SD=4.36) Race/Ethnicity: NR BMI: 26.8 (SD=5.9) Natural menopause: 90%
Breeze 3: Pinkerton et al. (2014) [72] N=593 Gabapentin gastroretentive 1800 mg (n=299) Placebo (n=294)	Healthy postmenopausal women who experienced 7 or more moderate-to-severe hot flashes per day during a 14-day baseline.	Current treatment with MHT; history of gastric reduction; substance abuse within the past year; or any serious medical condition. Concomitant treatment of hot flashes except antidepressant with unchanged dosages were permitted.	Age: 54 (SD=6.1) Race/Ethnicity: 69.5% White, 26.3% Black, 1.2% Other, 3% Hispanic BMI (<30): 61.7% (range: 16.3-59.4) Natural menopause: 74.7%

BMI: body mass index, CR: Controlled release, DRSP: Drospirenone, DVT: deep vein thrombosis, DYD: dydrogesterone, FSH: follicle stimulating hormone, mg: milligram, MHT: Menopausal hormone therapy, MI: myocardial infarction, N: number, NETA: Norethisterone acetate, NR: Not Reported, SD: Standard Deviation, SERM: selective estrogen receptor modulators, SNRI: serotonin–norepinephrine reuptake inhibitor, SSRI: selective serotonin reuptake inhibitor, US: United States, VMS: vasomotor symptoms

* Includes both SNRI and MHT. In MHT, arm a progesterone taper was provided for 14 days after estrogen.

†Includes both MHT low and standard dose. N=1,835 for full study and N = 1,411 for 4 out of 5 trial arms excluding estradiol 0.25 mg and progesterone 50 mg.

‡Includes both MHT low and standard dose.

§N=2,118 for full study and N=365 for VMS substudy.

#Reddy et al. included a trial arm of 0.625 mg conjugated estrogens (N=20). Women were randomly allocated to the doses and no information was provided on the number of women without a uterus and thus we excluded this arm from our review.

¤Trial arm of 0.25 mg estrogen/50 mg progesterone was excluded due to it being <0.5 mg estradiol

3.2. Results

In this main report, we describe the change in VMS frequency and severity (primary outcomes) from baseline to 8 – 16 weeks after treatment initiation. To aid in comparison to the fezolinetant trials, we first report results from trials that measured changes in moderate-severe VMS then we report results from trials that measured changes in any VMS. In Table 3.2. we provide conversions to daily VMS and note if the change met threshold for minimum clinically important difference (MCID) for all interventions of interest. Specifically, we consider MCIDs for the following outcomes: VMS frequency (≥25 per week or 3.57 per day), VMS severity (≥0.225), and Menopause-Specific Quality of Life Questionnaire (MENQoL) (≥1).[73,74] MCID were derived from two studies that used response to the Clinical Global Impression (CGI) scale to detect minimal clinically important changes in VMS frequency, VMS severity, and MENQoL after randomization to MHT or placebo.[73,74] MCID >1 for MENQoL was also reported in other sources.[75] Additionally, we report changes in MENQoL scores, sleep disturbance, urogenital symptoms, and mood. Finally, harms and discontinuation rates are summarized with long-term harms for MHT. Additional studies/outcomes are available in Section D of the Supplement.

Clinical Benefits

Fezolinetant

The efficacy of fezolinetant compared with placebo for the treatment of moderate-severe VMS associated with menopause is presented for three Phase 3 trials (Skylight 1 and 2, and Moonlight 1).[25-28,55]

VMS Frequency

In the Skylight 1 and 2 trials, participants achieved significant reduction in moderate-severe daily VMS frequency at 12 weeks (Skylight 1: mean reduction versus placebo of -2.39 (SE=0.44), p<0.001 in the 30 mg dose and -2.55 (SE=0.43), p<0.001 in the 45 mg dose; Skylight 2: mean reduction versus placebo of -1.86 (SE=0.55), p<0.001 in the 30 mg dose and -2.53 (SE=0.55), p<0.001 in the 45 mg dose). For Moonlight 1, a press release from March 2022 reported fezolinetant 30 mg daily did not meet the pre-defined endpoint (change in moderate-severe VMS frequency and severity).[28] No additional data from Moonlight 1 were available at the time of publication of this report.

VMS Severity

Participants treated with fezolinetant had a significant reduction in moderate-severe VMS severity at 12 weeks (Skylight 1: mean reduction versus placebo of -0.24 (SE=0.08), p=0.002 in the 30 mg dose and -0.20 (SE=0.08), p=0.007 in the 45 mg dose; Skylight 2: mean reduction versus placebo of -0.16 (SE=0.08), p=0.049 in the 30 mg dose and -0.29 (SE=0.08), p<0.001 in the 45 mg dose)[25,54].

MENQoL

The efficacy of fezolinetant compared with placebo for changes on MENQoL was evaluated in the three Phase 3 trials (Skylight 1 and 2, and Moonlight 1). No published data for MENQoL is available from these trials. MENQoL data from Phase 2 trials is described in Section D of the supplement.

Menopausal Hormone Therapy

Across the MHT trials, the efficacy of MHT was clinically similar between the standard dose (1 mg) and low dose estrogen (0.5 mg), thus we included both doses in our review. On average, MHT results in statistically and clinically significant reductions in VMS frequency and severity, along with improvements in quality of life as measured by the MENQoL.

VMS frequency

In the five trials that reported changes in moderate-severe VMS frequency, all reported a significantly greater reduction in VMS frequency in the MHT group compared to placebo. Stevenson et al. (2010) reported a daily mean difference in MHT low dose versus placebo of -1.19; 95% CI: 0.52 to 1.86; p<0.001.[35] This reduction met the threshold for MCID but was a smaller difference than reported in the Skylight 1 and 2 trials (-2.55 to -1.86). Archer et al. (2013)[39] reported greater improvements in weekly moderate-severe VMS frequency in all the low estrogen dose arms at week 12 compared to placebo (mean difference to placebo: -22.2; 95% CI:-27.8 to -16.6; p<0.0001 and -27.6; 95% CI:-33.2 to -22.0; p<0.001); equating to a daily mean difference of -4.1 and -3.3. These difference met the MCID threshold and were larger than the difference reported in the Skylight 1 and 2 trials. In the REPLENISH trial, there was a greater reduction in weekly moderate-severe VMS at week 12 in the MHT groups (1 mg or 0.5 mg) (ranging from -55.1 to -49.85) compared to placebo (-40.2) (all p<0.05)[33]; equating to a daily mean difference of -2.2 to -1.4. More participants in the MHT doses had reductions that met MCID criteria at week 12 (68%-73% vs 52% placebo; p<0.05)[73] and had more days per week without moderate-severe VMS (p<0.05).[60] In the CHOICE trial, there was a larger decrease in weekly moderate-severe VMS in the low dose estrogen arms at week 12 compared to placebo, p<0.001 (daily mean difference of -3.3 and -3.0).[40] The reductions were larger than reported in the Skylight trials. Endrikat et al. (2007)[34] reported a larger percent reduction in the number of moderate-severe VMS per week in the MHT group (-80.8%) during weeks 5-12 compared to placebo (-41.5%), p<0.0001.[34], with a daily mean difference of -3.1 and participants in the MHT arm were experiencing around 1 VMS per day, compared to 4 per day in the placebo group.

There were five trials that reported changes in all (mild, moderate, and severe) VMS frequency. MsFLASH 03 trial reported a greater reduction in daily VMS frequency by week 8 in participants who received MHT compared to placebo (mean difference from placebo: -2.3; 95% CI: -3.4 to -1.3; p<0.001) which was comparable to the Skylight trials.[38] Schürmann et al. (2004)[36] reported a larger

percent reduction in the weekly frequency of VMS in all MHT doses (-88.0% to -84.5%) at week 16 as compared to placebo (-47.0%) (all p<.0001); equating to a daily mean difference of -4.1 to -3.0. Lin et al. (2011)[37] reported significantly greater percent reduction in VMS frequency at week 16 in MHT group (percentage change difference from placebo: -28.5%, p=0.0001); equating to a daily mean difference of -2.9. Finally, in the two estrogen-only trials, Speroff et al.[41] reported a greater reduction in weekly VMS frequency at week 12 in those who received two transdermal systems (0.04 mg per day) (study 1: -50.7 and study 2: -48.4 VMS per week) compared to placebo (study 1: -41.9 and study 2: -31.2 VMS per week) (p<0.001 and p=0.004, respectively). The converted daily mean difference ranged from -3.7 (0.04 mg) to -0.7 (0.02 mg). However, the reduction in VMS frequency for those who received 0.02 mg was only significantly different to placebo in study 2 (p=0.006), not study 1 (p=0.088).[41] See Table 3.2. for values converted into daily changes and Tables D3.4.-.5. for values reported in manuscripts.

VMS severity

In the five trials that reported VMS severity, four trials reported significantly greater improvements in weekly VMS severity in the MHT groups, compared to placebo. Archer et al. (2013)[39] reported greater improvements in VMS severity in the two low dose estrogen arms compared to placebo (mean difference to placebo: -0.80; 95% CI:-1.01 to -0.59; p<0.0001 and -1.07; 95% CI:-1.28 to -0.86; p<0.0001). This mean difference was larger than that reported in the Skylight trials and met threshold for MCID. In the REPLENISH trial, there were larger reductions in moderate-severe VMS severity at week 12 in the MHT doses (ranging from -1.12 to -0.76) compared to placebo (-0.56) (all p<0.05)[33] and more participants in the MHT doses had reductions that met MCID criteria at week 12, compared to placebo (39%-56% vs 29% placebo; p<0.05).[74] In the CHOICE trial, there was a greater decrease in VMS severity in the two MHT arms at week 12, compared to placebo (p<0.001).[40] In the MsFLASH 03 trial, there was a significantly greater reduction in VMS severity by week 8 in the MHT group as compared to placebo (mean difference: -0.3; 95% CI: -0.4 to 0.1; p=0.02). Conversely, Lin et al. (2011)[37], an RCT conducted exclusively in Chinese participants, reported no significant difference in reduction of moderate-severe VMS severity at week 16 in the MHT arm compared to placebo (p=0.103).

MENQoL

MENQoL scores had a potential range from 0 (no symptoms) to 174 (extremely bothered) and negative values for the MENQoL indicate improvement. In the two studies that examined MENQoL, there were greater reductions in total MENQoL scores in the MHT arms as compared to placebo at week 8 and 12. In the MsFLASH 03 study, there was a mean difference to placebo at week 8 of -0.5; 95% CI: -0.7 to 0.2; p<0.001 and only the MHT group met MCID threshold.[76,77] See Table D.3.11. In the REPLENISH trial, there was a mean reduction in the MHT arms from -1.62 to -1.92 (all <0.05), compared to -1.39 in placebo at week 12[62], with all groups meeting criteria for MCID. Though not

measured by MENQoL, Stevenson et al. reported a larger improvement in health-related quality of life in those in the MHT doses.[35]

Other outcomes

Three MHT trials reported a greater reduction in vaginal dryness in the MHT arms, compared to placebo.[36,37,40] Six trials reported improvements in sleep outcomes in the MHT groups compared to placebo.[35,36,39,40,59,78] But two trials, including MsFLASH 03, reported no improvements in insomnia.[37,78] There were inconsistent effects on depression.[36,37] Full details are found in Tables D3.6.-7. in the Supplement.

SSRIs/SNRIs

There were overall inconsistent results in the review of SNRI /SSRIs for the treatment of VMS. Although some trials reported statistically significant improvements in VMS, none of the antidepressants reviewed achieved clinically meaningful improvements in VMS frequency nor MENQoL when compared to placebo. Of the SNRI/SSRIs reviewed, desvenlafaxine had the most sizeable and consistent treatment effects on VMS severity when compared to placebo, but these trials included mild in their assessment of VMS severity.

VMS Frequency

The two paroxetine trials reported greater reductions in weekly VMS frequency in those who received paroxetine compared to placebo (Study 1: paroxetine: -43.5 and placebo: -37.3, p=0.009 and Study 2: paroxetine: -37.2 and placebo: -27.6, p=0.0001). [42] Although the converted daily difference (-1.4 and -0.9) was smaller than that reported in the Skylight trials (See Table 3.2). Escitalopram was associated with a greater reduction in daily VMS frequency by week 8 in the MsFLASH 01 trial (mean difference to placebo: -1.41; 95% CI: -2.69, -0.13; p<0.001).[43] There were limited and mixed findings for venlafaxine. MsFLASH 03 trial reported greater reductions in daily VMS frequency by week 8 in participants who received venlafaxine compared to placebo (mean difference versus placebo: -1.8; 95% CI: -2.7 to -0.8; p=0.005).[38] Evans et al. (2005)[44], reported no significant difference between participants in the extended-release venlafaxine and placebo group in reduction of daily VMS frequency at week 12 (p=0.20), though this trial included only 40 participants. The most consistent evidence for improvements in VMS frequency was reported in the desvenlafaxine trials. Four of the five main RCTs of desvenlafaxine reported significantly greater reductions in moderate-severe VMS frequency at week 12 in desvenlafaxine 100 mg compared to placebo, with change from baseline values ranging from -6.3 to -7.8.[45,46,48,49] Daily mean difference from placebo ranged from -2.8 to -1.3 for the 100 mg dose. Efficacy for 150 mg and 200 mg doses were less consistent. See Table 3.2. However, Bouchard et al. (2012)[47] reported no significant difference in the change in moderate-severe VMS frequency in the desvenlafaxine 100 mg group compared to placebo, p=0.92. The authors note that a potential explanation could be that this trial

was conducted outside of the US (in Europe, Mexico, and South Africa). However, this trial was mostly White (93%) and non-Hispanic (93%) and the other agent in the trial, tibolone (not examined in this current report) did demonstrate significant differences in VMS frequency and severity compared to placebo. Efficacy data at longer follow-up for some of these trials are reported in Section D3 of the Supplement.

VMS Severity

Evidence for paroxetine was mixed. Study 2 of Simon et al. (2011) reported greater reductions in weekly VMS severity for paroxetine compared to placebo at week 12 (-0.12 and -0.07, respectively; p=0.011), but study 1 reported no significant difference between the paroxetine and placebo group (p=0.29). Escitalopram was associated with a greater reduction in VMS severity at week 8 in the MsFLASH 01 trial (mean difference to placebo: −0.22; 95% CI:−0.40 to −0.05; p<0.001). Evidence on venlafaxine was limited to two trials and mixed. MsFLASH 03 trial reported greater reduction in VMS severity by week 8 in the venlafaxine group as compared to placebo (mean difference versus placebo: -0.2; 95% CI: -0.3 to 0.0; p=0.02), but Evans et al. (2005) reported no significant difference between the groups (p=0.30).[44] Again, the most consistent evidence for improvements in VMS severity was reported in the desvenlafaxine trials. Four out of five trials reported greater improvements in VMS severity in the desvenlafaxine 100 mg arms compared to placebo, with change from baseline values ranging from -0.54 to -0.88 and mean difference was -0.33, larger than the difference in the Skylight trials. Again, Bouchard et al. (2012)[47] reported no significant difference in the change in moderate-severe VMS severity in the desvenlafaxine 100 mg group compared to placebo (p=0.94).

MENQoL

Two trials with SSRI/SNRIs examined changes in MENQoL. There were significantly greater improvements in MENQoL at week 8 in those who received escitalopram (mean difference to placebo: -0.4; 95% CI: -0.6 to -0.1)[76] or venlafaxine (mean difference to placebo: -0.2; 95% CI: -0.4 to 0.0; p=0.04).[76,77] Although in the venlafaxine trial, this difference was driven by one significant subdomain: psychosocial symptoms and vasomotor domain scores slightly increased from baseline to week 8.[77] See Table D.3.11. Improvements in both trials did not meet MCID.

Other Outcomes

There were improvements in sleep reported in the escitalopram trial[79], but evidence was mixed for paroxetine[69] and venlafaxine.[78] There were no improvements in sexual functioning in the paroxetine trial, measured using the sexuality subscore of the hot flash related daily interference scale (HFRDIS) and the Arizona Sexual Experiences Scale[70] Despite Evans et al. (2005) reporting no difference in VMS frequency or severity, this trial did report a greater reduction in patient-reported interference of VMS with daily life in the venlafaxine group (-51%) compared to the placebo group (-15%) (p<0.001). However, those in the venlafaxine arm had higher scores on this measure at baseline.[44] See Tables D3.6.-.7.

Gabapentin

Trials of gabapentin, although demonstrating statistical significance, also failed to show clinically meaningful differences in VMS frequency or severity. MENQoL was not assessed in these trials.

VMS frequency

Two gabapentin trials quantitatively reported changes in VMS frequency (mild, moderate, and severe) and both reported larger reductions in those who received gabapentin compared to placebo.[51,53] Pinkerton et al. (2014) reported a greater reduction in daily VMS frequency in the gabapentin group at week 12 (mean difference from placebo of -1.14; 95% CI, -1.8 to -0.48; p= 0.001).[51] Guttuso et al. (2003) reported a greater percent reduction in daily VMS frequency at week 12 in those who were receiving gabapentin compared to placebo (mean difference from placebo: -20.9%; 95% CI 2.7, 34.0; p=0.02), with a daily mean difference of -1.9.[53] The confidence intervals associated with these group differences were wide and suggest large variance across participants. Efficacy data at longer follow-up for these trials are reported in Section D3 of the Supplement.

VMS severity

In all three gabapentin trials, there was a significant reduction in VMS severity compared to placebo.[52,53,72] Pinkerton et al. (2014) reported that those in the gabapentin group had a significantly larger reduction in VMS severity compared to the placebo group (mean difference, -

0.19; 95% CI: -0.33 to -0.04; p=0.012). Guttuso et al. (2003) reported a greater percent reduction in daily VMS severity at week 12 in those who were receiving gabapentin compared to placebo (-25.5%; 95% CI 6.8, 42.3; p=0.01), and Reddy et al. (2006) reported that gabapentin reduced VMS severity score by 52% at week 12 compared to 20% in placebo (t=3.03, p=.004).

Other outcomes

There were no differences in quality of life[53] nor mood.[52,53] Effects on sleep were mixed. Sleep interference improved more in the gabapentin group compared to placebo (p=0.0001) in the Pinkerton et al. (2014) trial [51], but Guttuso et al. (2003) showed no improvements in total sleep outcome, measured by the Pittsburgh Sleep Quality Index (PSQI). However, a secondary analysis of the Guttuso trial focusing on sleep domains reported that gabapentin significantly improved the sleep quality subdomain (p=0.03), but not the sleep efficiency or daily disturbance subdomains.[71] See Tables D3.6.-.7.

Table 3.2. Key Trial Results with Converted Values to Daily Reduction

Trial Name/Author	Intervention	Arm Size	VMS Frequency		VMS Severity		MENQoL, DIFF from PBO, 95% CI; P value
			Change from baseline: Mean (SE)	Difference from Placebo: Mean Change (SE/95% CI), P Value	Change from baseline: Mean (SE)	Difference from Placebo: Mean Change (SE/95% CI), P Value	
Fezolinetant							
SKYLIGHT 1 Lederman et al. 2022[54]	Fezolinetant 30 mg	173	NR	–2.39 (0.44), p<0.001	NR	–0.2 (0.08), p=0.007	NR
	Fezolinetant 45 mg	174	NR	–2.55 (0.43), p<0.001	NR	–0.24§ (0.08), p=0.002	NR
	Placebo	175	NR	REF	NR	REF	NR
SKYLIGHT 2 Johnson et al. 2021, Johnston et al. 2022 [25,26]	Fezolinetant 30 mg	166	NR	-1.86 (0.55), p<0.001	NR	-0.16 (0.08), p=0.049	NR
	Fezolinetant 45 mg	167	NR	-2.53 (0.55), p<0.001	NR	-0.29§ (0.08), p<0.001	NR
	Placebo	167	NR	REF	NR	REF	NR
MHT: standard dose (1 mg estradiol)							
Schurmann et al. 2004[36]	Estradiol 1 mg/DRSP 1 mg	55	-7.6†§	-3.6†§, p<.0001	NR	NR	NR
	Estradiol 1 mg/DRSP 2 mg	52	-7.8*†§	-3.0†§, p<.0001	NR	NR	NR
	Estradiol 1 mg/DRSP 3 mg	57	-8.1*†§	-4.1†§, p<.0001	NR	NR	NR
	Placebo	61	-4.0*†§	REF	NR	NR	NR
Endrikat et al. 2007[34]	Estradiol valerate 1 mg/dienogest 2 mg	162	-6.1†§	-3.1†, p<0.0001	NR	NR	NR
	Placebo	162	-3.0†	REF	NR	NR	NR
REPLENISH: Lobo et al. 2018[33]	Estradiol 1 mg and progesterone 100 mg (TX-001HR)	415	-7.9†§	-2.2†, p<0.05	-1.12§ (NR)	-0.56†§, p<0.05	-1.92§

Trial Name/Author	Intervention	Arm Size	VMS Frequency		VMS Severity		MENQoL, DIFF from PBO, 95% CI; P value
			Change from baseline: Mean (SE)	Difference from Placebo: Mean Change (SE/95% CI), P Value	Change from baseline: Mean (SE)	Difference from Placebo: Mean Change (SE/95% CI), P Value	
	Estradiol 0.5 mg and progesterone 100 mg (TX-001HR)	424	-7.7†§	-2.0†, p<0.05	-0.90§ (NR)	-0.34†§, p<0.05	-1.62§
	Estradiol 0.5 mg and progesterone 50 mg (TX-001HR)	421	-7.1†§	-1.4†, p<0.05	-0.76§ (NR)	-0.20†§, p<0.05	-1.9§
	Placebo	151	-5.7†§	REF	-0.56§ (NR)	REF	-1.39§
Stevenson et al. 2010[35]	Estradiol 1 mg/DYD 5 mg	59	-6.2 (2.6)§	NR	NR	NR	NR
	Estradiol 0.5 mg/DYD 2.5 mg	122	-6.3 (3.4)§	-1.19 (95% CI 0.53, 1.86) p<0.001	NR	NR	NR
	Placebo	124	-4.9 (3.5)	REF	NR	NR	NR
Lin et al. 2011[37]	Estradiol 1 mg/DRSP 2 mg	183	-6.6*†§	-2.9†, p=0.0001	-0.57§ (0.85)	-0.30§, p=0.103	NR
	Placebo	61	-3.7*†§	REF	-0.28§ (0.58)	REF	NR
Speroff et al. 1996 (Study 1 and 2)[41]	Estradiol transdermal system: 0.02 mg	54	-6.7†§#	-0.7†, p=0.088	NR	NR	NR
	Estradiol transdermal system: 0.04 mg	53	-7.2†§#	-3.7†§, p<0.001	NR	NR	NR
	Placebo (single dose)	54	-6.0†§#	REF	NR	NR	NR
	Placebo (double dose)	52	-3.5†§#	REF	NR	NR	NR
	Estradiol transdermal system: 0.04 mg	37	-6.9†§#	-2.4†, p=0.004	NR	NR	NR
	Estradiol transdermal system: 0.02 mg	37	-6.6†§#	-2.1†, p=0.006	NR	NR	NR
	Placebo (double dose)	37	-4.5†§#	REF	NR	NR	NR

Trial Name/Author	Intervention	Arm Size	VMS Frequency		VMS Severity		MENQoL, DIFF from PBO, 95% CI; P value
			Change from baseline: Mean (SE)	Difference from Placebo: Mean Change (SE/95% CI), P Value	Change from baseline: Mean (SE)	Difference from Placebo: Mean Change (SE/95% CI), P Value	
MHT: low dose (0.5 mg estradiol)							
CHOICE: Panay et al. 2007[40]	Estradiol 0.5 mg/0.1 mg NETA	194	-8.2†§	-3.0† p<0.001	NR	p<0.001	NR
	Estradiol 0.5 mg/0.25 mg NETA	181	-8.5†§	-3.3†, p<0.001	NR	p<0.001	NR
	Placebo	200	-5.2†§	REF	NR	REF	NR
Archer et al. 2013[39]	Estradiol 0.5 mg/DRSP 0.25 mg	177	-7.9†§	-3.3†, p<0.0001	-1.21§ (1.08)	-0.80§ (95% CI:-1.01 to -0.59), p<0.0001	NR
	Estradiol 0.5 mg /DRSP 0.5 mg	178	-8.6†§	-4.1†§, p<0.001	- 1.45§ (1.12)	-1.07§ (95% CI:-1.28 to -0.86), p<0.0001	NR
	Placebo	176	-4.6†§	REF	-0.39§ (0.77)	REF	NR
MsFLASH 03: Joffe et al. 2014[38]	Estradiol 0.5 mg	97	-4.5*§ (95% CI: -5.4, -3.6)	-2.3 (95% CI: -3.4, -1.3), p<0.001	NR	-0.3*§ (95% CI: -0.4, 0.1), p=0.02	-0.5, 95% CI: -0.7, 0.2, P<0.001
	Venlafaxine 75 mg	96	-3.9*§ (95% CI: -4.7, -3.1)	-1.8 (95% CI: -2.7, -0.8), p=0.005	NR	-0.2* (95% CI: -0.3, 0.0), p=0.02	-0.2, 95% CI: -0.4, 0.0, p=0.04
	Placebo	146	NR	REF	NR	REF	REF
SNRIs							
Evans et al. 2005[44]	Venlafaxine 75 mg	40	NR*	p=0.20	NR*	p=0.30	NR
	Placebo	40	NR*	REF	NR*	REF	NR
	Desvenlafaxine 100mg	157	–7.23§ (0.38)	-1.76, p=0.003**	–0.80*§ (0.06)	-0.33§, p=0.006**	NR

Trial Name/Author	Intervention	Arm Size	VMS Frequency		VMS Severity		MENQoL, DIFF from PBO, 95% CI; P value
			Change from baseline: Mean (SE)	Difference from Placebo: Mean Change (SE/95% CI), P Value	Change from baseline: Mean (SE)	Difference from Placebo: Mean Change (SE/95% CI), P Value	
Speroff et al. 2008; Wywich et al. 2008¤ [46,50]	Desvenlafaxine 150 mg	163	–6.94§ (0.38)	-0.96, p=0.11¤	–0.59*§ (0.07)	-0.09, p=0.47¤	NR
	Desvenlafaxine 200 mg	155	–6.46§ (0.41)	-0.88, p=0.15¤	–0.74*§ (0.07)	-0.25§, p=0.04¤	NR
	Placebo	78	-5.50§ (0.46)	REF	–0.47*§ (0.09)	REF	NR
Archer et al. 2009 [45]	Desvenlafaxine 100 mg	150	-7.1§ (0.34)	-1.3, p=0.005	-0.65*§ (0.07)	-0.3†§, p<0.001	NR
	Desvenlafaxine 150 mg	151	-7.0§ (0.35)	-1.2, p=0.012	-0.66*§ (0.07)	-0.3†§, p<0.001	NR
	Placebo	150	-5.8§ (0.34)	REF	-0.33*§ (0.07)	REF	NR
Archer et al. 2009 [49]	Desvenlafaxine 100 mg	182	-6.3§ (0.34)	-1.4, p=0.002	-0.54*§ (0.07)	-0.3†§, p=0.002	NR
	Desvenlafaxine 150 mg	179	-7.0§ (0.35)	-2.1, p<0.001	-0.71*§ (0.07)	-0.4, †§, p<0.001	NR
	Placebo	180	-4.9§ (0.31)	REF	-0.28*§ (0.06)	REF	NR
Bouchard et al. 2012 [47]	Desvenlafaxine 100 mg	137	-5.78§ (0.33)	0.04, p=0.921	-0.61*§ (0.07)	0.0†, p=0.943	NR
	Placebo	150	-5.82§ (0.31)	REF	-0.61*§ (0.07)	REF	NR
Pinkerton et al. 2013 [48]	Desvenlafaxine 100 mg	158	-7.5§ (0.35)	-2.48 (95% CI:-3.47, -1.50), p<0.001	-0.63*§ (0.05)	-0.33§ (95% CI:-0.48, -0.18), p<0.001	NR
	Placebo	156	-5.0§ (0.35)	REF	-0.3*§ (0.05)	REF	NR
SSRIs							

Trial Name/Author	Intervention	Arm Size	VMS Frequency		VMS Severity		MENQoL, DIFF from PBO, 95% CI; P value
			Change from baseline: Mean (SE)	Difference from Placebo: Mean Change (SE/95% CI), P Value	Change from baseline: Mean (SE)	Difference from Placebo: Mean Change (SE/95% CI), P Value	
MsFLASH 01: Freeman et al. 2011[43]	Escitalopram 10 mg	104	-4.60*§ (95% CI:−5.47 to −3.74)	−1.41* (95% CI:−2.69 to −0.13), p<.001	−0.52*§ (95% CI:−0.64 to −0.40)	−0.22 (95% CI:−0.40 to −0.05), p<0.001	-0.4, (95% CI: -0.6, -0.1), P<0.001
	Placebo	101	−3.20* (95% CI:−4.15 to −2.24)	REF	−0.30*§ (95% CI:−0.42 to −0.17)	REF	REF
Simon et al. 2013 (Study 1)[42]	Paroxetine (7.5 mg)	301	-6.2*†§	-0.9†, p=0.009	-0.10* (NR)	-0.1†, p=0.29	NR
	Placebo	305	-5.3*†§	REF	-0.09* (NR)	REF	NR
Simon et al. 2013 (Study 2)[42]	Paroxetine (7.5 mg)	284	-5.3*†§	-1.4†, p=0.0001	-0.12* (NR)	-0.05†, p=0.01	NR
	Placebo	284	-3.9*†§	REF	-0.07* (NR)	REF	NR

Gabapentin							
Guttuso et al. 2003[53]	Gabapentin 900 mg	30	-4.9*†§	-1.9†, p=0.02	-24.0‡†	-11.8‡†, p=0.01	NR
	Placebo	29	-3.0*†	REF	-12.25‡†	REF	NR
Reddy et al. 2006[52]	Gabapentin 2400 mg	20	NR	NR	NR*	t=3.03, p=.004	NR
	Placebo	20	NR	NR	NR*	NR	NR
Pinkerton et al. 2014[51]	Gabapentin gastroretentive 1800 mg	299	-7.64*§ (NR)	-1.14 (95% CI: 1.8 to -0.48), p=0.0007	-0.65*§ (NR)	-0.19 (95% CI: -0.33 to -0.04), p=0.012	NR
	Placebo	294	-6.50*§ (NR)	REF	-0.46*§ (NR)	REF	NR

CI: Confidence Interval, DIFF from PBO: Difference from placebo, DRSP: Drospirenone, DYD: dydrogesterone, mg: milligrams, MHT: Menopausal hormone therapy, N: total number of participants, NETA: Norethisterone acetate, NR: Not Reported, REF: Reference group, SD: Standard Deviation, SE: Standard Error, SNRI: serotonin–norepinephrine reuptake inhibitor, SSRI: selective serotonin reuptake inhibitor, VMS: Vasomotor symptoms

*All VMS (mild, moderate, and severe)

†Values converted to daily changes

‡VMS severity was rated from 1 (mild) to 4 (very severe) and multiplied by the number of VMS events at that level.

§VMS frequency (≥25 per week or 3.57 per day), VMS severity (≥0.225), and Menopause-Specific Quality of Life Questionnaire (MENQoL) (≥1).[73,74]

#Change from week 1 to week 12.

¤Data from Wyrwich et al. (2008)

Harms

Fezolinetant

The most frequent adverse event in both the Skylight 1 and 2 trials was headache.[25,26] In the Skylight 2 trial, discontinuation was slightly higher in the fezolinetant arms (40 mg: 3% and 30 mg 1.2%) compared to placebo (0.6%); consistent with the Phase 2 trials. There were five reports of serious adverse events, all in the fezolinetant groups. At 52 weeks, more participants in the fezolinetant 45 mg group had increased alanine transaminase (ALT) or aspartate aminotransferase (AST) values >3 times the upper limit of normal compared to the 30 mg group (4.3% in fezolinetant 45 mg group vs. 1.8% in 30 mg fezolinetant group), with one participant in each dose group reporting AST or ALT >5 times the upper limit of normal. For participants who had received placebo for the first 12 weeks and then received fezolinetant (30 or 45 mg) for the final 40 weeks of the open-label extension phase, 2.7% in the 45 mg fezolinetant group had ALT/AST values >3 times the upper limit of normal compared to none in the 30 mg group.[26] See Table D3.8.-9. for detailed harms results. Limited detailed information was provided for adverse events for Skylight 1 and adverse event data was unavailable for Moonlight 1. Data on long-term harms is limited. A press release reported that Moonlight 3, a single-arm Phase 3 trial that examined safety and tolerability of fezolinetant 30 mg at 52 weeks, met its primary endpoint for frequency and severity of adverse events.[29]

Menopausal Hormone Therapy

Adverse Events from RCTs

There were no clear differences in the adverse events reported in the standard estrogen dose (1 mg) and low estrogen dose (0.5 mg) groups. Adverse events in both dose groups were mostly mild to moderate in severity and there were generally no significant differences in discontinuation between the MHT and placebo groups. In the two RCTs that evaluated transdermal MHT, there were few discontinuations due to adverse events, with no difference between MHT and placebo groups, and most skin reactions were mild or moderate in severity. Uterine bleeding and breast pain were more common in the MHT doses, but occurred infrequently overall and less often in the low-dose MHT trials.[33,35,38,40] Serious events in trials of less than one year in length were low and standard dose trials had slightly higher incidences of serious events. Full details of adverse events are described in Table D3.8-9. in the Supplement.

Long-Term Harms of MHT

To evaluate long-term adverse events for MHT, we identified two meta-analyses and two pooled analyses that provided risk estimates beyond 3 years.[80-83] The majority of studies included in the four identified sources were from the WHI trials and included a standard dose of MHT. See Table

D3.10. for details. Data from long-term outcomes of low dose MHT is limited. Including women of all ages (mean/median age: 48-76 years), Marjoribanks et al. (2017) conducted a meta-analysis including 22 trials with 43,637 participants from the WHI study. For the combined estrogen/progesterone doses, the mean follow-up ranged from 3-8 years and the risk ratios were >1 for: stroke, breast cancer, gallbladder disease, venous thromboembolism (VTE), and death (all-cause). Of note, the risk ratio for all-cause death was 3.58 for the standard dose but lower (1.00-1.06) for estimates based on higher doses of estrogen (greater than 1 mg), presumably as this estimate was based on more trials of longer duration (e.g., >7.9 years).[81] The risk ratios were <1 for all clinical fractures, suggesting a protective element of MHT. Data for cardiovascular outcomes were only available at one year. Thus, we supplemented Marjoribanks et al. with a meta-analysis from Kim et al. (2020) for these outcomes. Kim et al. (2020) included 26 RCTs of MHT with a median follow-up of 3.4 years and summary estimates based on RCTs were >1 for: stroke, VTE, pulmonary embolism, myocardial infarction, and coronary heart disease. All-cause death was 1.00 and estimates were <1 for cardiovascular disease, angina, and revascularization.[82] When examining only estrogen/progesterone in the observational studies, estimates were higher for venous thromboembolism (2.21), but lower for all-cause death (0.61) and myocardial infarction (0.77); though these estimates were based on fewer studies.

Risk estimates described above were for women across a broad age range, risk estimates in those aged 50-59 years may more accurately represent the group of women who are making treatment decisions for menopausal symptoms. Prentice et al. (2021) examined two RCTs including a total of 27,347 participants from the WHI trials who were aged 50-59 years. For the combined estrogen/progesterone trial, the median follow-up for the intervention phase was 5.6 years and 18 years for the cumulative phase. For the intervention phase, hazard ratios were >1 for: coronary heart disease, breast cancer, stroke, and pulmonary embolism, and were <1 for: colorectal cancer, hip fracture, and death. These hazard ratios did not change significantly for cumulative follow-up.[80] We supplemented Prentice et al. (2021) with an age subgroup analysis from WHI for coronary heart disease specifically. Rossouw et al. (2007) reported that CHD risk was <1 for women up to the age of 69, and 1.26 for women aged 70-79 years. Thus, risk was lower in women aged 50-59 years.[83] Full details of these estimates are found in the Table D3.10. in the Supplement.

SSRIs/SNRIs

Adverse events of any cause in the SSRI trials and venlafaxine trials were mostly mild or moderate in severity.[38,42-44] See Table 3.3. for most common adverse events. Across the majority of the desvenlafaxine trials, participants in the active treatment group had more adverse events than placebo, and these events were highest during the first week of treatment and with increasing dose.[46] Discontinuation rates differences between desvenlafaxine and placebo groups were mixed, but discontinuation differences were more prominent when comparing 150 mg and 200 mg dose with placebo.[45,46] Serious AEs were reported in all six desvenlafaxine trials. Additional details of

adverse events are described in Section D2 of the Supplement. Long-term harms data on SSRIs/SNRIs for treatment of VMS are limited but these agents have been used for other conditions (e.g., depression) and present no long-term safety concerns.

Gabapentin

There were more total adverse events in gabapentin compared to placebo.[51,53] However, these adverse events were mostly mild-moderate and there were very few serious AEs reported across the trials. Rates of discontinuation due to adverse events were marginally higher in the gabapentin arms compared to placebo.[51,53] See Table D3.8-9 in the Supplement for detailed harms results. Long-term harms data for treatment of VMS are limited but, when used for other conditions (e.g., seizures) present no long-term safety concerns.

Table 3.3. Adverse Events

Drug	Most Common Adverse Event Greater Than Placebo
Fezolinetant	Headache; in larger doses, elevated liver enzymes (ALT and AST)
MHT	Uterine bleeding (more reports of various serious adverse events)
SSRI/SNRI: Desvenlafaxine	Nausea, dry mouth, constipation, fatigue
SSRI/SNRI: Venlafaxine	Dry mouth, fatigue, decreased appetite
SSRI/SNRI: Paroxetine	Nausea, fatigue
SSRI/SNRI: Escitalopram	No adverse events greater than placebo arms
Gabapentin	Dizziness, headache, and somnolence

ALT: alanine transaminase, AST: aspartate aminotransferase, MHT: Menopausal hormone therapy, SNRI: serotonin–norepinephrine reuptake inhibitor, SSRI: selective serotonin reuptake inhibitor

Subgroup Analyses and Heterogeneity

In the trials of the agents in this review, few examined subgroup effects. In the Skylight 1 and 2 trials, there were no differences in efficacy of fezolinetant on VMS frequency and severity by subgroup defined by baseline age and VMS severity.[55] See Table D3.12. In the three MHT trials that reported subgroup effects, one reported no subgroup effect for age, race/ethnicity, and BMI[38] and two trials reported a subgroup effect of BMI on VMS.[58,63] In Stevenson et al. (2010), those who had a BMI of 30 kg/m^2 or greater and were prescribed MHT were not significantly different to placebo on change in VMS frequency at week 13.[63] In the REPLENISH trial[33], those with a BMI between 25kg/m^2 and 30 kg/m^2 who were prescribed low dose estrogen were not significantly different from placebo on change in VMS frequency and severity at week 12.[58] There was no subgroup effect of age in either trial.[58,63] See Tables D3.14-.15. Many trials recruited participants with both natural and surgical menopause, but no separate subgroup analyses were conducted on these groups. In the two SSRI/SNRI trials that recruited patients with a history of cancer, there appeared to be little difference in terms of safety and efficacy from trials in healthy postmenopausal women.[84,85]

We sought subpopulation data from the manufacturer on the effectiveness of the fezolinetant in subgroups of interest such as race and ethnicity. Data was not provided. We highlight the low representation of Black participants in the SSRI trials and, given the incidence of VMS in Black participants, the low representation of these participants across all trials. This is concerning because individuals from minority groups, particularly Black and Native American women, typically have the longest duration of symptoms.[52,53,54]

Uncertainty and Controversies

Comparability of Study Populations and Generalizability of Study Results

The included studies of fezolinetant and comparators are broadly similar in demographics and clinical characteristics (Table 3.1.). Across included intervention and comparator studies, the mean age and BMI were in a narrow range of 50 to 57 years of age and 25 to 29 kg/m^2, respectively. All included fezolinetant trials and all but two comparator trials[43,86] were conducted among a predominantly (60% to 100%) White population. However, as of the writing of this report, baseline characteristics of the Moonlight 1 study, which was conducted in China, Korea and Taiwan have not yet been reported. Baseline daily moderate to severe VMS frequency was similar between fezolinetant (11) and comparators (7 to 12). However, fezolinetant trials included fewer women with natural menopause. 58% and 63% of women in Skylight 1 and 2 had natural menopause. In contrast, the included comparator studies were comprised of approximately 75% to 100% women with natural menopause, aside from one study conducted among women with breast cancer[87] and a RCT conducted in Germany[34] which had lower rates of women with natural menopause.

The demographics cited above are expected to be similar to women who are likely to receive treatments of interest for VMS in the US. In a large national wide observational study, [13] the mean age of women with frequent VMS was 50 and more than half were overweight or obese. Similarly, approximately 62% of Americans identify as White race only[88].

Uncertainty Regarding the Comparability of Outcomes Measures Across Trials

While the population characteristics were largely comparable across trials, the definitions of our primary outcomes of VMS frequency and severity differed across trials, making cross-trial comparisons more difficult. For example, some studies defined their frequency outcome as a change in moderate-severe VMS only (fezolinetant, desvenlafaxine, and around half of the MHT trials) while the others defined their frequency outcome as a change in VMS of any severity (mild, moderate, severe). In this report, we aimed to examine trials that were most comparable in terms of VMS measurement. However, this limited the number of trials we could include in the evidence assessment, particularly for SSRI/SNRIs. Where trials differed in measurement, we noted this explicitly, and these differences added to the uncertainty of the results compared to the fezolinetant trials. For example, while most studies used the mild, moderate, and severe scale for VMS severity, some studies (e.g., of paroxetine and venlafaxine) had an additional category of very severe.[44,89,90] Only one of these studies defined this 4-point scale. Additionally, the categories of severity were defined by duration, physical and emotional symptoms, and action required and are therefore not comparable in each category to the mild, moderate, and severe scale.[91] Additionally, as described in Table A.1., most studies that calculated a VMS severity score did so in a way that was different from the fezolinetant trials. For example, only one included MHT trial [39] resembled fezolinetant in VMS calculation. Trials also reported changes from baseline using different values (e.g., absolute vs. relative reduction) and, due to limited reporting, in some cases it was not feasible to convert these into the same calculation. Furthermore, the mode of data collection for VMS frequency/severity differed across trials. For example, electronic diaries to record VMS in real-time[68] or retrospectively recording VMS one or twice a day using daily diaries.[43,64] Future studies may utilize electronic real-time assessments, which may overcome recall bias, but consistency in the measurement of outcomes is also critical to increasing the comparability of trials.

Uncertainty Regarding Efficacy of Fezolinetant

Astellas, the manufacturer of fezolinetant, is seeking approval for fezolinetant 45 mg based primarily on efficacy data from the Skylight 1 and Skylight 2 trials.[32] These trials were conducted in the U.S., Canada and Europe with 30 mg and 45 mg treatment arms, with findings of statistically significant improvements in VMS frequency and severity in both arms compared to placebo. However, the Moonlight 1 trial, another Phase 3 study that was conducted in China, Korea and Taiwan, failed to find statistically significant improvements in the VMS frequency and severity compared for the 30 mg dose compared to placebo. It is unclear whether the difference in findings between the Moonlight and Skylight trials at the 30 mg dose were due to type 1 error (that is, the statistically significant finding for the Skylight trials were spurious) or due to differences in characteristics of the enrolled populations such as demographics (e.g. the Skylight trials enrolled fewer than 1% Asian participants), diet, health and health behaviors. However, the discrepancy in trial outcomes results in increased uncertainty about the efficacy of fezolinetant.

Further, although the change in VMS frequency was above the MCID threshold in the Skylight trials (MCID threshold: ≥25 per week or 3.57 per day)[73], the change in VMS severity was not (MCID threshold: ≥0.225).[74] Therefore, it is unknown to what degree the observed improvements in VMS frequency and severity translate to improved patient quality of life. Supporting this concern is that the improvement in the Menopause-Specific Quality of Life (MENQOL) for fezolinetant (30 mg) in the Phase 2 trial was small, not significantly different to placebo, and did not meet MCID threshold.

Uncertainty Regarding Safety of MHT

The Women's Health Initiative (WHI) trials established the increased risk of hormone therapy, including coronary heart disease, stroke, venous thromboembolism, and breast cancer among women with a mean age of 63.[92-94] Since the publication of these studies in the early 2000's, there has been a substantial and sustained reduction in use of hormone therapy among women over the age of 40.[95] However, women who experience VMS are on average younger than women enrolled in the WHI studies, with the average age of onset of VMS being 47.1 years.[1] In our review of the literature and in our discussions with clinical experts we observed that the risks with MHT are in general lower among this younger age group. Hence, the many professional organizations [5 16 96,97 98] currently support offering MHT as first line or gold standard treatment of VMS especially for women who are younger than 60 years of age or are less than 10 years post-menopause, who do not have contraindications or excess cardiovascular or breast cancer risks, and after discussions of risks and benefits with the patient. Other factors may also influence MHT risk, such as route of administration, dose, combination of hormones, and duration of symptoms, and these factors have not been examined in detail within longitudinal studies.

More evidence is needed on the safety of MHT in real world usage as such populations who may be receiving lower doses, be younger and have fewer comorbidities than individuals enrolled in the WHI studies.

Lack of Evidence on Long Term Efficacy and Safety of Fezolinetant

There is uncertainty regarding long term efficacy and safety of most non-hormonal treatments. The median total duration of moderate to severe VMS is 9.4 years. In comparison, the longest placebo-controlled trials were relatively short for fezolinetant[99] (Skylight 1 and 2: 12 weeks) and gabapentin[72] (BREEZE 3: 12 weeks). However, because gabapentin has been used extensively for other indications, long-term safety data are available and no concerning findings have arisen. Although fezolinetant now has 52 weeks of uncontrolled data to inform efficacy and safety, it is a first-in-class medication. Since there are no FDA approved selective neurokinin-3 (NK3) receptor antagonists, we cannot rely on data from medications in the same therapeutic class or from other indications. Post-market safety events (e.g., black box warnings, REMS) for new therapeutics are common, occurring in about one-third of new approvals[100], underscoring the need for long-term safety data.

Lack of evidence on efficacy and safety among clinical and racial and ethnic subgroups

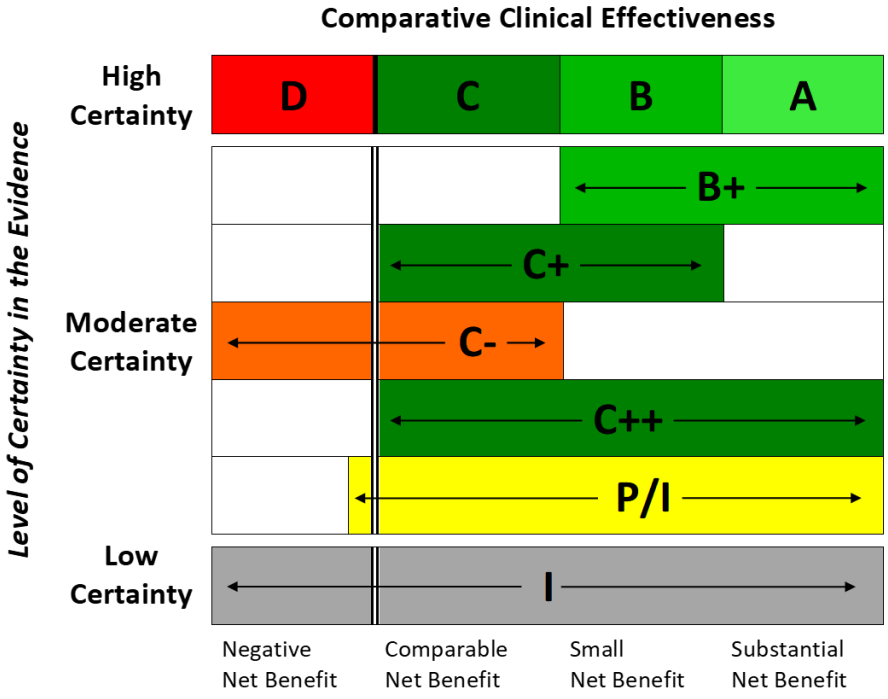
Another source of uncertainty is the lack of evidence on efficacy and safety among subgroups for the treatments of interest. Individuals with natural and surgically induced menopause differ in age and comorbidity. While many trials included participants with both natural and surgically induced menopause, no subgroup analyses were conducted to evaluate whether there were differences in efficacy or safety based on type of menopause.

VMS duration and severity are known to differ by race and ethnicity.[3] However, no studies included in our review reported on treatment effects by race and ethnicity. The difference in efficacy findings between the 30 mg fezolinetant dose among Asian populations (Moonlight trial) and predominantly non-Asian populations (Skylight trials) raises the potential that such differences may exist.

3.3. Summary and Comment

An explanation of the ICER Evidence Rating Matrix (Figure 3.1) is provided here.

Figure 3.1. ICER Evidence Rating Matrix



A = "Superior" - *High certainty of a substantial (moderate-large) net health benefit*
B = "Incremental" - *High certainty of a small net health benefit*
C = "Comparable"- *High certainty of a comparable net health benefit*
D= "Negative"- *High certainty of an inferior net health benefit*
B+= "Incremental or Better" – *Moderate certainty of a small or substantial net health benefit, with high certainty of at least a small net health benefit*
C+ = "Comparable or Incremental" - *Moderate certainty of a comparable or small net health benefit, with high certainty of at least a comparable net health benefit*
C- = "Comparable or Inferior" – *Moderate certainty that the net health benefit is either comparable or inferior with high certainty of at best a comparable net health benefit*
C++ = "Comparable or Better" - *Moderate certainty of a comparable, small, or substantial net health benefit, with high certainty of at least a comparable net health benefit*
P/I = "Promising but Inconclusive" *- Moderate certainty of a small or substantial net health benefit, small (but nonzero) likelihood of a negative net health benefit*
I = "Insufficient" – *Any situation in which the level of certainty in the evidence is low*

Fezolinetant vs. No Pharmacologic Treatment (Prescription nor Non-prescription)

In two large, unpublished RCTs (Skylight 1 and 2), fezolinetant showed improvements in VMS frequency and severity. The improvements were consistent across subgroups defined by age and baseline frequency and severity. However, the non-statistically significant results from the Moonlight 1 trial increases the uncertainty about the efficacy of fezolinetant. Further, even in the Skylight trials, fezolinetant failed to achieve a MCID in improvement of VMS severity above no

pharmacologic treatment. Also, since VMS typically lasts many years, long-term efficacy of treatments is relevant, but there are only 12 weeks of RCT data for fezolinetant. Finally, while there were significant improvements in other patient important outcomes (sleep inference, total climacteric symptoms, and global functioning) in Phase 2 trials, such information has not yet been reported for Phase 3 trials.

In terms of safety, fezolinetant was well tolerated and liver injury only occurred in higher doses (≥60 mg). Unblinded non-controlled extensions of Skylight 1 and 2 totaling 52 weeks and a 52-week single arm study of fezolinetant 30 mg support the longer term safety of fezolinetant. However, we point out that fezolinetant possesses a unique mechanism of action and there is no safety data from other drugs in its class to further support its long-term safety. For these reasons, we have assigned an ICER Evidence Rating for the overall net health benefits of fezolinetant versus no pharmacologic treatment of "Promising but inconclusive" (P/I).

Fezolinetant vs. MHT

In several large, peer-reviewed RCTs, MHT showed improvements in VMS frequency and severity. The point estimates for VMS frequency improvements tended to be larger for MHT than for fezolinetant, with MHT tending to reduce moderate to severe VMS by approximately one additional episode per day compared to fezolinetant. In the one study where the VMS severity score was calculated in a comparable manner, [39] MHT provided approximately 0.6 to 0.8 further reduction VMS severity score compared to fezolinetant. Further, MHT may provide additional benefits in terms of improving sleep, decreasing vaginal dryness, and preventing fractures. In terms of subgroups, those with higher BMI may have less improvement on MHT than the average MHT user.

Over the short-term, there were few adverse events in RCTs and most were mild-moderate. Longer term use of MHT may result in serious increased risks including coronary heart disease, stroke, venous thromboembolism, breast cancer and mortality, particularly among women ≥60 years old. We expect that most women who are using MHT would be <60 years old since the mean age of onset of VMS being 47.1 years[1], the median age of menopause is 51.4 and the median duration of moderate to severe VMS after menopause is 4.0 years.[1] We recognize that some women may be at higher risk of harms from MHT due to underlying conditions or older age, and in such cases, fezolinetant may be an alternative given its balance of benefits and harms. However, there are no studies that directly compare fezolinetant to MHT and only limited data are currently available to assess the overall effectiveness and long-term of safety of fezolinetant. For these reasons, we have assigned an ICER Evidence Rating for the overall net health benefits of fezolinetant vs. MHT of "Insufficient" (I).

Because of heterogeneity in studies and lack of high quality data for SSRI/SNRI, gabapentin, and pregabalin, we have not assigned ICER ratings for these drugs.

Table 3.4. Evidence Ratings

Treatment	Comparator	Evidence Rating
Fezolinetant	No pharmacologic treatment	P/I
Fezolinetant	MHT	I

MHT: menopausal hormone therapy

4. Long-Term Cost Effectiveness

4.1. Methods Overview

We developed a *de novo* decision analytic model for this evaluation, informed by key clinical trials and prior relevant economic models.

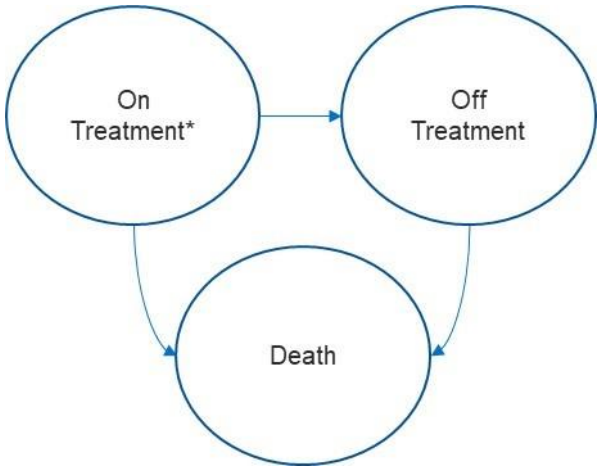
The model was focused on an intention-to-treat analysis, with a hypothetical cohort of women with VMS associated with menopause being treated with fezolinetant and comparators. Emphasis was placed on women who cannot or do not wish to take menopausal hormone therapy (MHT). The health outcome of each intervention was evaluated in terms of symptom improvements (e.g., using the Menopause-Specific Quality of Life [MENQoL] Questionnaire), life-years gained, quality-adjusted life years (QALYs) gained, and equal value of life years (evLYs) gained. The base-case comparison of this analysis focused on fezolinetant versus no pharmacologic treatment (prescription nor non-prescription) as estimated by the placebo arm of fezolinetant clinical trials. Key scenario analyses included MHT versus no pharmacologic treatment. Due to low quality or insufficient evidence, we did not compare SNRI/SSRI or gabapentin to no pharmacologic treatment, and we did not compare fezolinetant to other active comparators. Results were expressed in terms of the incremental cost per QALY gained, cost per evLY gained, and cost per symptom-free day. 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 thereafter. Costs and outcomes were discounted at 3% per year.

The cohort of patients was assigned to three mutually exclusive and exhaustive health states (Figure 4.1.): 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. All patients in the model began on treatment and responding to treatment. The first model cycle included treatment costs for all patients regardless of response to treatment until discontinuation due to the intervention occurred at the end of the response assessment period (e.g., 1 year). Long-term discontinuation due to symptom resolution was based on duration of VMS over the menopause transition.[13] Specifically, health state occupancy was derived using survival extrapolation methods of the proportion of women with and without VMS during the menopause transition using Kaplan-Meier curves. Frequency of VMS at baseline and reductions in frequency of VMS from treatment were tracked in the model to calculate symptom frequency and reductions in symptom frequency from treatment as a supplement to the health-related quality of life benefits estimated by the QALY and evLY. Patients remained in the model until death. All patients transitioned to death from all causes from any of the alive health states.

Key model inputs included clinical probabilities, quality of life values, and health care costs. Treatment effectiveness was estimated using evidence from trials of fezolinetant trials and relevant comparators derived from the clinical effectiveness section.

Health outcomes and costs were dependent on time spent in each health state, clinical events, adverse events (AEs), and direct medical costs. Quality-of-life weights were applied to each health state, including quality-of-life decrements for reasons such as serious adverse events. The model included direct medical costs, including but not limited to costs related to drug administration, drug monitoring, condition-related care, and serious adverse events. In addition, productivity changes and other indirect costs were included in a separate analysis representative of a modified societal perspective.

Figure 4.1. Model Schematic



*In some cases, there may be assignment of on treatment and not responding where treatment and health state costs are incurred with no gain in health benefits.

*Target Popul*ation

The population of focus for the economic evaluation includes patients comparable to those in fezolinetant clinical studies (Table 4.1.). This population includes women with an approximate age of 54 years with a wide range of months since onset of VMS.

Table 4.1 Baseline Population Characteristics

Characteristic	Total
Mean Age (SD), years	54.3 (5.0)
Black/African American (%)	20%
Hispanic or Latina (%)	22%
Mean, median duration of VMS (years)	8.8, 9.4
Baseline daily VMS frequency per 24 hours (range of mean)	9-12
Source	Kimball et al. Skylight 2, ENDO 2022; Fraser et al. Menopause 2020; Freeman et al. [25]

SD: standard deviation, VMS: vasomotor symptoms

Interventions

The list of interventions was developed with input from patient organizations, clinicians, manufacturers, and payers on which treatments to include. The primary intervention for this analysis is:

- Fezolinetant (Astellas Pharma, Inc.)

Comparators

The base-case comparison focused on no pharmacologic treatment as estimated by the placebo arm of fezolinetant clinical trials. For the purposes of adding context to the fezolinetant versus no pharmacologic treatment comparison, treatments currently being used for reducing VMS associated with menopause were also compared to no pharmacologic treatment, including MHT. MHT was included as a key scenario that may aid in the interpretation of fezolinetant's cost effectiveness. Comparators are as follows:

- No pharmacologic treatment (prescription nor non-prescription)
- MHT

4.2. Key Model Assumptions and Inputs

The base-case analysis took a health care system perspective and focused on direct medical care costs only. Outcomes were estimated over a lifetime time horizon to capture the potential impacts of short-term and ongoing morbidity and mortality. Costs and outcomes were discounted at 3% per year. Model assumptions are described in Table 4.2.

Table 4.2. Key Model Assumptions

Assumption	Rationale
Duration of treatment in the model will be consistent with VMS duration and assumed the same for all treatments.	There is no evidence available on the expected duration of treatment with fezolinetant. Assuming the same treatment duration for all treatments will not impact incremental health outcomes or costs.
The effectiveness of fezolinetant and comparators does not wane over time. Fezolinetant and its comparators have no residual benefits after stopping therapy.	There is no evidence suggesting treatment effects would change over longer treatment durations assuming response to treatment. Additionally, there is no evidence on durability of benefit beyond discontinuation of fezolinetant and comparators.
Relative treatment effects are consistent across baseline VMS severity and frequency.	There is limited evidence on relative treatment effects across starting levels of VMS severity and frequency, therefore we will not adjust treatment effects in subgroup analyses that may change the baseline severity or frequency of VMS.
Patients not responding to fezolinetant or other active treatments will not be treated/re-treated with the other treatments for VMS.	The purpose of this analysis is to isolate the cost-effectiveness of first line active treatment with fezolinetant and MHT. Although some patients may opt to try alternative active treatments after first treatment failure, effectiveness evidence is lacking within this subpopulation. Reverting to no pharmacologic treatment should not adversely impact incremental costs or outcomes for fezolinetant or comparators.
Patients can discontinue from lack of efficacy in the short-term (through response assessment in early cycles) and when underlying symptoms resolve.	By assuming that patients who respond to therapy remain on therapy until their underlying symptoms resolve, we anticipate that the costs versus the benefits of treatment will remain consistent with those observed in the fezolinetant trials.
Risks of MHT will be modeled as aggregate events that impact costs, utility, and mortality but not included as health states.	Such events can be tracked outside of health states without loss of generality.
Starting age does not impact trajectory of VMS or the VMS-related benefit of treatment.	The menopause transition can occur at various ages and the trajectory of symptoms are not impacted by starting age.[14]

Model inputs were estimated from the clinical review, published literature, and information from stakeholders. Key model inputs are shown in Table 4.3. These model inputs include changes in the MENQoL which are inputs to utility scores, utility scores by treatment, discontinuation during the first year of the model, and model-wide inputs such as duration of VMS and treatment.

Table 4.3. Key Model Inputs

Parameter	Fezolinetant	MHT	Placebo	Sources
Change in total MENQoL score versus placebo (95% CI)	[redacted]	-0.42 (-0.23, -0.51)	Reference group	Joffe et al. 2014[28,38,43,62] †
On treatment health state utility (95% CI)		0.829 (0.82, 0.83)	0.811	Coon et al. 2018[101]
Discontinuation in first year of model	3.6%	6%	1.3%	See Table D3.9
Annual price for therapy	$6,000*	$104.83‡	N/A	Placeholder price; IBM Micromedex
Model wide inputs				
Duration of VMS and treatment	Median VMS symptoms: 7.4 years Modeled median = 7 years using exponential distribution with rate parameter = 9.99			Avis et al. 2015[13]
Mean VMS-related direct costs per treated person per year	$1,731			Sarrel et al. 2015[102]
Mean VMS-related direct costs per untreated person per year	$2,300			Sarrel et al. 2015[102]
Off treatment with no symptoms health state utility	0.851			Jiang et al. 2021[103]

MENQoL: menopause-specific quality of life, VMS: vasomotor symptoms
*Fezolinetant price is a placeholder; interpret any model findings based on this placeholder price with caution
†Weighted mean difference from placebo was calculated for each intervention
‡Represents sum of the lowest available WAC prices for oral estradiol 1 mg and oral progesterone 100 mg; lowest available WACs for generics are chosen to approximate maximum allowable costs reimbursed by third party payers.

Clinical Inputs

Transition probabilities for moving from on treatment to off treatment were informed by long-term evidence on VMS duration.[13] 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 quantify how long frequent VMS persist after the final menstrual period. The study produced

estimates on the proportion of women with frequent VMS by stage (e.g., perimenopausal, postmenopausal) over multiple years. Base-case duration of VMS was derived from parametric curves fit to observed Kaplan-Meier curves. Transition probabilities were calculated for each time period in the model (annual cycles). We assumed the same duration of VMS across all treatments in the model.

The key treatment effects from fezolinetant clinical studies include the reduction in frequency and severity of VMS as compared to placebo using the 12-week endpoint. In order to generate outcomes related to reductions in frequency from treatment, the model summed average VMS frequency per cycle and over the model time horizon. The difference in VMS episodes on average per cycle (annual) was compared between interventions 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. This metric does not imply patients will avoid entire days without VMS, but is a reflection of the total amount of relief from VMS symptoms a patient may experience in one year compared with not being on treatment.

Without direct elicitation of utility values comparing fezolinetant to no pharmacologic treatment/placebo, we relied on evidence of patient-reported outcome instruments with known utility mappings. 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.[101] The mapping instrument derived a linear relationship of EQ-5D = 0.992 – 0.042*MENQoL. Other patient-reported outcome instruments were used in the fezolinetant trials including the patient-reported outcomes measurement information system (PROMIS) sleep disturbance-short form. However, the total MENQoL score and changes in the total MENQoL scores were chosen given the multiple domains measuring quality of life and changes in quality of life associated with menopause.

Safety endpoints were derived from fezolinetant trials and relevant trials for MHT. We found no evidence on serious adverse events (grade 3 or 4) versus placebo in fezolinetant trials. Since there are known long-term risks such as myocardial infarction from long-term use of MHT, we included the risks and the associated costs and decrements to health-related quality of life. These are described in the supplement section E.

Economic Inputs

As no publicly available list or net price exists for fezolinetant, we used a placeholder price of $6,000 per year for estimates of cost-effectiveness based on analyst market projections and uptake assumptions. This price was used for base-case assessments in the absence of a list price being furnished by the manufacturer; however, this placeholder price was not used to estimate any potential discounts necessary to achieve cost-effectiveness. We referenced generic utilization for MHT. Thus, the lowest available WAC prices with no additional rebates or discounts were used for

the proxy products chosen to represent the respective therapeutic class (annual price of $104.83 for MHT, assuming a prescription for oral estrogen and progesterone).

Other health care utilization unit costs were used in the model for both treated and untreated VMS. Sarrel et al. used the OptumHealth Reporting and Insights Database to estimate direct and indirect costs for women with VMS and stratifying the analysis by treated and untreated among over 250,000 women in the United States. Unit costs for health care utilization were applied to each arm of the model based on assignment of treatment. For example, untreated VMS annual costs were applied to the no pharmacologic treatment/placebo arm of the model whereas the treated VMS annual costs were applied to each active treatment arm's "on treatment" health state. Because the data was derived from administrative claims, this approach assumes no direct link between the magnitude of benefit in terms of reductions in severity or frequency and the change in resource utilization.

Finally, costs for managing and treating future unrelated complications to menopause were modeled for all arms in the model. For the MHT arm of the model, the increased risk of complications associated with MHT in some cases increased (e.g., breast cancer) or decreased (e.g., fractures) the total costs of non-intervention costs. Note that no differences in costs or health-related quality of life associated with these risks were modeled for fezolinetant versus no pharmacologic treatment, as there are currently no data available on the association of fezolinetant with such events and thus we assumed the risk was the same across those arms.

4.3. Results

Base-Case Results

The base-case comparison was fezolinetant versus no pharmacologic treatment/ placebo in patients with menopause-associated VMS. The total discounted costs, life years (LYs), quality-adjusted life years (QALYs), equal value of life years (evLYs) gained and the average VMS episodes per day are detailed in Table 4.4. Using a placeholder price of $6,000 annually, fezolinetant had a total discounted cost of $200,000 with discounted QALYs, LYs, evLYs of 16.43, 19.88, and 16.43, respectively. No pharmacologic treatment had a total discounted cost of $160,000 with discounted QALYs, LYs, evLYs of 16.33, 19.88, and 16.33, respectively.

Table 4.4. Results for the Base-Case for Fezolinetant Compared to No Pharmacologic Treatment

Treatment	Intervention cost	Other non-intervention costs*	Total cost	QALYs	Life Years	evLYs	Average VMS Episodes per Day
Fezolinetant	$50,000†	$150,000	$200,000	16.43	19.88	16.43	7.54
No Pharmacologic Treatment	$0	$160,000	$160,000	16.33	19.88	16.33	10.0
Incremental (Fezolinetant versus No Pharmacologic Treatment)	$50,000†	-$10,000	$40,000	0.10	0.00	0.10	-2.46

evLYs: equal value life year, QALYs: quality-adjusted life year, VMS: vasomotor symptoms

* Other non-intervention costs include long-run unrelated health state costs and do not differ between treatment arms in this base-case analysis

† Based on annual placeholder price of $6,000. Interpret cost findings with caution.

All costs rounded to nearest $10,000.

Table 4.5. presents the discounted lifetime incremental results from the base-case analysis, which include incremental cost-effectiveness ratios for incremental cost per QALY gained, cost per evLY gained, and cost per symptom-free day. Total discounted costs for fezolinetant were approximately $40,000 greater than no pharmacologic treatment; gains in QALYs and evLYs were 0.10 and 0.10, respectively. The cost to avoid one symptom-free day with fezolinetant, equivalent of 10 VMS episodes on average, is approximately $500 for fezolinetant.

Table 4.5. Incremental Cost-Effectiveness Ratios for the Base Case

Treatment	Comparator	Cost per QALY Gained	Cost per evLY Gained	Cost per Symptom-free Day
Fezolinetant*	No pharmacologic treatment	$390,000	$390,000	$500

evLY: equal value life years, MHT: menopausal hormone therapy, QALY: quality-adjusted life year

*Based on annual placeholder price of $6,000. Interpret findings with caution.

Sensitivity Analyses

To demonstrate effects of uncertainty on both costs and health outcomes, we varied input parameters using available measures of parameter uncertainty (e.g., standard errors or plausible ranges). Uncertainty in inputs was derived from literature-based sources inclusive of 95% confidence intervals and/or standard errors. Where uncertainty was not available, we varied inputs by percentages with reasonable lower and upper bounds. Evidence-based distributions were assigned to each input parameter for sensitivity analyses.

One-way sensitivity analysis results for fezolinetant compared to no pharmacologic treatment and MHT compared to no pharmacologic treatment are illustrated in Supplement E3. In terms of incremental QALYs, the model was also sensitive to the proportion of those who discontinue fezolinetant. On the costs side, the model was most sensitive to the cost of treated VMS per year and also to the proportion who discontinue fezolinetant during the first year.

Probabilistic sensitivity analyses were performed by jointly varying all model parameters over 1,000 simulations, then calculating 95% credible range estimates for each model outcome based on the results. Tables 4.6 and 4.7 present the probability of reaching certain cost-effectiveness thresholds for fezolinetant versus no pharmacologic treatment. A total of 14% and 14% of iterations for fezolinetant versus no pharmacologic treatment were beneath a threshold of $150,000 per QALY and $150,000 per evLY, respectively.

Table 4.6. Probabilistic Sensitivity Analysis Cost per QALY Gained Results: Fezolinetant vs. No Pharmacologic Treatment

Treatment	Cost Effective at $50,000 per QALY gained	Cost Effective at $100,000 per QALY gained	Cost Effective at $150,000 per QALY gained	Cost Effective at $200,000 per QALY gained
Fezolinetant*	1%	5%	14%	25%

QALY: quality-adjusted life year

*Price used in this analysis for fezolinetant was a placeholder price

Table 4.7. Probabilistic Sensitivity Analysis Cost per evLY Gained Results: Fezolinetant vs. No Pharmacologic Treatment

Treatment	Cost Effective at $50,000 per evLY gained	Cost Effective at $100,000 per evLY gained	Cost Effective at $150,000 per evLY gained	Cost Effective at $200,000 per evLY gained
Fezolinetant*	1%	5%	14%	25%

evLY: equal value of life year

*Price used in this analysis for fezolinetant was a placeholder price

Scenario Analyses

If data allowed, we considered conducting scenario analyses that included:

1. Comparison between MHT and no pharmacologic treatment.

2. Modified societal perspective that includes components such as productivity losses.

During the scope phase, we considered subgroup analyses stratifying analyses by duration of symptoms but given there is no impact on an incremental level (see Table 4.2 for this assumption) we did not include this scenario in the report. Separately, we considered stratifying risks associated with MHT by age and these inputs are already built into Scenario 1.

Scenario 1 is presented in Table 4.8 and Table 4.9 and provides a comparison between MHT and no pharmacologic treatment. This scenario included changes in risks associated with MHT (e.g., breast cancer among other risks) that are detailed in the supplement.

Table 4.8. Results for MHT Compared to No Pharmacologic Treatment

Treatment	Intervention Cost	Other Non-intervention Costs*	Total cost	QALYs	Life Years	evLYs	Average VMS Episodes per Day
MHT	$900	$158,000	$159,000	16.45	19.88	16.45	6.25
No Pharmacologic Treatment	$0	$157,000	$157,000	16.33	19.88	16.33	10.0
Incremental (MHT vs. No Pharmacologic Treatment)	$900	$700	$1,600	0.125	0.00	0.125	-3.75

evLYs: equal value life year, MHT: menopausal hormone therapy, QALYs: quality-adjusted life year, VMS: vasomotor symptoms

*Other non-intervention costs include long-run unrelated health state costs and differ between treatment arms in this base-case analysis because of increased risks of complications associated with MHT.

Table 4.9. Incremental Cost-Effectiveness Ratios for MHT Compared to No Pharmacologic Treatment

Treatment	Comparator	Cost per QALY Gained	Cost per evLY Gained	Cost per Symptom-free Day
MHT	No pharmacologic treatment	$13,000	$13,000	$12

evLY: equal value life year, MHT: menopausal hormone therapy, QALY: quality-adjusted life-year

Scenario 2 is shown in Table 4.10. Table E2.6 in the Supplement describes the inputs (indirect costs) used for the modified societal perspective.

Table 4.10. Incremental Cost-Effectiveness Ratios for the Modified Societal Perspective for Fezolinetant Compared to No Pharmacologic Treatment

Treatment	Comparator	Cost per QALY Gained	Cost per evLY Gained	Cost per Symptom-free Day
Fezolinetant*	No pharmacologic treatment	$360,000	$360,000	$400

evLY: equal value life year, MHT: menopausal hormone therapy, QALY: quality-adjusted life-year
*Price used in this analysis for fezolinetant was a placeholder price

Threshold Analyses

The annual drug costs at which fezolinetant would reach cost-effectiveness thresholds ranging from $50,000 to $200,000 per QALY gained as well as per evLYG, compared to no pharmacologic treatment, are presented below in Table 4.11.

Table 4.11. Cost per Outcome Threshold Analysis Results for Fezolinetant vs No Pharmacologic Treatment

	Net Price per Unit	Annual Price to Achieve $50,000 per Outcome*	Annual Price to Achieve $100,000 per Outcome*	Annual Price to Achieve $150,000 per Outcome*	Annual Price to Achieve $200,000 per Outcome*
QALY-Based (95% credible range)	To be determined	$1,500	$2,000	$2,500	$3,000
evLY-Based (95% credible range)	To be determined	$1,500	$2,000	$2,500	$3,000

evLY: equal value life year, QALY: quality-adjusted life year
*Rounded to the nearest $500
Note: Price used in this analysis for fezolinetant was a placeholder price

Model Validation

We used several approaches to validate the model. First, we provided preliminary model structure, methods and assumptions to manufacturers, patient groups, and clinical experts. Based on feedback from these groups, we refined data inputs used in the model, as needed. Second, we varied model input parameters to evaluate face validity of changes in results. We performed model

verification for model calculations using internal reviewers. As part of ICER's efforts in acknowledging modeling transparency, we are also sharing the model with the relevant manufacturers for external verification around the time of publishing the draft report for this review. Finally, we compared results to other cost-effectiveness models in this therapy area. The outputs from the model are validated against the trial/study data of the intervention and also any relevant observational datasets.

Uncertainty and Controversies

This cost-effectiveness analysis for fezolinetant was limited by several factors. The price of fezolinetant is currently a placeholder price based on market projections for similar technologies and thus estimates of its cost-effectiveness must be interpreted with caution. Because of inconsistency in trial endpoints, there was no indirect treatment comparison were performed on any outcome, including MENQoL, and therefore no conclusions should be made directly comparing the cost-effectiveness of fezolinetant versus other comparators such as MHT.

Health-related quality of life was derived using a mapping algorithm between the MENQoL and EQ-5D. Without direct utility scores, we relied on this mapping instrument and the total MENQoL scores to produce utility differences across treatment arms. However, the changes in utility scores are a function of the total changes in MENQoL as opposed to the VMS subdomain. This assumption allows for health-related quality of life to be associated not only with VMS but also other symptoms correlated with VMS.[104]

We acknowledge that women with VMS may attempt multiple treatments over the duration of the menopausal transition. The model did not include treatment switching or further attempts at treatment if patients discontinued due to adverse events or lack of efficacy during the first year. This assumption is in line with the objective of the analysis which is to isolate the value of first line usage of therapy. Beyond the first year, discontinuation was associated with resolution of symptoms. The assumptions on discontinuation beyond the first cycle (year) were the same across all treatments and comparators and does not impact the incremental findings as the same assumption was made across all arms of the model.

Finally, there were observed no treatment effect differences leading to cost offsets both for direct and indirect costs. We did, however, incorporate an associate between pharmacologic treatment and reduced direct health care costs.[102] There was a lack of literature that directly linked reductions in VMS frequency and severity with potential cost offsets and therefore we applied the same cost offsets for all treated patients, regardless of the treatment selected.

4.4 Summary and Comment

Assuming a placeholder price for fezolinetant, the base-case findings suggest that fezolinetant provides clinical benefit in terms of gains in QALYs and evLYs over no pharmacologic treatment but does so with increased costs to the health system. Of both pharmacologic treatments assessed, MHT had the greatest gains in QALYs and evLYs with the least amount of cost increase to the health system. Given the focus on VMS improvement in the economic model, the key drivers were health-related quality of life on fezolinetant and cost savings from treatment on fezolinetant. In line with its modest observed clinical and cost offset benefits, the threshold prices for fezolinetant ranged from $1,500 to $3,000 per QALY or evLY at a variety of commonly accepted cost-effectiveness thresholds.

5. Contextual Considerations and Potential Other Benefits

Our reviews seek to provide information on potential other benefits offered by the intervention to the individual patient, caregivers, the delivery system, other patients, or the public that was not available in the evidence base nor could be adequately estimated within the cost-effectiveness model. These elements are listed in the table below, with related information gathered from patients and other stakeholders. Following the public deliberation on this report the appraisal committee will vote on the degree to which each of these factors should affect overall judgments of long-term value for money of the intervention in this review.

Table 5.1. Contextual Considerations

Contextual Consideration	Relevant Information
Acuity of need for treatment of individual patients based on short-term risk of death or progression to permanent disability	Patients, caregivers, and clinical experts all identified a need for new therapeutic options for patients with VMS, especially for those who have contraindications for MHT.
Magnitude of the lifetime impact on individual patients of the condition being treated	VMS is a condition lasts a median of 9.4 years and can continue for more than a decade in many women. It can affect sleep, workplace performance and intimate relationships.
There is uncertainty about long-term efficacy	Whereas the duration of VMS is typically many years and there is evidence of the efficacy of MHT for multiple years, the primary outcomes in key fezolinetant trials only assessed efficacy up to 12 weeks.

Table 5.2. Potential Other Benefits or Disadvantages

Potential Other Benefit or Disadvantage	Relevant Information
Patients' ability to achieve major life goals related to education, work, or family life	Unpredictable flushing and sweating along with insomnia can adversely affect work performance.
Caregivers' quality of life and/or ability to achieve major life goals related to education, work, or family life	VMS is mainly managed by the patient and is not expected to impose substantial caregiver burdens in the traditional sense, but household members or intimate partners may be impacted by certain aspects of VMS such as sleep disruption, mood swings, or concerns related to sexual activity.
Patients' ability to manage and sustain treatment given the complexity of regimen	All treatments are administered orally or transdermally and so there is not expected to be a difference in complexity of regimen between treatments.
Society's goal of reducing health inequities	VMS associated with menopause disproportionally impacts certain racial and ethnic groups, in terms of both symptom frequency and severity and symptom duration, and in terms of underlying comorbidities that may impact treatment choices. Additionally, there are differences among individuals in their ability to access health care as well as surrounding social norms or stigma. This may exacerbate existing health inequities by selectively limiting therapy, including medications, to those patients who have fewer comorbidities, are able to afford them and have access to health care providers who can prescribe them. In highlighting inequalities in the VMS associated with menopause space, ICER calculated the Health Improvement Distribution Index, looking at the relative proportion of any health gains from treatment of VMS associated with menopause for the following groups who have a higher prevalence than the general US population.[105] Importantly, the key racial and ethnic data used for this analysis were longitudinal in nature and thus captured the composite prevalence associated with VMS in menopause. For more information on how we calculate the Health Improvement Distribution Index, refer to the Supplement. • African American/Black women: 1.3

6. Health Benefit Price Benchmarks

ICER does not provide health benefit price benchmarks as part of draft reports because results may change with revision following receipt of public comments. We therefore caution readers against assuming that the values provided in the Threshold Prices section of this draft report will match the health benefit price benchmarks that will be presented in the next version of this Report.

7. Potential Budget Impact

7.1. Overview of Key Assumptions

ICER uses results from the cost-effectiveness model to estimate the potential total budgetary impact of fezolinetant for women with moderate to severe VMS associated with menopause. We additionally use its placeholder price ($6,000 annually) and the three threshold prices (at $50,000, $100,000, and $150,000 per QALY) ($1,500, $2,000, and $2,500 per year, respectively) in our estimates of fezolinetant's budget impact.

Potential budget impact is defined as the total differential cost of the new therapy (fezolinetant) than relevant existing therapy (no pharmacologic therapy) for the treated population, calculated as differential health care costs (including drug costs) minus any offsets in these costs from averted health care events. All costs are undiscounted and estimated over a five-year time horizon. For 2022-2023, the five-year annualized potential budget impact threshold that should trigger policy actions to manage access and affordability is calculated to be approximately $777 million per year for new drugs.

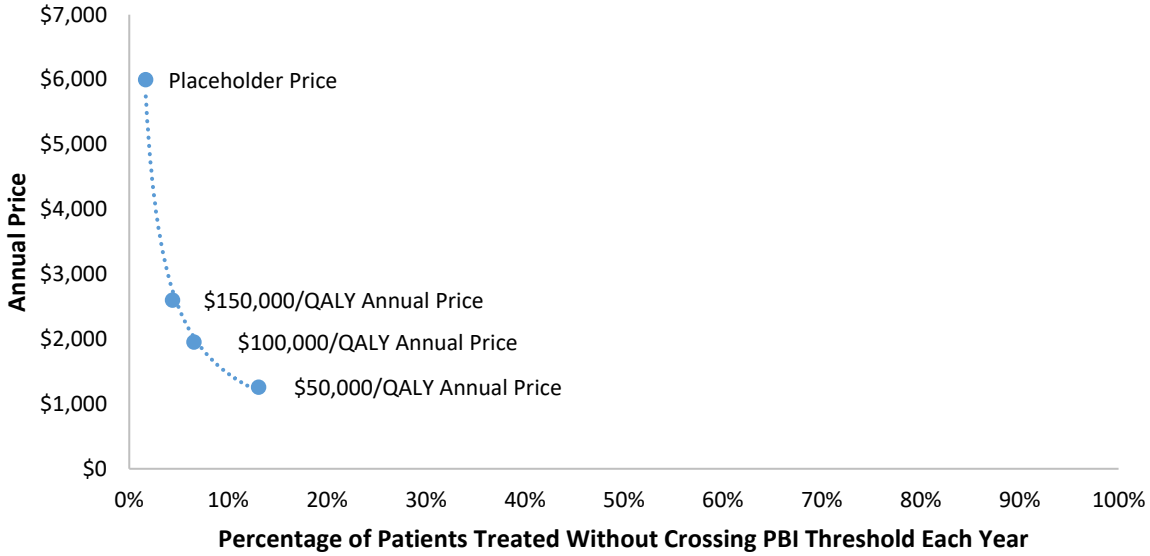
This potential budget impact analysis includes the estimated number of individuals in the US who would be eligible for treatment with fezolinetant. From relevant sources (see supplemental section F), we derive an estimate of 16,700,000 women eligible for treatment with fezolinetant in the US. For the purposes of this analysis, we assume that 20% of these patients would initiate treatment in each of the five years, or 3,340,000 patients per year.

As fezolinetant has been evaluated in a population of VMS associated with menopause patients who cannot or will not take MHT, we have chosen to model all patients belonging to a no pharmacologic treatment arm at baseline. Additionally, we did not conduct a budget impact analysis of MHT as it has been available for patients for several years.

7.2. Results

The primary findings of fezolinetant's budget impact in the US population of women with VMS associated with menopause is depicted in Figure 7.1. below. At fezolinetant's placeholder price of $6,000 annually, approximately 1.7% of women could be treated without surpassing the potential budget impact threshold of $777 million. In contrast, 4.4%, 6.5%, and 13.1% of the total population could be treated at the annual threshold prices of $150,000/QALY ($2,500), $100,000/QALY ($2,000), and $50,000/QALY ($1,500), respectively. Refer to the supplement section F for additional findings described at the per-individual level.

Figure 7.1. Budgetary Impact of Fezolinetant in Women with VMS Associated with Menopause



PBI: potential budget impact, QALY: quality-adjusted life-year

References

1. 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(9):924-932.
2. 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(2):153-159.
3. Miller SR, Gallicchio LM, Lewis LM, et al. Association between race and hot flashes in midlife women. *Maturitas.* 2006;54(3):260-269.
4. Sarrel P, Portman D, Lefebvre P, et al. Incremental direct and indirect costs of untreated vasomotor symptoms. *Menopause.* 2015;22(3):260-266.
5. "The 2022 Hormone Therapy Position Statement of The North American Menopause Society" Advisory Panel. The 2022 hormone therapy position statement of The North American Menopause Society. *Menopause.* 2022;29(7):767-794.
6. Astellas Submits Fezolinetant New Drug Application to U.S. FDA [press release]. June 23rd, 2022 2022.
7. Rance NE, Dacks PA, Mittelman-Smith MA, Romanovsky AA, Krajewski-Hall SJ. 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(3):211-227.
8. Jayasena CN, Comninos AN, Stefanopoulou E, et al. Neurokinin B Administration Induces Hot Flushes in Women. *Sci Rep.* 2015;5(1):8466.
9. 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(2):148-157.
10. 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(7):1226-1235.
11. Gold EB, Sternfeld B, Kelsey JL, et al. Relation of Demographic and Lifestyle Factors to Symptoms in a Multi-Racial/Ethnic Population of Women 40–55 Years of Age. *American Journal of Epidemiology.* 2000;152(5):463-473.
12. Politi MC, Schleinitz MD, Col NF. Revisiting the Duration of Vasomotor Symptoms of Menopause: A Meta-Analysis. *Journal of General Internal Medicine.* 2008;23(9):1507-1513.
13. Avis NE, Crawford SL, Greendale G, et al. Duration of Menopausal Vasomotor Symptoms Over the Menopause Transition. *JAMA Internal Medicine.* 2015;175(4):531-539.
14. Tepper PG, Brooks MM, Randolph JF, Jr., et al. Characterizing the trajectories of vasomotor symptoms across the menopausal transition. *Menopause.* 2016;23(10):1067-1074.
15. Reed SD, Lampe JW, Qu C, et al. Premenopausal vasomotor symptoms in an ethnically diverse population. *Menopause.* 2014;21(2):153-158.
16. Stuenkel CA, Davis SR, Gompel A, et al. Treatment of Symptoms of the Menopause: An Endocrine Society Clinical Practice Guideline. *J Clin Endocrinol Metab.* 2015;100(11):3975-4011.
17. Cramer H, Peng W, Launche R. Yoga for menopausal symptoms—A systematic review and meta-analysis. *Maturitas.* 2018;109:13-25.
18. Pockaj BA, Gallagher JG, Loprinzi CL, et al. Phase III Double-Blind, Randomized, Placebo-Controlled Crossover Trial of Black Cohosh in the Management of Hot Flashes: NCCTG Trial N01CC. *Journal of Clinical Oncology.* 2006;24(18):2836-2841.

19. Administration FaD. BRISDELLE® (paroxetine) Prescribing Information. Food and Drug Administration (FDA). https://www.accessdata.fda.gov/drugsatfda_docs/label/2014/204516s003lbl.pdf. Published 1992. Accessed2022.
20. A.P. H, Williams M. The efficacy and tolerability of SSRI/SNRIs in the treatment of vasomotor symptoms in menopausal women: a systematic review. *Journal of the American Association of Nurse Practitioners.* 2015;27(1):54-61.
21. Shams T, Firwana B, Habib F, et al. SSRIs for Hot Flashes: A Systematic Review and Meta-Analysis of Randomized Trials. *Journal of general internal medicine.* 2013;29(1):204-213.
22. Investigators WGftWsHI. Risks and Benefits of Estrogen Plus Progestin in Healthy Postmenopausal WomenPrincipal Results From the Women's Health Initiative Randomized Controlled Trial. *JAMA.* 2002;288(3):321-333.
23. Brunner RL, Aragaki A, Barnabei V, et al. Menopausal symptom experience before and after stopping estrogen therapy in the Women's Health Initiative randomized, placebo-controlled trial. *Menopause.* 2010;17(5):946-954.
24. Administration FaD. *Guidance for Industry :Estrogen and Estrogen/Progestin Drug Products to Treat Vasomotor Symptoms and Vulvar and Vaginal Atrophy Symptoms — Recommendations for Clinical Evaluation.* 2003.
25. Johnson KA, Lademacher C, Nappi RE, et al. A Phase 3, randomized, placebo-controlled, 12-week, double-blind study, plus a non-controlled extension treatment period, to assess efficacy and safety of fezolinetant, a neurokinin-3 receptor antagonist, in women with moderate-to-severe vasomotor symptoms associated with menopause. Paper presented at: North American Menopause Society (NAMS) Annual Meeting2021.
26. Johnson KA, Nappi RE, Neal-Perry G, et al. Fezolinetant for Treatment of Moderate-to-severe Vasomotor Symptoms Associated with Menopause: Results from a 52-week Study (SKYLIGHT 2). ENDO; 2022; Atlanta, Georgia.
27. Lederman S, Shapiro M, Stute P, Lee M, Wang X, G. N-P. Phase 3 Study of Fezolinetant for Treatment of Moderate-to-Severe Vasomotor Symptoms Associated With Menopause [A132]. Paper presented at: Obstetrics & Gynecology2022.
28. Astellas Announces Topline 12-week Results from Phase 3 Study of Fezolinetant for the Nonhormonal Treatment of Vasomotor Symptoms in Women in Asia [press release]. Mar 15, 2022 2022.
29. Astellas Pharma Inc. Astellas Announces Topline Results from Phase 3 Long-Term Safety Study of Fezolinetant in Mainland China. https://www.astellas.com/en/news/26281. Published 2022. Accessed September 14, 2022.
30. Clinicaltrials.gov. A Study to Find Out if Fezolinetant Helps Reduce Moderate to Severe Hot Flashes in Women in Asia Going Through Menopause (Moonlight 1). https://clinicaltrials.gov/ct2/show/NCT04234204. Published 2020. Accessed September 9, 2022, 2022.
31. Clinicaltrials.gov. A Study to Assess the Safety and Tolerability of Fezolinetant in Women Seeking Treatment for Relief of Vasomotor Symptoms (VMS) Associated With Menopause (Moonlight 3). https://clinicaltrials.gov/ct2/show/NCT04451226. Published 2020. Accessed September 16, 2022.
32. Astellas Pharma Inc. U.S. FDA Accepts Astellas' New Drug Application for Fezolinetant. 2022. https://www.astellas.com/en/news/26126. Accessed September 9, 2022.

33. Lobo RA, Archer DF, Kagan R, et al. A 17beta-Estradiol-Progesterone Oral Capsule for Vasomotor Symptoms in Postmenopausal Women: A Randomized Controlled Trial. *Obstetrics and gynecology.* 2018;132(1).
34. Endrikat J, Graeser T, Mellinger U, Ertan K, Holz C. A multicenter, prospective, randomized, double-blind, placebo-controlled study to investigate the efficacy of a continuous-combined hormone therapy preparation containing 1mg estradiol valerate/2mg dienogest on hot flushes in postmenopausal women. *Maturitas.* 2007;58(2).
35. Stevenson JC, Durand G, Kahler E, Pertyński T. Oral ultra-low dose continuous combined hormone replacement therapy with 0.5 mg 17β-oestradiol and 2.5 mg dydrogesterone for the treatment of vasomotor symptoms: results from a double-blind, controlled study. *Maturitas.* 2010;67(3).
36. Schürmann R, Holler T, Benda N. Estradiol and drospirenone for climacteric symptoms in postmenopausal women: a double-blind, randomized, placebo-controlled study of the safety and efficacy of three dose regimens. *Climacteric.* 2004;7(2).
37. Lin SQ, Sun LZ, Lin JF, et al. Estradiol 1 mg and drospirenone 2 mg as hormone replacement therapy in postmenopausal Chinese women. *Climacteric.* 2011;14(4).
38. Joffe H, Guthrie KA, LaCroix AZ, et al. Low-dose estradiol and the serotonin-norepinephrine reuptake inhibitor venlafaxine for vasomotor symptoms: A randomized clinical trial. *JAMA Intern Med.* 2014;174(7).
39. Archer DF, Schmelter T, Schaefers M, Gerlinger C, Gude K. A randomized, double-blind, placebo-controlled study of the lowest effective dose of drospirenone with 17β-estradiol for moderate to severe vasomotor symptoms in postmenopausal women. *Menopause.* 2013;21(3).
40. Panay N, Ylikorkala O, Archer DF, Gut R, Lang E. Ultra-low-dose estradiol and norethisterone acetate: effective menopausal symptom relief. *Climacteric : the journal of the International Menopause Society.* 2007;10(2):120-131.
41. Speroff L, Whitcomb RW, Kempfert NJ, Boyd RA, Paulissen JB, Rowan JP. Efficacy and local tolerance of a low-dose, 7-day matrix estradiol transdermal system in the treatment of menopausal vasomotor symptoms. *Obstet Gynecol.* 1996;88(4 Pt 1).
42. Simon JA, Portman DJ, Kaunitz AM, et al. Low-dose paroxetine 7.5 mg for menopausal vasomotor symptoms: Two randomized controlled trials. *Menopause.* 2013;20(10).
43. Freeman EW, Guthrie KA, Caan B, et al. Efficacy of escitalopram for hot flashes in healthy menopausal women: a randomized controlled trial. *JAMA.* 2011;305(3).
44. Evans ML, Pritts E, Vittinghoff E, McClish K, Morgan KS, Jaffe RB. Management of postmenopausal hot flushes with venlafaxine hydrochloride: a randomized, controlled trial. *None.* 2005;105(1).
45. 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(2).
46. Speroff L, Gass M, Constantine G, S. O. Efficacy and tolerability of desvenlafaxine succinate treatment for menopausal vasomotor symptoms: a randomized controlled trial. *Obstetrics & Gynecology.* 2008;111(1):77-87.
47. Bouchard P, Panay N, De Villiers TJ, et al. Randomized placebo-and active-controlled study of desvenlafaxine for menopausal vasomotor symptoms. *Climacteric.* 2012;15(1):12-20.
48. Pinkerton JV, Archer DF, Guico-Pabia CJ, Hwang E, Cheng R-fJ. Maintenance of the efficacy of desvenlafaxine in menopausal vasomotor symptoms: a 1-year randomized controlled trial. *Menopause (New York, NY).* 2013;20(1):38-46.

49. 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(3).
50. Wyrwich KW, Spratt DI, Gass M, Yu H, JD. B. Identifying meaningful differences in vasomotor symptoms among menopausal women. *Menopause.* 2008;15(4):698-705.
51. Pinkerton JV, Kagan R, Portman D, Sathyanarayana R, Sweeney M. Phase 3 randomized controlled study of gastroretentive gabapentin for the treatment of moderate-to-severe hot flashes in menopause. *Menopause.* 2014;21(6).
52. Reddy SY, Warner H, Guttuso Jr T, et al. Gabapentin, estrogen, and placebo for treating hot flushes: A randomized controlled trial. *Obstet Gynecol.* 2006;108(1).
53. Guttuso Jr T, Kurlan R, McDermott MP, Kieburtz K. Gabapentin's effects on hot flashes in postmenopausal women: A randomized controlled trial. *Obstet Gynecol.* 2003;101(2).
54. Lederman S, Shapiro Cm M, 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 [A132]. *Obstetrics & Gynecology.* 2022;139(1).
55. Nappi R. Fezolinetant for the treatment of moderate-to-severe vasomotor symptoms (VMS) associated with menopause: pooled subgroup analyses by baseline age and baseline VMS severity from the SKYLIGHT 1 and SKYLIGHT 2 studies. *None.* 2022;None(None).
56. Johnson K, Lademacher C, Nappi R. A Phase 3, randomized, placebo-controlled, 12-week, double-blind study, plus a non-controlled extension treatment period, to assess efficacy and safety of fezolinetant, a neurokinin-3 receptor antagonist, in women with moderate-to-severe vasomotor symptoms associated with menopause. *Menopause.* 2021;None(None).
57. Johnson K, Nappi R, Neal-Perry G. Fezolinetant for Treatment of Moderate-to-severe Vasomotor Symptoms Associated with Menopause: Results from a 52-week Study (SKYLIGHT 2). *None.* 2022;None(None).
58. Black DR, Minkin MJ, Graham S, Bernick B, Mirkin S. Effects of combined 17&bgr;-estradiol and progesterone on weight and blood pressure in postmenopausal women of the REPLENISH trial. *None.* 2020;28(1).
59. Kagan R, Constantine G, Kaunitz AM, Bernick B, Mirkin S. Improvement in sleep outcomes with a 17β-estradiol-progesterone oral capsule (TX-001HR) for postmenopausal women. *Menopause.* 2019;26(6).
60. Kaunitz AM, Bitner D, Constantine GD, Bernick B, Graham S, Mirkin S. 17beta-estradiol/progesterone in a single, oral, softgel capsule (TX-001HR) significantly increased the number of vasomotor symptom-free days in the REPLENISH trial. *None.* 2020;27(12).
61. Mirkin S, Goldstein SR, Archer DF, Pickar JH, Graham S, Bernick B. Endometrial safety and bleeding profile of a 17β-estradiol/progesterone oral softgel capsule (TX-001HR). *Menopause.* 2020;27(4).
62. Simon JA, Kaunitz AM, Kroll R, Graham S, Bernick B, S. M. Oral 17β-estradiol/progesterone (TX-001HR) and quality of life in postmenopausal women with vasomotor symptoms. *Menopause.* 2019;26(5).
63. Tsiligiannis S, Wick-Urban BC, van der Stam J, Stevenson JC. Efficacy and safety of a low-dose continuous combined hormone replacement therapy with 0.5 mg 17β-estradiol and 2.5 mg dydrogesterone in subgroups of postmenopausal women with vasomotor symptoms. *Maturitas.* 2020;139(None).
64. Archer DF, Seidman L, Constantine GD, Pickar JH, S. O. A double-blind, randomly assigned, placebo-controlled study of desvenlafaxine efficacy and safety for the treatment of vasomotor

symptoms associated with menopause. *American journal of obstetrics and gynecology.* 2009;200(2):172-e171.
65. Archer DF, Dupont CM, Constantine GD, Pickar JH, Olivier S, Investigators S. Desvenlafaxine for the treatment of vasomotor symptoms associated with menopause: a double-blind, randomized, placebo-controlled trial of efficacy and safety. *American journal of obstetrics and gynecology.* 2009;200(3):238-e231.
66. Archer DF, Pinkerton JV, Guico-Pabia CJ, Hwang E, Cheng RFJ. Cardiovascular, cerebrovascular, and hepatic safety of desvenlafaxine for 1 year in women with vasomotor symptoms associated with menopause. *Menopause.* 2013;20(1).
67. Diem SJ, Joffe H, Larson JC, et al. Effects of escitalopram on markers of bone turnover: A randomized clinical trial. *J Clin Endocrinol Metab.* 2014;99(9).
68. Simon JA, Portman DJ, Kaunitz AM, et al. Low-dose paroxetine 7.5 mg for menopausal vasomotor symptoms: two randomized controlled trials. *Menopause.* 2013;20(10):1027-1035.
69. Pinkerton JV, Joffe H, Kazempour K, Mekonnen H, Bhaskar S, Lippman J. Low-dose paroxetine (7.5 mg) improves sleep in women with vasomotor symptoms associated with menopause. *Menopause.* 2015;22(1).
70. Portman DJ, Kaunitz AM, Kazempour K, Mekonnen H, Bhaskar S, Lippman J. Effects of low-dose paroxetine 7.5 mg on weight and sexual function during treatment of vasomotor symptoms associated with menopause. *Menopause.* 2014;None(None).
71. Yurcheshen ME, Guttuso T, McDermott M, Holloway RG, Perlis M. Effects of gabapentin on sleep in menopausal women with hot flashes as measured by a Pittsburgh sleep quality index factor scoring model. *J Women's Health.* 2009;18(9).
72. Pinkerton JV, Kagan R, Portman D, Sathyanarayana R, Sweeney M, Investigators. B. Phase 3 randomized controlled study of gastroretentive gabapentin for the treatment of moderate-to-severe hot flashes in menopause. *Menopause.* 2014;21(6):567-573.
73. Constantine GD, Revicki DA, Kagan R, et al. Evaluation of clinical meaningfulness of estrogen plus progesterone oral capsule (TX-001HR) on moderate to severe vasomotor symptoms. *Menopause.* 2019;26(5).
74. Constantine GD, Simon JA, Kaunitz AM, et al. TX-001HR is associated with a clinically meaningful effect on severity of moderate to severe vasomotor symptoms in the REPLENISH trial. *Menopause.* 2020;27(11).
75. Lewis JE, Hilditch JR, Wong CJ. Further psychometric property development of the Menopause-Specific Quality of Life questionnaire and development of a modified version, MENQOL-Intervention questionnaire. *Maturitas.* 2005;50(3):209-221.
76. Diem SJ, LaCroix AZ, Reed SD, et al. Effects of Pharmacologic and Non-pharmacologic Interventions on Menopause-related Quality of Life: A Pooled Analysis of Individual Participant Data from Four MsFLASH Trials. *Menopause.* 2020;27(10).
77. Caan B, LaCroix AZ, Joffe H, et al. Effects of estrogen or venlafaxine on menopause related quality of life in healthy postmenopausal women with hot flashes: a placebo-controlled randomized trial. *Menopause.* 2015;22(6).
78. Ensrud KE, Guthrie KA, Hohensee C, et al. Effects of estradiol and venlafaxine on insomnia symptoms and sleep quality in women with hot flashes. *Sleep.* 2015;38(1).
79. Guthrie KA, Larson JC, Ensrud KE, et al. Effects of pharmacologic and nonpharmacologic interventions on insomnia symptoms and self-reported sleep quality in women with hot flashes: A pooled analysis of individual participant data from four MsFLASH trials. *Sleep.* 2018;41(1).

80. Prentice RL, Aragaki AK, Chlebowski RT, et al. Randomized Trial Evaluation of the Benefits and Risks of Menopausal Hormone Therapy among Women 50-59 Years of Age. *Am J Epidemiol.* 2021;190(3).
81. Marjoribanks J, Farquhar C, Roberts H, Lethaby A, Lee J. Long-term hormone therapy for perimenopausal and postmenopausal women. *Cochrane Database Syst Rev.* 2017;2017(1).
82. Kim JE, Chang JH, Jeong MJ, et al. A systematic review and meta-analysis of effects of menopausal hormone therapy on cardiovascular diseases. *Scientific reports.* 2020;10(1):1-3.
83. Rossouw JE, Prentice RL, Manson JE, et al. Postmenopausal hormone therapy and risk of cardiovascular disease by age and years since menopause. *Jama.* 2007;297(13):1465-1477.
84. Boekhout AH, Vincent AD, Dalesio OB, et al. Management of hot flashes in patients who have breast cancer with venlafaxine and clonidine: A randomized, double-blind, placebo-controlled trial. *J Clin Oncol.* 2011;29(29).
85. Capriglione S, Plotti F, Montera R, et al. Role of paroxetine in the management of hot flashes in gynecological cancer survivors: Results of the first randomized single-center controlled trial. *Gynecol Oncol.* 2016;143(3).
86. Grady D, Cohen B, Tice J, Kristof M, Olyaie A, Sawaya GF. Ineffectiveness of sertraline for treatment of menopausal hot flushes: a randomized controlled trial. *Obstet Gynecol.* 2007;109(4).
87. Boekhout AH, Vincent AD, Dalesio OB, et al. Management of hot flashes in patients who have breast cancer with venlafaxine and clonidine: a randomized, double-blind, placebo-controlled trial. *J Clin Oncol.* 2011;29(29):3862-3868.
88. United States Census Bureau. 2020 Census Illuminates Racial and Ethnic Composition of the Country. 2021.
89. Stearns V, Beebe KL, Iyengar M, Dube E. Paroxetine Controlled Release in the Treatment of Menopausal Hot Flashes. *JAMA.* 2003;289(21).
90. Loprinzi CL, Qin R, Baclueva EP, et al. Phase III, Randomized, Double-Blind, Placebo-ControlledEvaluation of Pregabalin for Alleviating Hot Flashes, N07C1. *Journal of clinical oncology.* 2010;28(4).
91. Sloan JA, Loprinzi CL, Novotny PJ, Barton DL, Lavasseur BI, Windschitl H. Methodologic lessons learned from hot flash studies. *J Clin Oncol.* 2001;19(23):4280-4290.
92. Rossouw JE, Anderson GL, Prentice RL, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. *JAMA.* 2002;288(3):321-333.
93. Anderson GL, Limacher M, Assaf AR, et al. Effects of conjugated equine estrogen in postmenopausal women with hysterectomy: the Women's Health Initiative randomized controlled trial. *JAMA.* 2004;291(14):1701-1712.
94. Manson JE, Chlebowski RT, Stefanick ML, et al. Menopausal hormone therapy and health outcomes during the intervention and extended poststopping phases of the Women's Health Initiative randomized trials. *JAMA.* 2013;310(13):1353-1368.
95. Sprague BL, Trentham-Dietz A, Cronin KA. A sustained decline in postmenopausal hormone use: results from the National Health and Nutrition Examination Survey, 1999-2010. *Obstet Gynecol.* 2012;120(3):595-603.
96. Croke LM. ACOG Releases Clinical Guidelines on Management of Menopausal Symptoms. *Am Fam Physician.* 2014;90(5):338-340.
97. The American College of Obsetricians and Gynecologists. Management of Menopausal Symptoms. *Practice Bulletin.* 2014;Number 141. https://www.acog.org/clinical/clinical-

guidance/practice-bulletin/articles/2014/01/management-of-menopausal-symptoms. Accessed August 30, 2022.
98. National Institute for Health and Care Excellence. Menopause: diagnosis and management. 2015. https://www.nice.org.uk/guidance/ng23. Accessed August 30, 2022.
99. Santoro N, Waldbaum A, Lederman S, et al. Effect of the neurokinin 3 receptor antagonist fezolinetant on patient-reported outcomes in postmenopausal women with vasomotor symptoms: results of a randomized, placebo-controlled, double-blind, dose-ranging study (VESTA). *Menopause.* 2020;27(12).
100. Downing NS, Shah ND, Aminawung JA, et al. Postmarket Safety Events Among Novel Therapeutics Approved by the US Food and Drug Administration Between 2001 and 2010. *Jama.* 2017;317(18):1854-1863.
101. 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(6):566-573.
102. Sarrel P, Portman D, Lefebvre P, et al. Incremental direct and indirect costs of untreated vasomotor symptoms. *Menopause.* 2015;22(3):260-266.
103. Jiang R, Janssen MFB, Pickard AS. US population norms for the EQ-5D-5L and comparison of norms from face-to-face and online samples. *Qual Life Res.* 2021;30(3):803-816.
104. Mirkin S, Graham S, Revicki DA, Bender RH, Bernick B, GD. C. Relationship between vasomotor symptom improvements and quality of life and sleep outcomes in menopausal women treated with oral, combined 17b-estradiol/progesterone. *Menopause.* 2019;26(6).
105. 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. *Am J Public Health.* 2006;96(7):1226-1235.
106. Harlow SD, Gass M, Hall JE, et al. Executive summary of the Stages of Reproductive Aging Workshop + 10: addressing the unfinished agenda of staging reproductive aging. *J Clin Endocrinol Metab.* 2012;97(4):1159-1168.
107. Modi M, WS. D. Neurokinin 3 receptor antagonism: a novel treatment for menopausal hot flushes. *Neuroendocrinology.* 2019;109(3):242-248.
108. 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(12).
109. Health Measures. PROMIS Sleep Disturbance - Calibrated Items. https://www.healthmeasures.net/index.php?option=com_instruments&view=measure&id=70&Itemid=992. Published 2016. Accessed September 21, 2022.
110. Hurst H, J. B. Assessing the clinical significance of change scores recorded on subjective outcome measures. *J Manipulative Physiol Ther.* 2004;27:26-35.
111. Hilditch JR, Lewis J, Peter A, et al. A menopause-specific quality of life questionnaire: development and psychometric properties. *Maturitas.* 1996;24(6):161-175.
112. Radtke JV, Terhorst L, SM. C. The Menopause-Specific Quality of Life (MENQOL) questionnaire: psychometric evaluation among breast cancer survivors. *Menopause.* 2011;18(3).
113. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. *Lancet Oncol.* 2012;13(11):1141-1151.
114. US Census Bureau. 2021.

115. 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(4).
116. Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. *Ann Intern Med.* 1997;126(5):376-380.
117. Higgins JP. Cochrane Collaboration Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011]. 2008.
118. Sarri G, Pedder H, Dias S, Guo Y, Lumsden MA. Vasomotor symptoms resulting from natural menopause: a systematic review and network meta-analysis of treatment effects from the National Institute for Health and Care Excellence guideline on menopause. *BJOG: An International Journal of Obstetrics & Gynaecology.* 2017;124(10):1514-1523.
119. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ.* 2021;372:n71.
120. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ.* 2019;366.
121. Higgins J, Savović J, Page M, Sterne J. *Revised Cochrane risk-of-bias tool for randomized trials (RoB 2).* 2019.
122. Ollendorf DA PS. An integrated evidence rating to frame comparative effectiveness assessments for decision makers. *Medical Care.* 2010.
123. Ollendorf D P, SD. ICER Evidence Rating Matrix: A User's Guide. 2020.
124. Lobo RA, Kaunitz AM, Santoro N, Bernick B, Graham S, Mirkin S. Metabolic and cardiovascular effects of TX-001HR in menopausal women with vasomotor symptoms. *Climacteric.* 2019;22(6).
125. Kalay AE, Demir B, Haberal A, Kalay M, O. K. Efficacy of citalopram on climacteric symptoms. *Menopause.* 2007;14(2):223-229.
126. Butt DA, Lock M, Lewis JE, Ross S, Moineddin R. Gabapentin for the treatment of menopausal hot flashes: A randomized controlled trial. *Menopause.* 2008;15(2).
127. McClung MR, Kagan R, Graham S, Bernick B, Mirkin S, Constantine G. Effects of E2/P4 oral capsules on bone turnover in women with vasomotor symptoms. *Menopause (New York, NY).* 2022;29(3).
128. Berhan Y, A. B. Is desvenlafaxine effective and safe in the treatment of menopausal vasomotor symptoms? A meta-analysis and meta-regression of randomized double-blind controlled studies. *Ethiopian Journal of Health Sciences.* 2014;24(3):209-218.
129. GD S, PJ N, A B, al. e. Recommendations for Conduct, Methodological Practices, and Reporting of Cost-effectiveness Analyses: Second Panel on Cost-Effectiveness in Health and Medicine. *Jama.* 2016;316(10):1093-1103.
130. Prentice RL, Aragaki AK, Chlebowski RT, et al. Randomized trial evaluation of the benefits and risks of menopausal hormone therapy among women 50–59 years of age. *American journal of epidemiology.* 2021;190(3):365-375.
131. D'Agostino RB, Sr., Vasan RS, Pencina MJ, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. *Circulation.* 2008;117(6):743-753.
132. Heit JA. Epidemiology of venous thromboembolism. *Nat Rev Cardiol.* 2015;12(8):464-474.
133. Hendrick RE, Monticciolo DL, Biggs KW, Malak SF. Age distributions of breast cancer diagnosis and mortality by race and ethnicity in US women. *Cancer.* 2021;127(23):4384-4392.
134. Blume SW, Curtis JR. Medical costs of osteoporosis in the elderly Medicare population. *Osteoporos Int.* 2011;22(6):1835-1844.
135. Sullivan PW, Ghushchyan V. Preference-Based EQ-5D index scores for chronic conditions in the United States. *Med Decis Making.* 2006;26(4):410-420.

136. Kazi DS, Moran AE, Coxson PG, et al. Cost-effectiveness of PCSK9 inhibitor therapy in patients with heterozygous familial hypercholesterolemia or atherosclerotic cardiovascular disease. *Jama.* 2016;316(7):743-753.
137. O'Sullivan AK, Rubin J, Nyambose J, Kuznik A, Cohen DJ, Thompson D. Cost estimation of cardiovascular disease events in the US. *PharmacoEconomics.* 2011;29(8):693-704.
138. Grosse SD, Nelson RE, Nyarko KA, Richardson LC, Raskob GE. The economic burden of incident venous thromboembolism in the United States: A review of estimated attributable healthcare costs. *Thromb Res.* 2016;137:3-10.
139. McGarvey N, Gitlin M, Fadli E, Chung KC. Increased healthcare costs by later stage cancer diagnosis. *BMC Health Serv Res.* 2022;22(1):1155.
140. Velentzis LS, Salagame U, Canfell K. Menopausal hormone therapy: a systematic review of cost-effectiveness evaluations. *BMC Health Serv Res.* 2017;17(1):326.
141. US Census Bureau. 2017 National Population Projections Tables: Main Series. 2017.
142. US Census Bureau. National Population by Characteristics: 2010-2020. 2019.
143. Astellas Pharma Inc. Data on File.
144. Institute for Clinical and Economic Review. 2020-2023 Value Assessment Framework. https://icer.org/our-approach/methods-process/value-assessment-framework/. Published 2020. Accessed.
145. Pearson SD. The ICER Value Framework: Integrating Cost Effectiveness and Affordability in the Assessment of Health Care Value. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research.* 2018:258-265.

Supplemental Materials

A. Background: Supplemental Information

A1. Definitions

Menopause Definitions

Menopause: Natural menopause is defined as the permanent cessation of menstruation. Menopause is defined retrospectively, after women have experienced 12 months of cessation of menstruation without any other obvious pathologic or physiologic cause. Surgical menopause occurs after bilateral oophorectomy (removal of the ovaries) which removes the main source of estrogen in the body thus triggering the onset of menopause after surgery.

Perimenopause: Time that encompasses the menopausal transition plus one year after final menstrual period.[106]

Postmenopause: Begins at the final menstrual period and continues throughout the individual's remaining life span.[106]

Intervention Definitions

Fezolinetant: Fezolinetant is a once daily oral nonhormonal therapy being investigated for the treatment of moderate to severe VMS associated with menopause. It acts by regulating neuronal activity in the hypothalamus thereby affecting temperature regulation. If approved, it would be the first selective neurokinin-3 (NK3) receptor antagonist available in the US. On June 23rd, 2022, Astellas submitted a New Drug Application for fezolinetant 45 mg to the FDA.

Neurokinin-3 (NK3) receptor antagonist: A small molecule that blocks the NK3 receptor. Theories suggest that VMS are caused by a loss of thermoregulatory control that coincides with altered kisspeptin, neurokinin B, and dynorphin (KNDy) signaling. KNDy neurons are stimulated by NKB, a hypothalamic neuropeptide that regulates the female reproductive axis[107], and inhibited by estrogen.[108] During menopause, the decline in estrogen levels disrupts the KNDy neurons and thus a NK3 receptor antagonist may regulate KNDy neurons and prevent the reduction in core body temperature experienced during VMS.

Gabapentin: Gabapentin is a medication used off label as a nonhormonal treatment for VMS associated with menopause and is typically administered at a dose of 300 mg three times per day.[16]

Pregabalin: Pregabalin is a medication used off label as a nonhormonal treatment for VMS associated with menopause and is typically administered at a dose of 300 mg per day.[16]

Selective serotonin reuptake inhibitors (SSRIs) and Serotonin-norepinephrine reuptake inhibitors (SNRIs): SSRIs and SNRIs are two class of medications used as a nonhormonal treatment for VMS associated with menopause.[16] The only medication in these two classes that has an FDA approved indication for the treatment of VMS associated with menopause is paroxetine (Brisdelle, a SSRI).[19] However, several medications, such as escitalopram, venlafaxine, desvenlafaxine, citalopram, etc., are used off label.

Menopausal hormone therapy (MHT): MHT includes estrogen alone and estrogen and progestin or progesterone drug products that have FDA approved indications for the treatment of moderate to severe VMS associated with menopause. Estrogen and progestin or progesterone are typically administered for women who have an intact uterus. For women who have had a hysterectomy, estrogen alone can be used.

Outcome Measure Definitions

Vasomotor symptoms (VMS): VMS is characterized by hot flashes and night sweats. VMS is thought to be brought on by decreased estrogen levels and increased NKB activity acting on the hypothalamus, a region of the brain which regulates body temperature. The change in hypothalamic thermoregulation increases blood flow to the skin, resulting in the VMS.

Severity of Vasomotor Symptoms: The severity of VMS are defined clinically by the Food and Drug Administration (FDA) as follows:

- 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

Hot flashes: Hot flashes, also known as hot flushes are the sudden onset of heat in the upper chest and face which spreads throughout the body, and they typically last two to four minutes. Hot flashes are often accompanied by profuse sweating which can occur at night (night sweats) and cause sleep disruption and negatively affect mood.

Climacteric symptoms: Climacteric symptoms are symptoms experienced shortly before and during menopause. Climacteric is the period of life starting from the decline in ovarian activity until after the end of ovarian function. Climacteric symptoms typically cover vasomotor symptoms (hot flashes, diaphoresis) and vaginal dryness but they can also include sleep disturbances, mood changes, urinary tract symptoms, and sexual problems (loss of libido, dyspareunia, etc.).

Frequency of Moderate to Severe VMS: Typically compares differences from baseline to a follow-up time point (e.g., week 8-16). Baseline frequency is measured as the daily or weekly average number of moderate (sensation of heat with sweating and able to continue activity) to severe (sensation of heat with sweating causing cessation of activity) VMS based on the non-missing values. Follow-up timepoint frequency were measured as the daily or weekly frequency at week 8-16.

Mean Severity of Moderate to Severe VMS Per Day: Mean severity is measured variably across trials. A description of the different measurements used across trials is reported in Table A.1. of this supplement.

Patient-reported Outcomes Measurement Information System Sleep Disturbance - Short Form 8b (PROMIS SD SF 8b)[109]: Assesses self-reported sleep disturbance over the past 7 days and includes perceptions of restless sleep; satisfaction with sleep; refreshing sleep; difficulties sleeping, getting to sleep or staying asleep; amount of sleep; and sleep quality. Responses to each of the 8 items range from 1 (no disturbed sleep) to 5 (disturbed sleep), and the range of possible summed raw scores is 8 to 40. Higher scores on the PROMIS SD SF 8b indicate more of the disturbed sleep.

Patient's Global Impression of Change (PGIC)[110]: PGIC in VMS is a 1 item instrument that asks: "Compared to the beginning of this study, how would you rate your hot flashes/night sweats now?" Subject ratings range from (1) much better to (7) much worse with 4 indicating no change.

Menopause-Specific Quality of Life Questionnaire (MENQoL)[111]: the MENQoL is a 29-item tool used to assess health-related quality of life in the immediate post-menopausal period, covering four domains of menopausal symptoms (vasomotor, psychosocial, physical, and sexual domains).[112] The MENQOL is self-administered and asks the subject if they have experienced the specific problem in the past month and, if so, how bother have they been by the problem on a scale of 0 (not at all bothered) to 6 (extremely bothered). Higher scores indicate worse symptoms.

Urogenital menopausal symptoms: Urogenital menopausal symptoms refer to changes to the labia, clitoris, vagina, urethra, and bladder that occur in menopausal women due to reduced estrogen levels. This estrogen deficiency leads to a decrease in blood flow to the vagina and vulva resulting in atrophy, decreased vaginal lubrication, discharge, itching. Such symptoms are a major cause of pain during or after sexual intercourse for menopausal women.

Aspartate aminotransferase (AST)/alanine aminotransferase (ALT): AST and ALT are liver enzymes that serve as biomarkers of liver damage.

Other Relevant Definitions

Health Improvement Distribution Index: The Health Improvement Distribution Index identifies a subpopulation that has a higher prevalence of the disease of interest and therefore, creates an opportunity for proportionately more health gains within the subpopulation. This opportunity may be realized by achieving equal access both within and outside the identified subpopulation to an intervention that is known to improve health. The Health Improvement Distribution Index is defined as the disease prevalence in the subpopulation divided by the disease prevalence in the overall population. For example, if the disease prevalence was 10% in poor Americans whereas the disease prevalence across all Americans was 4%, then the Health Improvement Distribution Index would be 10%/4% = 2.5. For interventions known to increase health in this disease and that accomplish equal access across the entire population, poor Americans would receive 2.5 times the health improvements as compared to the same sized group of Americans without regard to economic status. Health Improvement Distribution Indexes above 1 suggest that more health may be gained on the relative scale in the subpopulation of interest when compared to the population as a whole. This statistic may be helpful in characterizing a treatment's contextual considerations and potential other benefits (Section 5).

For this calculation, we used data from a longitudinal analysis of VMS and race and ethnicity measured within a Study of Women's Health Across the Nation (SWAN) from 1996 to 1997.[105] The study consisted of individuals 42 to 52 years of age with an intact uterus, with 3,288 women ultimately included in the analysis. The racial and ethnic groups included in this analysis based on available data included White, African American/Black, Hispanic, Chinese, and Japanese. The provided odds ratios were converted to risk ratios using data provided within the publication and one additional source [113] that ultimately fed into an estimate of baseline risk; baseline risk for VMS in at-risk menopausal women was estimated at 50.8%. The resulting risk ratios were then weighted by population weights available within US Census data. [114] These adjusted risk ratios functionally equate to the Health Improvement Distribution Index, as they both describe risk (or prevalence) in an at-risk population divided by the risk for the general population. We performed calculations for all reported races/ethnicities, but only report here that subgroup with a risk ratio greater than 1:

- African American/Black: = (1.63/((1-0.508)+0.508*1.63))/0.966 = 1.3

Table A.1. Definitions of Severity Across the Main Trials

Title	Trial Name, First Author, Date	VMS Severity Definition
Fezolinetant	Phase 2a: Hot flashes[108]	The mean daily total VMS score during a given period was calculated by multiplying the number of mild, moderate, or severe VMS episodes during the period by 1, 2, or 3, respectively, summing the values and dividing by the number of days in the period.
	VESTA: Fraser, G.L., Santoro, N. 2020 SKYLIGHT 1: Lederman et al., 2022 SKYLIGHT 2: Johnston et al. 2021, 2022[25,26,54,99,115]	The moderate/severe VMS severity per day was determined by the following calculation: [(number of moderate VMS X 2) + (number of severe VMS X 3)] / (number of moderate + number of severe VMS).
Standard-dose Estrogen	Lin et al. 2011[37]	VMS were rated from 0 to 3, where 0 indicated no symptoms and 3 represented a sensation of heat with sweating causing cessation of activity (severe) and an average score was taken from only moderate-severe VMS.
Low-dose Estrogen	MsFLASH 03: Joffe et al. 2014[38]	VMS severity was rated as 1=mild to 3=severe and a daily average was calculated.
	Archer et al. 2013[39]	The mean severity of moderate to severe VMS on each day was calculated as: [(2 number of moderate VMS) + (3 number of severe VMS)] / (total number of moderate to severe VMS). A daily mean daily severity was calculated by averaging the daily severity of moderate to severe HF across the week.
	Panay et al. 2007[40]	Hot Flush Weekly Weighted Score (HFWWS) is calculated by summing: (number of mild VMS X 1) + (number of VMS x 2) + (number of severe VMS x 3)
SSRIs/SNRIs	Archer et al. 2009[49]	Daily VMS severity = [(number of mild VMS *1) + (number of moderate VMS *2) + (number of severe VMS *3)] divided by the total number of VMS on that day.
	Archer et al. 2009[45]	Daily VMS severity = [(number of mild VMS *1) + (number of moderate VMS *2) + (number of severe VMS *3)] divided by the total number of VMS on that day.
	Speroff et al. 2008[46]	The average daily VMS severity score was calculated as follows: ([number of mild VMS*1] + [number of moderate VMS*2] + [number of severe VMS*3])/total number of VMS on that day
	Pinkerton et al. 2013[48]	The average daily severity of VMS for each week was the sum of the number of VMS weighted by severity (1, mild; 2, moderate; 3, severe) divided by the number of days in that week with data.
SSRIs/SNRIs	Bouchard et al. 2012[47]	The sum of the daily severity scores divided by the number of days with data, with the daily severity score calculated as: (number of mild VMS*1) + (number of moderate VMS*2) + (number of severe VMS*3) divided by the total number of VMS.
	Evans et al. 2005[44]	Scoring was on a scale from 1 to 4, with 1 being mild, 2 being moderate, 3 being severe, and 4 being very severe. VMS scores are defined by multiplying the VMS frequency times the average VMS score (scaled from 1 through 4 by patient report) with 1 through 4 severities, respectively, being applied to definitions of mild, moderate, severe, and very severe.

Title	Trial Name, First Author, Date	VMS Severity Definition
	Simon et al. 2013 (Study 1 & 2)[42]	Weekly VMS severity score was calculated as: [(2 number of moderate VMS) + (3 number of severe VMS)] / (total number of moderate to severe VMS)
	MsFLASH 01: Freeman et al. 2011[43]	VMS severity was rated from 1 to 3 (mild, moderate, severe) and a daily average was provided.
Gabapentin	Pinkerton et al. 2014[51]	VMS severity was defined as follows: 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). A daily average score was calculated.
	Reddy et al. 2006[52]	VMS were rated from 1 to 4, was designated for each VMS based on the level of severity (1 for mild to 4 for very severe). The composite score was the weekly summation of the number of VMS in each severity category multiplied by the severity score assigned to each VMS.
	Guttuso et al. 2003[53]	Each VMS was recorded by filling in the appropriate severity bubble on a scale of 1 to 7. A score was calculated by adding the VMS severity scores over a week and dividing by the number of days for which completed diaries were received.

HFWWS: Hot Flush Weekly Weighted Score, VMS: vasomotor symptoms.

A2. Potential Cost-Saving Measures for Moderate to Severe Vasomotor Symptoms Associated with Menopause

ICER includes in its reports information on wasteful or lower-value services in the same clinical area that could be reduced or eliminated to create headroom in health care budgets for higher-value innovative services (for more information, see https://icer.org/our-approach/methods-process/value-assessment-framework/). These services are ones that would not be directly affected by therapies for VMS, as these services will be captured in the economic model. Rather, we are seeking services used in the current management of VMS beyond the potential offsets that arise from a new intervention. During stakeholder engagement and public comment periods, ICER encouraged all stakeholders to suggest services (including treatments and mechanisms of care) currently used for patients with VMS that could be reduced, eliminated, or made more efficient. No suggestions were received.

B. Patient Perspectives: Supplemental Information

B1. Methods

We spoke with five patients experiencing VMS, both from surgical and natural menopause, and representatives from three patient advocacy organizations.

The conversations were informed by a semi-structured interview guide, which focused the conversation on several themes:

1. General experience with VMS associated with menopause
2. Experience seeking relief and treatment
3. Patient preferences regarding potential future treatments

After each of these conversations, patient comments were transcribed, collated, organized, and summarized. We drew upon themes that emerged from our conversations and summaries are included in the Patient and Caregiver Perspectives section of the report.

We supplemented the interviews with data from an online survey conducted by the National Menopause Foundation. The survey was conducted for 2 weeks starting on August 14, 2019 targeting 5,000 women in the National Menopause Foundation database and via SurveyMonkey list purchase. The target age was 45-65 and there were 229 responses.

C. Clinical Guidelines

Clinical practice guidelines for the treatment of VMS have been issued by several US and non-US-based organizations. These guidelines are summarized below.

The North American Menopause Society (NAMS)[5]

In 2022, NAMS released a position statement and evidence ratings on the use of hormone therapy. Details of the evidence ratings can be found in the position statement. Key points are:

1. MHT is the gold standard for relief of VMS.
2. Various formulations, doses, and routes of prescription hormone therapy preparations have comparable high efficacy for relieving VMS (good and consistent [Level I] evidence).
3. Different MHT formulation, dose, and route of administration may have different effects on target organs (limited or inconsistent [Level II] evidence).
4. MHT choice should be determined individually through shared decision-making based on symptom relief, adverse events, and patient preferences (primarily on consensus and expert opinion [Level III]).
5. MHT use should be reassessed periodically (primarily on consensus and expert opinion [Level III]).
6. The increased absolute risks associated with MHT is low, including low increased risk for venous thromboembolism, gallbladder disease, stroke and breast cancer (good and consistent [Level I] evidence).
7. MHT reduces the absolute risks for all-cause mortality, fracture, diabetes mellitus (estrogen plus progestogen therapy and estrogen therapy), and breast cancer (estrogen therapy) in women aged younger than 60 years (good and consistent [Level I] evidence).[94]

The Endocrine Society[16]

In their most recent (2015) practice guideline and evidence ratings (using the GRADE framework), the Endocrine Society makes the following statements:

1. Suggest initiating MHT for the treatment of VMS for menopausal women <60 years of age or <10 years post-menopause, who do not have contraindications or excess cardiovascular or breast cancer risks (based on low quality (Grade 2) evidence).

2. For women with mild VMS, the Endocrine Society suggests non-medication approaches such as such as turning down the thermostat, dressing in layers, avoiding alcohol and spicy foods, and reducing obesity and stress (low quality (Grade 2) evidence).

3. For women with moderate to severe VMS who have a contraindication to MHT or who refuse MHT, the Endocrine Society suggests nonhormonal treatments: SSRIs, SNRIs, clonidine, gabapentin, or pregabalin (very low quality (Grade 1) evidence).

4. The Endocrine Society also suggests that providers counsel women on the lack of consistent evidence for over-the counter (OTC) or complementary medicine therapies (low quality (Grade 2) evidence).

American College of Obstetricians and Gynecologists (ACOG)[96,97]

In their most recent (2014) practice guideline and rating of evidence, ACOG has provided recommendations for the treatment of VMS.

1. Systemic MHT is the most effective treatment for VMS and that patients be treated with the lowest dose and for the shortest period possible (good or consistent evidence).

2. Nonhormonal treatments that are effective include: SSRIs, SNRIs, clonidine, and gabapentin (good or consistent evidence).

3. There is limited or inconsistent evidence for progestin-only medications, testosterone, compounded bioidentical hormones, phytoestrogens, herbal supplements, or lifestyle modifications but particular lifestyle modifications may be considered: layering clothing, maintaining a lower ambient temperature, drinking cool liquids, and avoiding alcohol and caffeine.

The National Institute for Health and Care Excellence (NICE)[98]

In their most recent (2019) guideline, NICE has provided assessments regarding the long-term risks or benefits of recommending MHT.

1. The risk of venous thromboembolism:
    a. is increased over baseline population risk with oral MHT.
    b. is not increased over baseline population risk with transdermal MHT.
2. The risk of cardiovascular disease or death from cardiovascular disease does not increase with MHT for women under the age of 60.
3. Estrogen and progesterone is associated with increased risk of breast cancer but estrogen only is not associated with increased risk.
4. MHT is not associated with risk of developing type 2 diabetes.
5. MHT is associated with reduced bone fractures.

NICE makes the following recommendations for the treatment of VMS:

1. Women should be offered MHT after discussing the short- and long-term risks and benefits.
2. SSRIs, SNRIs and clonidine should not be offered as first line treatment.
3. Isoflavones and black cohosh have some evidence but there is substantial uncertainty due to the different preparations on the market, unknown safety, and potential drug-drug interactions.

D. Comparative Clinical Effectiveness: Supplemental Information

D1. Detailed Methods

PICOTS

Population

The population of focus for the review is women seeking relief from VMS associated with menopause.

Data permitting, we will evaluate the evidence for subpopulations defined by:

- Sociodemographic factors (e.g., age, race, and ethnicity)
- Weight/Body Mass Index (BMI)
- Women who are not eligible for menopausal hormone therapy due to contraindications (e.g., history of breast cancer, blood clots, etc.)
- Women who have experienced surgical menopause

Interventions

The intervention of interest for this review is:

- Fezolinetant (Astellas Pharma Inc.)

Comparators

Data permitting, we intend to compare fezolinetant to:

- No pharmacologic treatment (prescription nor non-prescription, as estimated by the placebo arm of clinical trials)
- Menopausal Hormone Therapy (MHT)
    - Estrogen and progestin or progesterone
    - Estrogen only
- SSRIs/SNRIs that have been studied for VMS symptoms only
- Gabapentin
- Pregabalin

Outcomes

Data permitting, we will evaluate the outcomes described in the list below.

- Patient-Important Outcomes
    - Frequency and severity of vasomotor symptoms
    - Sleep quality
    - Quality of life
    - Interference of symptoms with daily life
    - Functional impairment (e.g., work productivity)
    - Urogenital menopausal symptoms that may be addressed by the intervention or comparators
    - Other patient-reported outcomes (e.g., mood changes)
- Adverse events (AEs) including but not limited to
    - Serious AEs
    - Discontinuation due to AEs
    - Other AEs including but not limited to
        - Endometrial hyperplasia or cancer
        - Bone density markers (e.g., fractures, osteoporosis)
        - Breast cancer
        - Coronary heart disease
        - Venous thromboembolism (e.g., pulmonary embolism)
        - Stroke
        - Colorectal cancer
        - Liver toxicity (e.g., AST and ALT levels)
        - All-cause mortality
        - Suicidality

Timing

Evidence on intervention effectiveness and evidence on harms will be derived from studies of any duration.

Settings

Vasomotor symptoms associated with menopause are generally treated in outpatient and/or clinic settings, which will be the focus of our review.

Table D1.1 PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item
TITLE		
Title	1	Identify the report as a systematic review.
ABSTRACT		
Abstract	2	See the PRISMA 2020 for Abstracts checklist.
INTRODUCTION		
Rationale	3	Describe the rationale for the review in the context of existing knowledge.
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.
METHODS		
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.

Section and Topic	Item #	Checklist item
Synthesis methods	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.
RESULTS		
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.
Study characteristics	17	Cite each included study and present its characteristics.
Risk of bias in studies	18	Present assessments of risk of bias for each included study.
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.
Results of syntheses	20a	For each synthesis, briefly summarize the characteristics and risk of bias among contributing studies.
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g., confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.
	20c	Present results of all investigations of possible causes of heterogeneity among study results.
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.
DISCUSSION		
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.
	23b	Discuss any limitations of the evidence included in the review.
	23c	Discuss any limitations of the review processes used.
	23d	Discuss implications of the results for practice, policy, and future research.
OTHER INFORMATION		
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.

Section and Topic	Item #	Checklist item
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.
	24c	Describe and explain any amendments to information provided at registration or in the protocol.
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.
Competing interests	26	Declare any competing interests of review authors.
Availability of data, code, and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.

From: Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *PLoS Med.* 2021;18(3):e1003583.

Data Sources and Searches

Procedures for the systematic literature review assessing the evidence on fezolinetant and comparators (e.g., no pharmacologic treatment, MHT [e.g., estrogen and progestin or progesterone, or estrogen alone], SSRIs/SNRIs, gabapentin, and pregabalin) for moderate to severe VMS associated with menopause followed established best methods.[116,117] During the scoping phase, we identified two network meta-analyses for SSRIs[21] and menopausal hormone therapy[118] that matched our protocol. We abstracted data from these two network meta-analyses for trials that met our inclusion criteria and conducted an updated literature search for new evidence published since the last search. The review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[119] The PRISMA guidelines include a list of 27 checklist items.

We searched MEDLINE, EMBASE, Cochrane Database of Systematic Reviews, and Cochrane Central Register of Controlled Trials for relevant studies. Each search was limited to English language studies of human subjects and excluded articles indexed as guidelines, letters, editorials, narrative reviews, case reports, or news items. All search strategies were generated utilizing the Population, Intervention, Comparator, and Study Design elements described above. The search strategies included a combination of indexing terms (MeSH terms in MEDLINE and EMTREE terms in EMBASE), as well as free-text terms. We also conducted a targeted search for longer-term adverse event outcomes for MHT.

To supplement the database searches, we performed a manual check of the reference lists of included trials and reviews and invite key stakeholders to share references germane to the scope of this project. We also supplemented our review of published studies with data from conference proceedings, regulatory documents, and information submitted by manufacturers, and other grey literature when the evidence met ICER standards (for more information, see https://icer.org/policy-on-inclusion-of-grey-literature-in-evidence-reviews/. Where feasible and deemed necessary, we also accepted data submitted by manufacturers "in-confidence," in accordance with ICER's published guidelines on acceptance and use of such data (https://icer.org/guidelines-on-icers-acceptance-and-use-of-in-confidence-data-from-manufacturers-of-pharmaceuticals-devices-and-other-health-interventions/).

Table D1.2. Search Strategy of OVID MEDLINE® Epub Ahead of Print 1946 to Present and Cochrane Central Register of Controlled Trials and Systematic Reviews (fezolinetant, gabapentin, pregabalin, SNRIs)

#	Search Term
1	exp menopause/
2	"change of life, female/" or "postmenopause" or "perimenopause" or "flashes, hot" or "climacteric" or "systems, vasomotor"
3	1 or 2
4	(fezolinetant or esn364).ti,ab.
5	3 and 4
6	(Pregnenedione or neurontin or convalis or "gabapentin hexal" or "gabapentin stada" or novogabapentin or gabapentin).ti,ab
7	(Lyrica or "CI 1008" or "CI-1008" or CI1008 or pregabalin).ti,ab
8	("Serotonin and Norepinephrine Reuptake Inhibitors" or "SSRIs and NRIs" or "NRIs and SSRIs" or "SSRIs and SNRIs" or "SNRIs and SSRIs" or "Serotonin and Noradrenaline Uptake Inhibitors" or "SNRIs" or "SNRI" or "SSNRI" or "Serotonin and Norepinephrine Uptake Inhibitors" or "serotonin noradrenalin reuptake inhibitor" or "venlafaxine" or "desvenlafaxine").ti,ab
9	6 or 7 or 8
10	3 and 9
11	("clinical trial" or "comparative study" or "randomized controlled study" or "multicenter study" or "clinical trial, phase III" or "controlled clinical trial" or "meta analysis" or "meta-analysis" or "RCT" or "systematic literature review" or "SLR" or "randomized controlled trial" or "systematic review").pt.
12	10 and 11
13	(animals not (humans and animals)).sh.
14	12 not 13
15	Limit 14 to English Language
16	5 or 15

Table D1.3. Search Strategy of EMBASE SEARCH (fezolinetant, gabapentin, pregabalin, SNRIs)

#	Search Term
1	menopause'/exp OR menopause OR 'postmenopause' OR 'perimenopause' OR 'hot flashes' OR 'climacteric' OR 'climacterum' OR 'vasomotor nervous system'
2	('fezolinetant' OR 'as347269300' OR 'esn364' OR 'a2693'):ti,ab
3	#1 and #2
4	('ci945' OR 'dineurin' OR 'dm1796' OR 'dm5689' OR 'gabalept' OR 'gabaliquid' OR 'gabapen' OR 'gabatin' OR 'go3450' OR 'goe3450' OR 'neurotoni' OR 'gabapentin'):ti,ab
5	('ci 1008' OR 'ci1008' OR 'lyrica' OR 'lyrica cr' OR 'pd 144723' OR 'pd144723' OR 'pregabalin'):ti,ab

6	('serotonin noradrenalin reuptake inhibitor' OR 'serotonin norepinephrine reuptake inhibitor' OR 'serotonin norepinephrine uptake inhibitor' OR 'SNRI' OR 'SNRIs' OR 'SSNRI' OR 'NRI' OR 'venlafaxine' OR 'desvenlafaxine'):ti,ab
7	#4 or #5 or #6
8	#1 and #7
9	#8 AND ([cochrane review]/lim OR [systematic review]/lim OR [meta analysis]/lim OR [randomized controlled trial]/lim OR 'controlled clinical trial'/de OR 'Randomized Clinical Trial'/de)
10	#9 AND [english]/lim
11	#10 AND [medline]/lim
12	#3 or #11

Table D1.4. Search Strategy of OVID MEDLINE® Epub Ahead of Print 1946 to Present and Cochrane Central Register of Controlled Trials (SSRIs updated since 2013 Network Meta-Analysis[21])

#	Search Term
1	exp menopause/
2	"change of life, female/" or "postmenopause" or "perimenopause" or "flashes, hot" or "climacteric" or "systems, vasomotor"
3	1 or 2
4	("Uptake Inhibitors" or "5-HT Uptake Inhibitors" or "5 HT Selective Serotonin Reuptake Inhibitors" or "Selective Serotonin Reuptake Inhibitor" or "SSRI" or "paroxetine" or "escitalopram" or "citalopram" or "fluoxetine" or "sertraline").ti,ab
5	3 and 4
6	(animals not (humans and animals)).sh.
7	5 not 6
8	limit 7 to yr="2013 -Current"

Table D1.5. Search Strategy of EMBASE search (SSRIs updated since 2013 Network Meta-Analysis[21])

#	Search Term
1	menopause'/exp OR menopause OR 'postmenopause' OR 'perimenopause' OR 'hot flashes' OR 'climacteric' OR 'climacterum' OR 'vasomotor nervous system'
2	('antidepressants, serotonin specific reuptake inhibitors' OR 'selective serotonin reuptake inhibitor' OR 'serotonin reuptake inhibitor' OR 'SSRI' OR 'SSRI antidepressant' OR 'paroxetine' OR 'escitalopram' OR 'citalopram' OR 'fluoxetine' OR 'sertraline'):ti,ab
3	#1 AND #2
4	#3 AND [medline]/lim
5	#4 AND [01-03-2013]/sd NOT [07-07-2022]/sd

Table D1.6. Search Strategy of OVID MEDLINE® Epub Ahead of Print 1946 to Present and Cochrane Central Register of Controlled Trials (Menopausal Hormone Therapy [MHT] updated since 2017 Network Meta-Analysis[118])

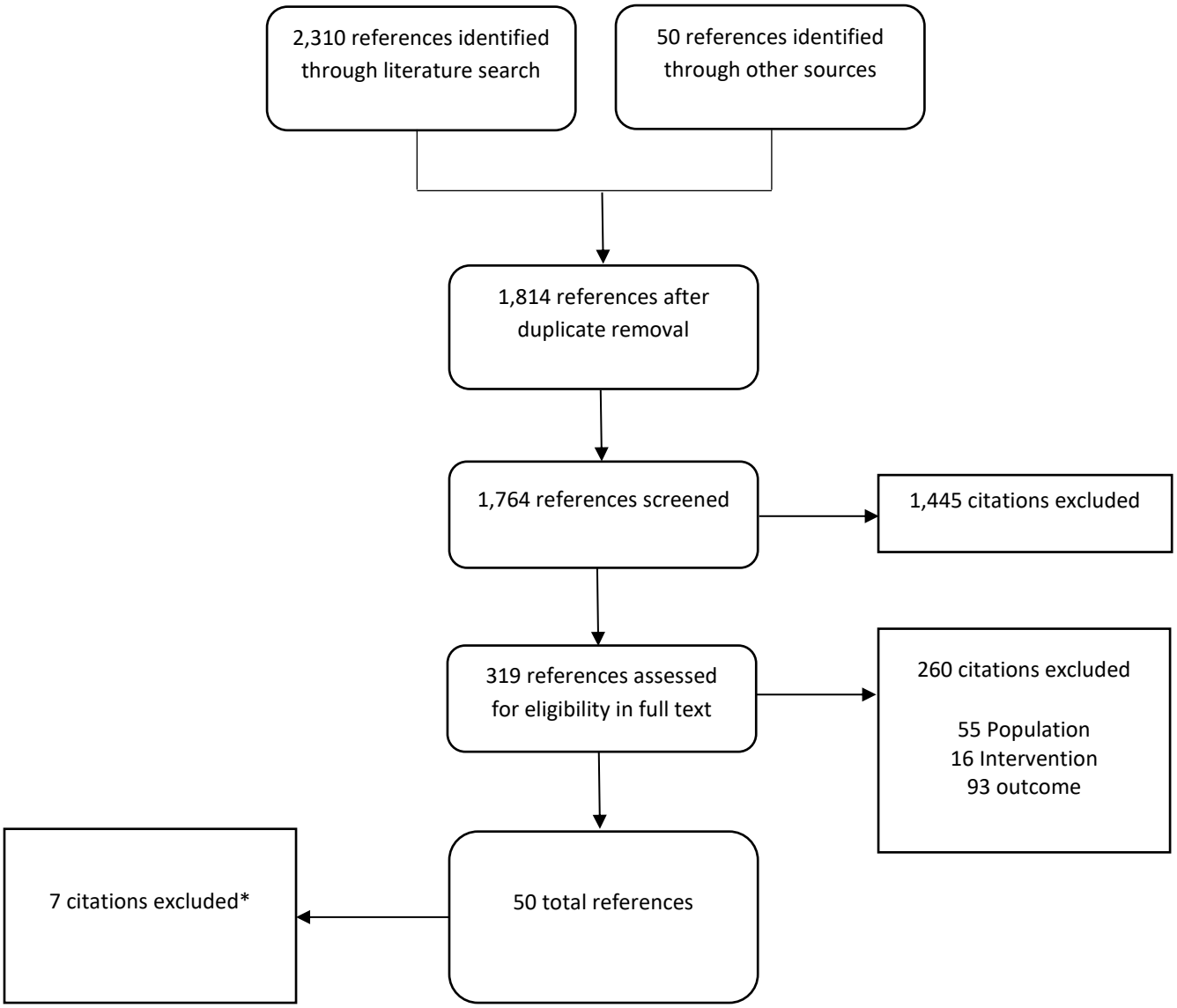
#	Search Term
1	exp menopause/
2	"change of life, female/" or "postmenopause" or "perimenopause" or "flashes, hot" or "climacteric" or "systems, vasomotor"
3	1 or 2
4	(angeliq or oestradiol or "estraderm TTS" or "estradiol valerate" or delestrogen or ovocyclin or loestrin or LoDOse or "Ethynyl Estradiol" or Lynoral or Estinyl or Ethinyloestradiol or "Estradiol 17 beta" or "Estradiol Anhydrous" or Oestradio or Estrace or "Estraderm TT").ti,ab
5	(microgynon or trikvilar or triregol or aviane or gynatrol or "HRP 102" or HRP102 or mesigyna or "estracomb TTS").ti,ab
6	("estrogen replacement therapies" or "estrogen replacement" or "estrogen replacements").ti,ab
7	(bedol or climaval or elleste or estraderm or estradot or evorel or femseven or oestrogel or progynoval or prognova or sandrena or zumenon or Estrace or Menest or Premarin or Prempro or Prefest or Activella or ambaelz or mimvey or FemHRT or "jevantique lo" or Jinteli or Duavee or Alora or Minivelle or "Vivelle-Dot" or Climara or Menostar or "Combi-Patch" or "Climara Pro" or EstroGel or Elestrin or Divigel or EvaMist or Femring or "Depo-Estradiol" or Delestrogen or Estring or Vagifem or Yuvafem or Estrace or climagest or climesse or clinorette or femoston or indivina or kliofem or kliovance or novofem or nuvelle or tridestra or trisequens or premique or premak or levonorgestrel or drospirenone).ti,ab
8	HORMONE REPLACEMENT THERAPY/ or ESTROGEN REPLACEMENT THERAPY/
9	(hormon$ adj3 substit$).ti,ab.
10	(HRT or HT or MHT or MPA).ti,ab.
11	('menopausal hormone' adj2 therap$).ti,ab.
12	exp ESTRIOL/
13	ESTROGENS/ or ESTROGENS, NON-STEROIDAL/
14	ESTRADIOL/ or ESTRAMUSTINE/
15	ETHINYL ESTRADIOL/ or ETHINYL ESTRADIOL-NORGESTREL COMBINATION/
16	(oestrogen? or estrogen? or estradiol?).ti,ab.
17	PROGESTOGENS/
18	progesta$.ti,ab.
19	MEDROXYPROGESTERONE ACETATE/
20	exp PROGESTERONE/
21	ETHISTERONE/
22	NORETHINDRONE/ or NORGESTREL/ or ETHINYL ESTRADIOL-NORGESTREL COMBINATION/ or LEVONORGESTREL/ or NORPROGESTERONES/
23	4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22
24	3 and 23
25	("clinical trial" or "comparative study" or "randomized controlled study" or "multicenter study" or "clinical trial, phase III" or "controlled clinical trial" or "meta analysis" or "meta-analysis" or "RCT" or "systematic literature review" or "SLR" or "randomized controlled trial" or "systematic review").pt.
26	24 and 25
27	(animals not (humans and animals)).sh.

28	26 not 27
29	Limit 28 to English Language
30	limit 29 to yr="2015-Current"

Table D1.7. Search Strategy of EMBASE (Menopausal Hormone Therapy [MHT] updated since 2017 Network Meta-Analysis)[118]

#	Search Term
1	menopause'/exp OR menopause OR 'postmenopause' OR 'perimenopause' OR 'hot flashes' OR 'climacteric' OR 'climacterum' OR 'vasomotor nervous system'
2	('gestogen' OR 'progestagen' OR 'progestational agent' OR 'progestational drug' OR 'progestational hormones' OR 'progestine' OR 'progestins' OR 'progestogen' OR 'alpha estrogen' OR 'alpha oestrogen' OR 'beta estrogen' OR 'beta oestrogen' OR 'estrogen uptake' OR 'estrogene' OR 'oestrogen' OR 'oestrogen uptake' OR 'oestrogene' OR 'oestrogenic agent'):ti,ab
3	('progestone' OR 'progestronaq' OR 'progiron' OR 'prolidon' OR 'prolutex' OR 'proluton' OR 'ultrogestan' OR 'uterogestan' OR 'utrogestan'):ti,ab
4	(bedol or climaval or elleste or estraderm or estradot or evorel or femseven or oestrogel or progynoval or prognova or sandrena or zumenon or 'Estrace' or 'Menest' or 'Premarin' or 'Prempro' or 'Prefest' or 'Activella' or 'ambaelz' or 'mimvey' or 'FemHRT' or 'jevantique lo' or 'Jinteli' or 'Angeliq' or 'Duavee' or 'Alora' or 'Minivelle' or 'Vivelle-Dot' or 'Climara' or 'Menostar' or 'Combi-Patch' or 'Climara Pro' or 'EstroGel' or 'Elestrin' or 'Divigel' or 'EvaMist' or 'Femring' or 'Depo-Estradiol' or 'Delestrogen' or 'Estring' or 'Vagifem' or 'Yuvafem' or 'Estrace' or limagest or climesse or clinorette or femoston or indivina or kliofem or kliovance or novofem or nuvelle or tridestra or trisequens or drospirenone or premique or premak):ti,ab
5	HORMONE SUBSTITUTION'/exp or 'ESTROGEN THERAPY'/exp
6	(HRT or HT or MHT):ti,ab
7	("menopausal hormone" adj2 therap$):ti,ab
8	ESTRIOL/exp
9	ESTROGEN/exp
10	(oestrogen? or estrogen? or estradiol?):ti,ab
11	GESTAGEN/exp
12	progest$:ti,ab
13	#2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12
14	#1 AND #13
15	#14 AND ([cochrane review]/lim OR [systematic review]/lim OR [meta analysis]/lim OR [randomized controlled trial]/lim OR 'controlled clinical trial'/de OR 'Randomized Clinical Trial'/de)
16	#15 AND [English]/lim
17	#16 AND [medline]/lim
18	#17 AND [01-01-2015]/sd NOT [07-06-2022]/sd

Figure D1. PRISMA flow Chart Showing Results of Literature Search for Fezolinetant, Gabapentin, Pregabalin, SNRIs, SSRIs, and HRT



*7 studies were excluded due to trial length or population and are described below in section D.

Study Selection

We performed screening at both the abstract and full-text level. Two reviewers independently screened the titles and abstracts of all publications identified using Nested Knowledge (Nested Knowledge, Inc, St. Paul, MN) and a third reviewer resolved any issues of disagreement through consensus. We did not exclude any study at abstract-level screening due to insufficient information. For example, an abstract that did not report an outcome of interest would be accepted for further review in full text. We retrieved the citations that were accepted during abstract-level screening for full text appraisal. Two investigators independently reviewed full

papers and provided justification for exclusion of each excluded study, according to the PICOTS elements.

Data Extraction and Quality Assessment

We examined the risk of bias for each trial using the revised Cochrane risk-of-bias tool for randomized trials (RoB 2)[120] and guidance criteria published by Higgins et al (2019).[121] Risk of bias was assessed for each of the following aspects of the trials: randomization process, deviation from the intended interventions, missing outcome data, measurement of the outcome, selection of the reported results, and overall risk of bias. To assess the risk of bias in trials in the report, we rated the categories as: "low risk of bias", "some concerns", or "high risk of bias". Guidance for risk of bias ratings using these criteria is presented below:

Low risk of bias: *The study is judged to be at low risk of bias for all domains for this result.*

Some concerns: *The study is judged to raise some concerns in at least one domain for this result, but not to be at high risk of bias for any domain.*

High risk of bias: *The study is judged to be at high risk of bias in at least one domain for this result* or *the study is judged to have some concerns for multiple domains in a way that substantially lowers confidence in the result.*

Assessment of Level of Certainty in Evidence

We used the ICER Evidence Rating Matrix to evaluate the level of certainty in the available evidence of a net health benefit among each of the interventions of focus (see Appendix D).[122,123]

Assessment of Bias

As part of our quality assessment, we evaluated the evidence base for the presence of potential publication bias. Given the emerging nature of the evidence base for these newer treatments, we scanned the ClinicalTrials.gov site to identify studies completed more than two years ago. Search terms include menopause or change of life and fezolinetant, as347269300, esn364, or a2693, or each comparator's generic name and, when available, brand name. We selected studies which would have met our inclusion criteria, and for which no findings have been published. We will provide qualitative analysis of the objectives and methods of these studies to ascertain whether there may be a biased representation of study results in the published literature.

Data Synthesis and Statistical Analyses

The studies were summarized in the text and in evidence tables of the Evidence Report. This summary is key to understanding the evidence base pertaining to the topic. Any key differences between the studies in terms of the study design, patient characteristics, interventions (including dosing and frequency), outcomes (including definitions and methods of assessments), and study quality was noted in text of the report. For all comparators, we only include trials/arms of trials that contained a dose considered to be clinically effective. For MHT specifically, we only included trials (or arms of clinical trials) that contained a dose of estrogen that was above 0.5 mg, regardless of progesterone dose, and, for estrogen-only MHT, we only included trials of those without a uterus to match clinical practice. We did not include trials without a placebo arm for comparison.

Table D1.8. Risk of Bias Assessment

Studies (Author, Year)*	Randomization process	Deviation from the intended interventions	Missing outcome data	Measurement of the outcome	Selection of the reported result	Overall Risk of Bias
Fezolinetant						
Depypere et al 2019[†108]	Low	Some concerns	Low	Low	Low	Some concerns
Fraser et al 2020[115]	Low	Low	Low	Low	Low	Low
MHT – standard dose estrogen 1 mg						
Schurmann et al 2004[36]	Low	Low	Low	Low	Low	Low
Endrikat et al 2007[34]	Low	Low	Low	Low	Low	Low
Lin et al 2011[37]	Low	Low	Low	Low	Low	Low
Lobo et al 2019[124]	Low	Low	Low	Low	Low	Low
MHT – low dose estrogen 0.5 mg						
Panay et al 2007[40]	Low	Low	Low	Low	Low	Low
Stevenson et al 2010[35]	Low	Low	Low	Low	Low	Low
Archer et al 2013[39]	Low	Low	Low	Low	Low	Low
Joffe et al 2014[ł38]	Low	Low	Low	Low	Low	Low
Speroff et al 1996[§41]	Low	Low	High	Low	Low	High

SNRIs						
Evans et al 2005[#44]	Low	Low	High	Low	Low	High
Speroff et al 2008[46]	Low	Low	Low	Low	Low	Low
Archer et al 2009a[64]	Low	Low	Low	Low	Low	Low
Archer et al 2009b[65]	Low	Low	Low	Low	Low	Low
Boekhout et al 2011[84]	Low	Low	Low	Low	Low	Low
Bouchard et al 2012[47]	Low	Low	Low	Low	Low	Low
Pinkerton et al 2013[48]	Low	Low	Low	Low	Low	Low
SSRIs						
Stearns et al 2003[89]	Low	Low	Low	Low	Low	Low
Kalay et al 2007[125]	Some concerns	Some concerns	Low	High	Low	High
Grady et al 2007[86]	Low	Low	Low	Low	Low	Low
Freeman et al 2011[43]	Low	Low	Low	Low	Low	Low
Simon et al 2013a[68]	Low	Low	Low	Low	Low	Low
Simon et al 2013b[68]	Low	Low	Low	Low	Low	Low
Capriglione et al 2016[85]	Low	Some concerns	Low	Low	Low	Some concerns

Gabapentin						
Guttuso et al 2003[53]	Low	Low	Low	Low	Low	Low
Reddy et al 2006[52]	Low	Low	Low	Low	Some concerns	Some concerns
Butt et al 2008[126]	Low	Low	Low	Low	Low	Low
Pinkerton et al 2014[72]	Low	Low	Low	Low	Low	Low

SNRI: serotonin and norepinephrine reuptake inhibitors, SSRI: selective serotonin reuptake inhibitors
*Risk of bias was only evaluated for published manuscripts of RCTs.
†The direction of the bias was unpredictable for both deviation from the intended interventions domain and overall risk of bias.
ŧJoffe et al includes both HRT and SSRI as interventions
§Intervention includes estrogen only
#The direction of the bias was unpredictable for measurement of the outcome domain.
For all other cases where bias was identified, it favored the experimental group over the comparator(s).

D2. Additional Clinical Evidence

The main report discusses primary sources of data to inform our review of fezolinetant for the treatment of moderate to severe VMS associated with menopause. In this supplement, we describe evidence for Phase 2 clinical trials of fezolinetant, additional data for trials included in the main report if they had additional follow-up time points beyond 16 weeks (See Table D.2.1.) and additional information on harms. We also describe two trials of SSRIs/SNRIs that exclusively included participants with a history of cancer and thus had contraindications to MHT.

As described in the main report, we only included trials of the comparators (MHT, SSRI/SNRI, gabapentin, and pregabalin) that were most comparable to the Phase 3 Skylight 1 and 2 trials for fezolinetant in terms of study design (i.e., Randomized Control Trial), relevant population (i.e., general healthy women experiencing frequent VMS associated with menopause), assessment of VMS (i.e., self-reported frequency and severity of VMS), and length of follow-up (i.e., between 8 to 16 weeks). The length of follow-up in the Skylight 1 and 2 trials were 12 weeks. We considered studies between 8 and 16 weeks to be comparable as many published trials on treatments for VMS report an early decrease in VMS that typically levels out before week 8. There were additional trials that fell outside of these criteria and thus were not included in our assessment. We describe these trials in this supplement, including reasons for exclusion and influence on our assessment. See Table D.2.2.

Phase 2 Trials

Fezolinetant versus Placebo

The Hot Flash trial was a Phase 2a randomized, placebo-controlled, double-blind trial[108] aimed to evaluate the efficacy of oral fezolinetant (90 mg twice a day) versus placebo for 12 weeks. The inclusion was the same as the Skylight 1 and 2 trials, except participants had to experience at least 49 per week and there were no restrictions on BMI. Baseline characteristics were fairly consistent with the Skylight 1 and 2 trials, except 99% of participants were White. This trial measured weekly VMS frequency and severity, instead of daily used in the Phase 2b and 3 trials. There was a greater reduction in weekly moderate-severe VMS frequency at week 12 in participants in the fezolinetant group compared to placebo (least squares mean reduction from placebo: -35.2; 95% CI: -47.6, -22.8; p=0.001), equating to a mean difference of -5.03 per day. There was a greater reduction in moderate-severe VMS severity at 12 weeks in the fezolinetant group compared to the placebo group (least squares mean reduction from placebo: -12.4; 95% CI: -17.0, -7.8; p=0.001).[108] See Table D2.1. There were significant improvements in other patient-reported outcomes in the fezolinetant group compared to placebo at week 12, such as sleep, inference of VMS in daily life, total climacteric symptoms, and global functioning. All adverse events were mild or moderate in severity and total adverse events were not significantly different between fezolinetant and placebo. The most common adverse event was gastrointestinal disorders. More participants in the fezolinetant

group had increased values for alanine transaminase (ALT), whereas the placebo group had increased values for aspartate aminotransferase (AST). All of these were mild. Full details of adverse events are found in Table D.3.9. and D.3.10. of this Supplement.

The VESTA trial was a Phase 2b randomized, double-blind, placebo-controlled, dose-ranging trial[115] aimed to evaluate the efficacy of seven doses of oral fezolinetant (15, 30, 60, or 90 mg twice daily, or 30, 60, or 120 mg once daily) versus placebo for 12 weeks. We presented results for all arms but focused mostly on the 30 mg and this was the only dose in this trial that was moved onto Phase 3. The inclusion was the same as the Skylight 1 and 2 trials and baseline characteristics were similar. The trial reported a significantly greater reduction in daily moderate-severe VMS frequency at week 12 in participants in all fezolinetant doses compared to placebo (least squares mean reduction from placebo ranged from -1.8 to -2.6). In the 30 mg dose, there was a significant reduction in daily VMS frequency compared to placebo, p=0.006. See Table D.2.1. There was also a greater reduction in daily moderate-severe VMS severity at week 12 in participants in the following fezolinetant doses: 60 mg and 90 mg twice a day and 60 mg and 120 mg once a day, compared to placebo (least squares mean reduction from placebo ranged from -0.9 to -1.4).[115] However, in the 30 mg dose, there was no significant difference in the change in VMS severity when compared to placebo, p=0.4647.

MENQoL improvements varied across the doses, with the mean difference from placebo ranging from 0 to -0.7. Three fezolinetant doses (30 mg, 60 mg, 90 mg twice daily) were associated with a significant improvement in MENQoL compared to placebo, but did not meet MCID. Participants in the 30 mg daily dose had a mean non-significant reduction in MENQoL versus placebo of -0.1; 95% CI: -0.6 to 0.3.[99] There were significantly larger improvements in hot flash related daily interference scale (HFRDIS) at week 12 in participants in three fezolinetant doses: 60 mg twice a day, 90 mg twice a day, and 120 mg once a day, and larger improvements in VMS, measured by GCS, in all fezolinetant does except 15 mg twice a day, compared to placebo.[99]. Adverse events were mostly mild to moderate in severity. However, incidences of adverse events did increase with increasing doses and more participants who received fezolinetant discontinued due to adverse events. The most common reason for discontinuing was elevated liver enzymes and there was one occurrence of drug-induced liver injury consisting of elevations in ALT in the fezolinetant 60 mg dose.[115] In addition, nine participants had ALT or AST at least 3 times the upper limit of normal, with three participants at least 8 times the upper limit of normal (fezolinetant 60 mg, 90 mg twice a day, and 60 mg once a day). Full details of adverse events are found in Table D.3.9. and D.3.10. of this Supplement.

Table. D2.1. VMS frequency and severity outcomes for Phase 2 fezolinetant trials.

Trial Name/Author	Exclusion from main report	Intervention	Arm Size	VMS Frequency		VMS Severity	
				Change from baseline: Mean (SE)	DIFF from PBO: Mean (SE/95% CI), P Value	Change from baseline: Mean (SE)	DIFF from PBO: Mean (SE/95% CI), P Value
Phase 2A: Depypere et al 2019[108]	Phase 2	Fezolinetant 90 mg	43	-76.1† (95% CI: -87.2, -65.0)	-35.2 (95% CI: -47.6, -22.8), p=0.001	-26.6† (95% CI: -31.1, -22.2)	-12.4 (95% CI: -17.0, -7.8), p=0.001
		Placebo	44	-35.3† (95% CI -46.9, -23.6)	REF	-12.1† (95% CI: -16.6, -7.7)	REF
VESTA: Fraser et al 2020 [115]	Phase 2	Fezolinetant 30 mg	43	-7.4 (0.58)	-2.1 (0.75) (95% CI: -3.52 to -0.58), p=0.0064	-0.9 (0.16)	-0.2 (0.21) (95% CI: -0.58 to 0.26), p=0.4647

Additional long-term outcomes for trials included in the main report

There were six trials that reported additional outcomes beyond 16 weeks. Five trials reported additional outcomes for VMS frequency and severity and these are reported in Table D2.2.

Menopausal Hormone Therapy versus Placebo

In the main report, we described the RCT published by Lobo et al. (2018) that evaluated low and standard-dose MHT versus placebo at week 12. This RCT also collected sleep outcomes at month 6 and 12 and the significant changes in sleep disturbance remained significant at month 6 and 12.[59]

SSRI/SNRI versus Placebo

In the main report, we described a manuscript published by Simon et al. (2013) of two RCTs that evaluated oral paroxetine (7.5 mg) versus placebo at week 12.[42] In Study 2 of Simon et al. (2011) outcomes were also assessed at 24 weeks. At week 24, the reduction in mean weekly VMS frequency remained significant between paroxetine and placebo, p=0.002, and there were significantly more responders (e.g., participants who achieved 50% or more reduction in VMS) in the paroxetine group compared to placebo, p=0.007. However, at week 24, the reduction in mean weekly VMS severity was no longer significantly different between paroxetine and placebo, p=0.053.[42] See Table D2.2.

In the main report, we described an RCT published by Pinkerton et al. (2013) that evaluated oral desvenlafaxine (100 mg) versus placebo at week 12. This RCT also had outcomes at month 6 and 12. At month 6 and 12, there were greater reductions from baseline in daily moderate-severe VMS frequency (month 6: -8.58 [SE=0.35] and month 12: -7.70 [SE=0.45], versus placebo, p<0.001 for both time points) and in daily moderate-severe VMS severity in the desvenlafaxine group as compared to placebo (month 6: -0.85 [SE=0.07] and month 12: -0.75 [SE=0.07], versus placebo, p<0.001 for both time points). Improvements in GCS scores for anxiety, depression, psychological symptoms, and vasomotor subscale were maintained at month 6 and 12 (p<0.001 for all outcomes).[48] See Table D2.2.

In the main report, we described an RCT published by Archer et al. (2009) that evaluated oral desvenlafaxine (100 mg) versus placebo at week 12. This RCT also had outcomes at week 26. At week 26, there was a significantly greater reduction in daily moderate-severe VMS frequency in participants who received desvenlafaxine 150 mg compared to placebo, p=0.001, but not those who received desvenlafaxine 100 mg, p=0.061. The significant reductions in daily moderate-severe VMS severity were maintained at week 26 for those received desvenlafaxine 150 mg compared to placebo, p=0.008, but not those who received desvenlafaxine 100 mg, p>0.05. [49] See Table D2.2.

Gabapentin versus Placebo

In the main report, we described an RCT published by Guttuso et al. (2003) that evaluated oral gabapentin versus placebo at week 12. This RCT also included an open-label Phase up to week 17 where all participants received gabapentin from week 12 to 17. Both groups (those who received gabapentin from week 1-17 and those who received placebo from week 1-12 and gabapentin from week 13-17) continued to decrease and were not significantly different from each other at week 17, p=0.82. Of note, there was a slight increase in VMS severity scores at week 13 in the gabapentin arm as they repeated the open-label titration.[53] See Table D2.2.

In the main report, we described an RCT published by Pinkerton et al. (2014) that evaluated gastroretentive gabapentin versus placebo at week 12. This RCT also had outcomes at week 24. The significant reductions in daily VMS frequency and severity in the gabapentin arm compared to placebo at week 12 was maintained at week 24 (mean difference from placebo in VMS frequency: -1.08 (95% CI: -1.98 to -0.19), p=0.017; mean difference from placebo in VMS severity: -0.22; 95%CI: -0.44 to -0.0, p=0.046). Improvement in sleep in those who received gabapentin compared to placebo was also maintained at week 24, p<0.0001.[51] See Table D2.2.

Table. D2.2. Additional long-term VMS frequency and severity outcomes for trials included in the main report.

Trial Name/Author	Intervention	Arm Size	VMS Frequency		VMS Severity	
			Change from baseline: Mean (SE)	DIFF from PBO: Mean (SE/95% CI), P Value	Change from baseline: Mean (SE)	DIFF from PBO: Mean (SE/95% CI), P Value
Archer et al. 2009[64] 26 weeks	Desvenlafaxine 100 mg	182	-61%‡ (NR)	p=0.061	-24%*‡ (NR)	p>0.05
	Desvenlafaxine 150 mg	179	-69%‡ (NR)	p=0.001	-29%*‡ (NR)	p=0.008
	Placebo	180	-51%‡ (NR)	REF	-13%*‡ (NR)	REF
Pinkerton et al. 2013[48] 6 and 12 months	Desvenlafaxine 100 mg	125, 112	6 months: -8.58 (0.35) 12 months: -7.7 (0.45)	6 months: p<0.001 12 months: -2.86 (95% CI: -4.14, -1.57), p<0.001	6 months: -0.85* (0.07) 12 months: -0.75* (0.07)	6 months: p<0.001 12 months: -0.31 (95% CI: -0.51, -0.11), p=0.003
	Placebo	124, 102	12 months: -4.8 (0.47)	REF	12 months: -0.44 (0.07)	REF
Simon et al. 2013 (Study 2)[68] Week 24	Paroxetine (7.5 mg)	284	NR*†	p=0.002	NR*	p=0.053
	Placebo	284	NR*†	REF	NR*	REF
Guttuso et al. (2003)[53]	Gabapentin 900 mg	30	-53.5% (22.0)†‡	-2.1 (-18.6, 15.6), p=0.82	67.3 (20.9)†	0 (-13.6, 16.8)
	Placebo	29	-53.7% (32.6)†‡	REF	61.3 (38.9)†	1.00
Pinkerton et al. 2014[72] Week 24	Gabapentin gastroretentive 1800 mg	299	-8.99* (0.37)	-1.08 (95% CI: -1.98 to -0.19), p=0.017	-0.86* (0.09)	-0.22 (95%CI: -0.44 to -0.0), p=0.046
	Placebo	294	-7.91* (0.36)	REF	-0.64* (0.09)	REF

CI: Confidence Interval, DIFF from PBO: Difference from placebo, NR: Not Reported, REF: Reference group, SD: Standard Deviation, SE: Standard Error, VMS: Vasomotor symptoms, mg: milligrams.

*All VMS (mild, moderate, and severe)

†Weekly score

‡Percentage change.

Additional Harms

In the main report, we broadly described harms. In this supplement, we provide additional information on harms.

Menopausal Hormone Therapy

Serious events in the trials of less than 1 year in length were low. The standard dose trials had slightly higher incidences of serious events including breast tenderness, headache, breast swelling, and ankle fracture in Lin et al., and one case of permanent bleeding due to adenomyosis uteri interna and several leiomyomata.[36] The REPLENISH trial reported adverse events at 12 months and reported occurrence of additional serious adverse events, such as acute pancreatitis, deep vein thrombosis, chronic obstructive pulmonary disease, infective cholecystitis, and breast cancer. The case of deep vein thrombosis occurred in a woman with a family history.[124] Gynecological changes were slightly more common in MHT than placebo but were not considered serious adverse events. There was little evidence of endometrial hyperplasia or malignancy[33,38-40], except one single case of endometrial hyperplasia in the REPLENISH low dose group.[61] There were reports that MHT lowered cholesterol (total and low-density lipoprotein)[33,37] and decreased bone turnover, as measured by bone-specific alkaline phosphatase (BSAP), N-terminal propeptide of type I procollagen (P1NP), and C-terminal telopeptide of type I collagen (CTX-1)[127], highlighting two potential benefits of MHT. Full details of adverse events are described in Table D3.9-10 in the Supplement.

SSRIs/SNRIs

For desvenlafaxine, serious AEs were reported in all six desvenlafaxine trials, including intestinal obstruction[47], increased liver function test values and cholecystitis[46], hypertension[45,46,49], and bronchospasm.[45] Speroff et al. (2008) reported two cardiovascular events in the desvenlafaxine arms: coronary occlusion with revascularization and myocardial infarction, but these occurred in women with cardiovascular risks at baseline.[46], and another RCT in a larger sample (N=2118) reported no evidence for risk of cardiovascular and cerebrovascular events in participants who received desvenlafaxine.[48,66]

Additional trials outside the scope of our review

There were seven trials that were determined to be outside of the scope of our review due to differences in population, measurement, and length of trial. We have included these below and explained how these many influence our assessment.

Trials with Participants with Contraindications to MHT

Capriglione et al. (2016)[85] specifically recruited patients with a history of gynecological cancer to evaluate oral paroxetine 7.5 mg daily versus placebo for 16 weeks. Participants were gynecological

cancer survivors aged between 18 and 80 years of age. Exclusion criteria for this RCT included current metastatic cancer or other pre-existing chronic conditions (e.g., hypertension, impaired kidney function), history of psychiatric disorder, and use of VMS medications/supplements, beta-blockers, warfarin, or anti-epileptic medication.[85] Participants in the trial of those with a history of gynecological cancer were similar to the trials in healthy samples but were mostly White (99.5%). Capriglione et al. reported a greater reduction in weekly moderate-severe VMS frequency at week 16 in those receiving paroxetine (−46.5) compared to placebo (−39.3), p=0.009, and a greater reduction in weekly moderate-severe VMS severity at week 16 in the paroxetine group (−0.09), compared to placebo (−0.05), p=0.005; reductions in VMS frequency met MCID thresholds but VMS severity did not. At week 16, there were no differences in sleep outcomes (GCS sleep measure and HFRDIS) between the two arms.[85] Adverse events appeared to be similar to trials in healthy women. Nausea was more common in paroxetine arm than placebo and there were no serious adverse events.[85] See Table D.3.4-D.3.10.

Boekhout et al. (2011)[84] specifically recruited patients with a history of breast cancer to evaluate oral venlafaxine (75 mg daily) versus placebo for 12 weeks. Participants were breast cancer survivors, older than 18 years of age, and had natural or chemotherapy-induced menopause or were premenopausal with ovarian function suppression, with at least two VMS per day. Exclusion criteria for this RCT included history of chronic heart conditions, had recently started treatment for SSRIs, or had planned a switch in endocrine treatment during the study period. Participants in this trial were slightly younger (mean age of 49 years) compared to trials in healthy postmenopausal women. This trial reported only VMS severity. To measure VMS severity, the investigators asked participants to rate each VMS from 1-4 (mild-very severe) then summed the values for VMS severity. There was a median decrease in both venlafaxine (13.3 to 7.6) and placebo (14.4 to 10.9) and there was no significant difference in the change from baseline to week 12 between venlafaxine and placebo, p=0.07.[84] Adverse events appeared to be similar to trials in healthy women. Nausea was more common in the venlafaxine group than placebo, along with constipation. These adverse events were not associated with discontinuation. However, discontinuation was higher in the venlafaxine group (56%) compared to placebo (20%).[84] See Table D.3.4-D.3.10.

Trials of Shorter Duration (<8 weeks)

Stearns et al. (2003)[89] conducted a randomized control trial to evaluate controlled-release paroxetine 12.5 mg or 25 mg daily versus placebo for 6 weeks.[89] The inclusion and exclusion was similar to other SSRI trials, except this RCT that had more lenient criteria of at least 14 VMS per week. Consequentially, this trial had a lower baseline VMS frequency compared to the other SSRI trials at around 6.7 per day which was lower than the other SSRI trials that reported a baseline daily VMS frequency of 8-12. The investigators reported that there were significantly greater reductions in daily VMS frequency (mild, moderate, and severe) for participants who received paroxetine controlled-release at week 6, compared to those who received placebo (paroxetine 12.5 mg: -1.55;

95% CI: -2.76 to -0.34; p=0.01; and paroxetine 25 mg: -1.50; 95% CI: -2.66 to -0.34; p= 0.01). These differences were smaller than those reported in the Skylight trials and did not meet MCID threshold. There were also significantly greater reductions in VMS severity for paroxetine at week 6, compared to placebo (paroxetine 12.5 mg: −4.7; 95% CI: −8.1 to −1.3; p=0.007, and paroxetine 25 mg: −3.6; 95% CI:−6.8 to −0.4; p=0.03). These findings were supported by greater improvements in VMS, as measured by the Greene Climacteric Scale (GCS), at week 6 in participants who received paroxetine. There were no differences in change from baseline between the groups for sleep disturbance (measured using GCS), disability, functioning (measured using Sheehan Disability Scale), or depression (measured using Beck Depression Inventory-II).[89]

Grady et al. (2007)[86] conducted a randomized control trial to evaluate the efficacy of oral sertraline (50 mg daily for two weeks then increased to 100 mg daily) versus placebo for six weeks. The inclusion criteria were similar to the other SSRI trials and like Stearns et al. had a more lenient criterion of at least 14 VMS per week. Baseline characteristics were similar to the other SSRI trials. This trial reported no difference in mean percentage change in VMS frequency at week six between sertraline and placebo (sertraline: -39.0 [SE=44.8] and placebo: -38.3 [SE=32.8], p=0.94), nor VMS severity at week six (sertraline: -42.2 [SE=48.0] and placebo: -40.6 [SE=36.5, p=0.86). There was no difference in change in sleep quality (measured using PSQI) or positive/negative affect. There was a greater worsening of sexual function, measured using the Female Sexual Function Index, in the sertraline arm versus placebo (p=0.001), and of quality of life, measured using the Medical Outcomes Study Short Form 36 (p=0.05). Similar to the other SSRI trials, adverse events of any cause were mostly mild or moderate in severity, and the most commonly reported adverse event in this sertraline trial was dry mouth.[86] See Table D.3.4-D.3.10. The three excluded SSRI/SNRI trials described above provided additional information on the uncertainty of evidence for SSRIs/SNRIs, given the mixed results.

Butt et al. (2008)[126] evaluated oral gabapentin 900 mg daily versus placebo for four weeks. Participants were postmenopausal women with 14 VMS per week. Exclusion criteria were similar to the other gabapentin trials, except Butt et al. excluded those who were using SSRIs, SNRIs, or antiseizure medications.[126] Baseline characteristics were similar to the other gabapentin trials. Butt et al. reported that daily VMS frequency decreased 45.7% at week four for those in gabapentin which was significantly greater than the 24.7% reduction in the placebo, p<0.001. There was also a greater decrease in VMS severity in those who received gabapentin (51% reduction) compared to placebo (26.5% reduction), p<0.001. MENQoL total score improved more in participants receiving gabapentin (-0.8) than those receiving placebo (-0.4), p=0.004. However, the difference in MENQoL did not meet MCID criteria and was primarily driven by the change in vasomotor domain. Adverse events were similar to other gabapentin trials, with the most frequent adverse events in the gabapentin arms including dizziness and abdominal bloating. This trial reported that dizziness was the primary reason for early withdrawals in the gabapentin arm (8 out of 14).[126] See Table D.3.4-D.3.10.

We identified one Phase 3 trial that examined pregabalin for the treatment of VMS but it was not included due to the duration of the trial being less than 8 weeks.[90] Loprinzi et al. study evaluated oral pregabalin 75 mg or 150 mg twice daily versus placebo for 6 weeks. Participants in Loprinzi et al.[90] were women with bothersome hot flashes occurring at least 28 times per week. Participants in this trial were mostly 50 years of age or older (79%) and White (93%), and 40% had a history of breast cancer. The efficacy of pregabalin compared with placebo for the treatment of VMS associated with menopause was evaluated in one RCT.[90] Participants achieved a significant reduction in daily VMS frequency (mild, moderate, and severe) at week 6 in both the 75 mg and 150 mg dose, compared to placebo (p=0.003 and p=0.005, respectively). These daily improvements met MCID and are larger than reported in the Skylight 1 and 2 trials, though they measure also mild VMS. There was a significant reduction in weekly VMS severity score at week 6 in both the 75 mg and 150 mg as compared to placebo (p=0.002 and p=0.007, respectively). There were also improvements in mood (75 mg dose and 150mg), sleep, and quality of life (150 mg dose).[90] In this trial, adverse events were mild or moderate and discontinuation rates were similar across the arms. Adverse events were increased with the higher dose (150 mg).[90]

Trials Not Including Standard Measure of VMS

Kalay et al. (2007)[125] conducted an RCT in India that evaluated oral citalopram (10 mg daily for the first week and increased to 20 mg daily) versus placebo for eight weeks. Ten women received 40 mg daily because of insufficient improvement. The inclusion criteria and baseline characteristics were similar to the other SSRI trials, but no data was provided for race/ethnicity. This RCT examined VMS using the Kupperman Index of Climacteric Symptoms, which measures broader menopausal symptoms than VMS frequency and severity. Kalay et al. reported a larger reduction in climacteric symptoms, measured via Kupperman index, in the citalopram group (from 41.85 to 24.97) compared to the placebo group (from 40.06 to 36.65), p=0.001. MENQoL scores were provided by subdomain only. There was a greater reduction in vasomotor, psychosocial, and physical symptoms in the citalopram group compared to placebo (citalopram versus placebo for all three subdomains: p=0.001), but not for the sexual subdomain (no changes from placebo in both groups). Adverse events were mild or moderate for this citalopram trial and the most reported adverse events in the citalopram group were somnolence, increased perspiration, palpitation, and dry mouth in the citalopram trial. [125] While this trial did not measure VMS in a comparable way, the assessment MENQoL provides evidence for improvements for those who were prescribed SSRIs/SNRIs.

Table. D2.3. Supplement Trial Results

Trial Name/Author	Exclusion from main report	Intervention	Arm Size	VMS Frequency		VMS Severity	
				Change from baseline: Mean (SE)	DIFF from PBO: Mean (SE/95% CI), P Value	Change from baseline: Mean (SE)	DIFF from PBO: Mean (SE/95% CI), P Value
Boekhout et al. 2011[84]	Population with contraindications to MHT	Venlafaxine 75 mg	41	NR	NR	-41%‡‖	p=0.07
		Placebo	20	NR	NR	-29%‡‖	REF
Stearns et al. 2003[89]	Shorter duration (<8 weeks)	Paroxetine (12.5 mg CR)	51	-3.3* (NR)	-1.55* (95% CI: -2.76 to -0.34), p=0.01	−8.52§ (1.27)	−4.7 (95% CI:−8.1 to −1.3), p=0.007
		Paroxetine (25 mg CR)	58	-3.2* (NR)	-1.50* (95% CI: -2.66 to -0.34); p= 0.01	−7.43§ (1.18)	−3.6 (95% CI:−6.8 to −0.4), p=0.03
		Placebo	56	-1.8* (NR)	REF	−3.82§ (1.17)	REF
Capriglione et al. 2016[85]	Population with contraindications to MHT	Paroxetine (7.5 mg)	42	-46.5† (NR)	p=0.009	-0.09 (NR)	p=0.005
		Placebo	38	-39.3† (NR)	REF	-0.05 (NR)	REF
Grady et al. 2007[86]	Shorter duration (<8 weeks)	Sertraline 100 mg	50	-39%*‡ (44.8)	-0.7% (95% CI:– 15.9 to 17.2%), p=0.94	-42%*‡ (48.0)	-1.6% (95% CI: – 16.4 to 19.6%), p=0.86
		Placebo	49	-38%*‡ (32.8)	REF	-41%*‡ (36.5)	REF
Butt et al. 2008[126]	Shorter duration (<8 weeks)	Gabapentin 900 mg	99	-46%*†‡ (95% CI: 38.7, 52.7	p<0.001	-51.0%*‡ (95% CI: 43.3, 58.5)	p<0.001
		Placebo	98	-25%*†‡ (95% CI: 17.3, 32.1)	REF	-26.5%*‡ (95% CI: 18.3, 34.7)	REF
Loprinzi et al. 2010[90]	Shorter duration (<8 weeks)	Pregabalin 75mg	69	-4.6* (95% CI: -5.6 to -3.9)	P=0.003	-9.7* (95% CI: -12.1 to -7.3)	P=0.002
		Pregabalin 150mg	69	-4.9* (95% CI: -6.1 to -4.0)	P=0.005	-9.6* (95% CI: -12.9 to -7.6)	P=0.007
		Placebo	69	-2.9* (95% CI: -3.6 to -1.4)	REF	-6.1* (95% CI: -7.9 to -2.9)	REF

Note: Kalay et al. was not reported in the table as data were not comparable (Kupperman Index). CI: Confidence Interval, DIFF from PBO: Difference from placebo, NR: Not Reported, REF: Reference group, SD: Standard Deviation, SE: Standard Error, VMS: Vasomotor symptoms, mg: milligrams.

*All VMS (mild, moderate, and severe)
†Weekly score
‡Percentage change
§Met criteria for MCID.

D3. Evidence Tables

Table D3.1. Study Design

Title	Author	Intervention	Study Design	Inclusion	Exclusion	Trial Duration
Fezolinetant[54-57,99,108,115]						
Phase 2A[108]	Depypere et al.	Placebo (N=44) Fezolinetant 90 mg (N=43)	Phase 2a, double-blind, placebo-controlled study	Healthy women aged 40 to 65 years who had reached menopause and were experiencing moderate or severe VMS, at least 7 moderate to severe hot flashes or night sweats per day over a period of 7 consecutive days	Any medical condition that could confound results, had a recent history of a psychological disorder such as current major depression	12 weeks
VESTA[99,115]	Fraser et al., Santoro et al.	Fezolinetant 30 mg (N=43) Placebo (N=43)	Randomized, double-blind, placebo-controlled, dose-ranging, parallel group study	Healthy postmenopausal women aged 40-65 years, with ≥ 50 moderate to severe VMS episodes per week based on seven consecutive days of VMS recordings from any point during	Recent use of VMS therapy that could interfere with the occurrence of VMS (antidepressant use was permitted if the dose had not changed within the 3 months before screening)	12 weeks

				the screening period		
SKYLIGHT 1 [54,55]	Lederman et al., Nappi et al	Fezolinetant 30mg (n=173) Fezolinetant 45 mg (n=174) Placebo (n=175)	NR	Women aged 40–65 years with moderate-to-severe VMS at least 7 hot flashes per day	NR	12 weeks
SKYLIGHT 2 [55-57]	Johnson et al., Nappi et al	Fezolinetant 30 mg (n=166) Fezolinetant 45 mg (n=167) Placebo (n=167)	Double-blind, placebo-controlled, multicenter phase 3 study	Women aged ≥ 40-65 years with moderate-to-severe VMS, at least 7 hot flashes per day, associated with menopause	NR	12 weeks (extension: 52 weeks)

MHT: standard dose (1 mg estradiol)[33,34,36,37,39,58-62]						
Estradiol and drospirenone for climacteric symptoms in postmenopausal women: a double-blind, randomized, placebo-controlled study of the safety and efficacy of three dose regimens.[36]	Schürmann et al.	Placebo (n= 61) Estradiol 1 mg/DRSP 1 mg (n=55) Estradiol 1 mg/DRSP 2 mg (n=52) Estradiol 1 mg/DRSP 3 mg (n=57)	Double-blind, randomized, placebo-controlled, multicenter study with four treatment groups.	Healthy post-menopausal Caucasian women aged 45–65 years, who complained of at least 5 moderate to severe hot flushes per day during the screening period	Contraindications for MHT, treatment with anticoagulant medications, recent use of oral, transdermal, or transvaginal hormonal preparations. Past medical history for cardiovascular disease, depression, diabetes, hypertension, or other diseases that could affect the study results.	4, 8 and 16 weeks
A multicenter, prospective, randomized, double-blind, placebo-controlled study to investigate the efficacy of a continuous-combined hormone therapy preparation containing 1mg estradiol valerate/2mg dienogest on hot flushes in postmenopausal women.[34]	Endrikat et al.	1 mg estradiol valerate/2 mg dienogest (n=162) Placebo (n=162)	Multicenter, prospective, randomized, double-blind, placebo-controlled study	Women aged 52–65 years in general or aged 40–51 years in case of previous bilateral oophorectomy, and had an intact uterus	Contraindications to HRT; any disease/conditions that compromised the function of the body systems; abnormal cervical smear; abnormal baseline lab values considered clinically significant; history of alcohol or drug abuse; current significant liver dysfunction; insulin-dependent	Week 12

					diabetes; hypertension; concomitant medication with drugs known to influence the study medication; any severe systemic disease that could interfere with the study.	
Estradiol 1 mg and drospirenone 2 mg as hormone replacement therapy in postmenopausal Chinese women.[37]	Lin et al.	Estradiol 1 mg/DRSP 2 mg (n=183) Placebo (n=61)	Double-blind, randomized, placebo-controlled, phase III study	Women who had 24 or more moderate to severe hot flushes over 7 consecutive days during the screening period, had a negative pregnancy test and negative bilateral mammography results	History of cardiovascular disease, uncontrolled thyroid disorders, clinical depression, malignant or premalignant disease, abnormal gynecologic findings, hepatic disease, adrenal insufficiency or renal failure, abnormal glucose tolerance and severe or congenital hyper-triglyceridemic; abnormal baseline laboratory findings; a history of alcohol/drug abuse or current smoking; recent hormonal therapy;	12 and 16 weeks

					use of herbal/other medicines for climacteric disorders.	
REPLENISH: Metabolic and cardiovascular effects of TX-001HR in menopausal women with vasomotor symptoms[33,39,58-61]	Lobo et al., Archer et al., Kagan et al., Simon et al., Kaunitz et al. 2020, Black et al. 2020, Mirkin et al.	Estradiol 1 mg and progesterone 100 mg (n=415) Estradiol 0.5 mg and progesterone 100 mg (n=424) Estradiol 0.5 mg and progesterone 50 mg (n=421) Placebo (n=151)	Phase 3, randomized, double-blind, placebo-controlled, multi-center trial	Healthy menopausal women aged 40–65 years with BMI 34 kg/m^2 or less, had an intact uterus and at least 12 months of spontaneous amenorrhea	Contraindications or allergy to MHT; a history of endometrial hyperplasia or undiagnosed vaginal bleeding; uterine fibroids diagnosed at screening; heavy smoking, or a history of drug or alcohol abuse; recent use of another therapy for VMS	VMS sub-study: 4 and 12 weeks. Overall trial: 12 months
Oral 17β-estradiol/progesterone (TX-001HR) and quality of life in postmenopausal women with vasomotor symptoms[62] Substudy of Lobo et al. (2018)	Simon et al.	Estradiol 1 mg and progesterone 100 mg (n=141) Estradiol 0.5 mg and progesterone 100 mg (n=149) Estradiol 0.5 mg and progesterone 50 mg (n=147) Placebo (n=135)	Phase 3, prospective, randomized, double-blind, placebo-controlled, multicenter trial	Women who had a minimum of 7 moderate to severe VMS daily or 50 per week before enrollment	NR	NR

MHT: low dose (0.5 mg estradiol or equivalent)[35,38-41,63]						
Ultra-low-dose estradiol and norethisterone acetate: effective menopausal symptom relief.[40]	Panay et al.	Estradiol 0.5 mg and NETA 1 mg (n=194) Estradiol 0.5 mg and NETA 0.25 mg (n=181) Placebo (n=200)	Randomized, placebo-controlled, double-blind, multicenter, multi-national, parallel-group evaluation	Women who had at least 50 moderate to severe hot flushes per week, no menses during the past year or 6 months spontaneous amenorrhea with FSH levels 440 mIU/ml and estradiol levels 525 pg/ml.	Recent exposure to MHT. Suspected or previous history of breast cancer or estrogen-dependent neoplasia, untreated endometrial hyperplasia and abnormal genital bleeding. History of diabetes mellitus, hypertension, any thrombo-embolic conditions and hepatic or renal impairment.	4, 8, 12 and 24 weeks
Oral ultra-low dose continuous combined hormone replacement therapy with 0.5 mg 17β-oestradiol and 2.5 mg dydrogesterone for the treatment of vasomotor symptoms: results from a double-blind, controlled study.[35,63]	Stevenson et al., Tsiligiannis et al.	Estrogen and progestin/progesterone (n=124) Placebo (n=127)	Double-blind, randomized, placebo-controlled, parallel group study	Non-hysterectomized, postmenopausal women aged 45–65 years who had been amenorrhoeic for ≥ 12 months, had serum estradiol and FSH levels within the post-menopausal range, had ≥ 50 moderate to severe hot flushes during	Endometrial biopsy showing clinically relevant abnormalities and/or bilayer endometrial thickness of ≥5 mm, recent abnormal vaginal bleeding, a history of or current estrogen	13 weeks, and a follow-up treatment period of 39 weeks.

				the screening period		
A randomized, double-blind, placebo-controlled study of the lowest effective dose of drospirenone with 17β-estradiol for moderate to severe vasomotor symptoms in postmenopausal women. [39]	Archer et al.	DRSP 0.25 mg/E2 0.5 mg (n=177) DRSP 0.5 mg/E2 0.5 mg (n- 178) Placebo (n=176)	Phase 3, double-blind, randomized, parallel-group, placebo-controlled study	Women aged 40 years or older, experienced spontaneous amenorrhea for 12 months or more, had a minimum of 7 to 8 moderate to severe HF per day, or 50 to 60 moderate to severe HF per week during the screening period	Recent use of oral hormonal products.	12-weeks
MsFLASH 03 [38]	Joffe et al.	Estradiol 0.5 mg (n=97) Venlafaxine 75 mg (n=96) Placebo (n=146)	Three arm, double-blind, randomized trial	Healthy women aged 40 to 62 years in the menopause transition, were postmenopausal, had FSH level exceeding 20 mIU/mL, and an estradiol level not exceeding 50 pg/mL in the	Pregnancy or breastfeeding; suicide attempt in the past 3 years; diagnosis of bipolar disorder or psychosis; psychotropic medications for VMS in the past month; major depressive episode	4 and 8 weeks

				absence of a reliable menstrual marker and were required to have at least 14 VMS per week	or drug or alcohol abuse in the past year, recent or current use of MHT; hormonal contraceptives, SERM or aromatase inhibitors, and some comorbidities	
Efficacy and local tolerance of a low-dose, 7-day matrix estradiol transdermal system in the treatment of menopausal vasomotor symptoms.[41]	Speroff et al.	Study 1: Estradiol transdermal system: 0.02 mg (n=54) Estradiol transdermal system: 0.04 mg (n=53) Placebo single dose (n=54) Placebo double dose (n =52) Study 2: Estradiol transdermal system: 0.02 mg (n=37) Estradiol transdermal system: 0.04 mg (n=37) Placebo double dose (n=37)	Double-blind, placebo-controlled trial	Women at least 50 years of age, undergone hysterectomy, had natural menopause or at least 35 years of age, had surgical menopause, and screened for baseline VMS (at least 56 per week)	Contraindications to MHT; those with a skin condition that may be exacerbated by use of transdermal system	12 Weeks

SNRIs[44,46-48,50,64,65,84]						
Management of postmenopausal hot flushes with venlafaxine hydrochloride: a randomized, controlled trial[44]	Evans et al.	Venlafaxine 75 mg (n=40) placebo (n=40)	Randomized controlled trial	Women with natural or surgical menopause and had more than 14 hot flushes per week	Receiving estrogens, progestins, androgens, antidepressants, or chemotherapy.	4, 8, and 12 weeks
Efficacy and tolerability of desvenlafaxine succinate treatment for menopausal vasomotor symptoms: a randomized controlled trial.[46,50]	Speroff et al., Wyrwich et al.	Desvenlafaxine 100 mg (n=157) Desvenlafaxine 150 mg (n=163) Desvenlafaxine 200 mg (n=155) Placebo (n=78)	Multi-center, randomized, double-blind, placebo-controlled trial	Healthy postmenopausal women with BMI 40 kg/m^2 or less who experienced at least 7 moderate-to-severe hot flushes per day (or 50 or more per week)	Recent use of MHT or therapies for VMS; history of seizure disorder; myocardial infarction; malignancy other than basal or squamous cell carcinoma; glaucoma or raised intraocular pressure; or hepatic, renal medical disease; current major depressive, bipolar, psychotic, or generalized anxiety disorder; other clinically important abnormalities at screening.	12 weeks

(12-week) A double-blind, randomly assigned, placebo-controlled study of desvenlafaxine efficacy and safety for the treatment of vasomotor symptoms associated with menopause[64]	Archer et al.	Desvenlafaxine 150 mg (n= 151) Desvenlafaxine 100 mg (n=150) Placebo (n=150)	Double-blind, randomly assigned, placebo-controlled study	Healthy postmenopausal women with BMI 40 kg/m^2 or less, and experienced at least 7 moderate to severe hot flushes per day or 50or more per week for 2 consecutive weeks at baseline	Recent use of any hormone-containing drug or VMS therapy; history of seizure disorder, myocardial infarction, or malignancy or treatment for malignancy other than basal or squamous cell carcinoma; hepatic, renal, or other medical disease; presence of psychiatric disease requiring therapy	4 and 12 weeks
(26-weeks) Desvenlafaxine for the treatment of vasomotor symptoms associated with menopause: a double-blind, randomized, placebo-controlled trial of efficacy and safety[65]	Archer et al.	Desvenlafaxine 100 mg (n=182) Desvenlafaxine 150 mg (n=179) Placebo (n=180)	Multicenter, randomized, double-blind, placebo-controlled trial	Healthy, postmenopausal women with BMI ≤40 kg/m^2 who experienced at least 7 moderate to severe hot flashes per day (or 50/ week) recorded by participants for 7 consecutive days during screening	Recent use of any hormone-containing drug or VMS therapy; history of seizure disorder, myocardial infarction, or malignancy or treatment for malignancy other than basal or squamous cell carcinoma; hepatic, renal, or other medical disease; presence of psychiatric disease requiring therapy	26 weeks

Management of hot flashes in patients who have breast cancer with venlafaxine and clonidine: A randomized, double-blind, placebo-controlled trial[84]	Boekhout et al.	Venlafaxine 75 mg (n=41) Placebo (n=20)	Randomized, Double-Blind, Placebo-Controlled Trial	Postmenopausal women aged 18 years and above with a history of breast cancer, and experienced at least two hot flashes per day	Treatment with antidepressants or SSRIs less than 4 weeks before assignment; recently used drugs that might affect study drug metabolism; had a history of uncontrolled hypertension, heart disease or angina pectoris, recent myocardial infarction, planned switch in endocrine treatment during the study period; were pregnant; or were breastfeeding	12 weeks
Randomized placebo- and active-controlled study of desvenlafaxine for menopausal vasomotor symptoms[47]	Bouchard et al.	Placebo (n=150) Desvenlafaxine 100 mg (n=137)	Phase-3, multicenter, randomized, double-blind, placebo-controlled, and active-comparator controlled trial	Healthy women who had completed their last natural menstrual period ≥12 months prior to screening, had an intact uterus, a BMI of ≤34.0 kg/m^2, and a minimum of 7 moderate and severe VMS per day, or ≥50 moderate and	Recent use of any hormone-containing drug or VMS therapy; estrogen-dependent neoplasia; seizure disorder; active or recent arterial thrombo-embolic disease; cerebrovascular accident or stroke; venous thromboembolism; malignancy or treatment for	12 weeks

				severe VMS per week recorded for 7 consecutive days during screening	malignancy within 2 years; hepatic, renal medical disease; major depressive, bipolar, psychotic, or generalized anxiety disorder requiring therapy; narrow-angle glaucoma or current raised intraocular pressure; undiagnosed vaginal bleeding.	
Maintenance of the efficacy of desvenlafaxine in menopausal vasomotor symptoms: a 1-year randomized controlled trial.[48,66]	Pinkerton et al, Archer et al.	Desvenlafaxine 100 mg (n=1066) Placebo (n=1052)	Phase 3, multicenter, randomized, double-blind, placebo-controlled, parallel-group trial	Women aged 45 years or older, had a BMI of 34.0 kg/m^2 or lower, and had confirmed menopause status. Efficacy sub-study: Approximately 20% of the enrolled participants met the additional criterion of 7 or more moderate and severe HFs per day (or ≥50 HFs per week) recorded for 2	Recent use of any VMS therapy; history of seizure disorder, myocardial infarction, narrow-angle glaucoma, or malignancy or treatment of malignancy other than basal or squamous cell carcinoma; important medical disease; major depressive, bipolar, psychotic, or generalized anxiety disorder requiring therapy; other clinically important abnormalities at screening.	12 months

				weeks before randomization		
SSRIs[68-70,85,86,89,125]						
Paroxetine Controlled Release in the Treatment of Menopausal Hot Flashes[89]	Stearns et al.	Paroxetine 12.5 mg CR (n=51) Paroxetine 25 mg CR (n= 58) Placebo (n=56)	Randomized, double-blind, placebo-controlled, parallel group study conducted across 17 US sites	Menopausal women aged 18 years or older who had experienced a minimum of 2-3 daily hot flashes or at least 14 bothersome hot flashes per week, discontinued any HRT at least 6 weeks before screening	Presented with signs of active cancer or were receiving current chemotherapy or radiation therapy	6 weeks

Efficacy of citalopram on climacteric symptoms.[125]	Kalay et al.	Citalopram 20 mg (n=25) Placebo (n=25)	Single-blind, randomized control trial	Women with natural or surgical menopause, had more than 7 to 8 hot flashes per day, and had a normal thyroid function	Psychotic disease and/or who were undergoing psychiatric therapy and those taking herbal products, dopaminergic or antidopaminergic drugs, or narcotic analgesics.	
FAST trial (Grady et al., 2007) [86]	Grady et al.	Sertraline 100 mg (n=50) Placebo (n=49)	Double blind, randomized, placebo-controlled trial	Healthy perimenopausal or postmenopausal women aged 40 to 60 years who reported experiencing at least 14 hot flushes per week	Women with a history of breast or ovarian cancer, depression, chronic kidney or liver disease, bipolar affective disorder, seizures, and known hypersensitivity to SSRIs, use of VMS therapies.	6 weeks
MSFLASH 01: Efficacy of escitalopram for hot flashes in healthy menopausal women: a randomized controlled trial.[43,67]	Freeman et al., Diem et al.	Escitalopram 10 mg (n=104) Placebo (n=101)	Multicenter, 8-week, randomized, double-blind, placebo-controlled, parallel group trial	Women who had at least 28 hot flashes or night sweats per week for 3 weeks where hot flashes or night sweats rated as bothersome or severe on 4 or more days per week and the frequency in week 3 did not decrease by more than 50%	Use of therapies for hot flashes in the past 30 days; current severe medical illness, major depressive episode, drug or alcohol abuse in the past year, suicide attempt in the past 3 years, lifetime diagnosis of bipolar disorder, or psychosis; or uncontrolled hypertension,	8 weeks

				from the mean weekly levels in weeks 1 and 2	history of endometrial or ovarian cancer, myocardial infarction, angina or cerebrovascular events, or other preexisting medical conditions	
(Study 1 & 2) Low-dose paroxetine 7.5 mg for menopausal vasomotor symptoms: Two randomized controlled trials[68-70]	Simon et al., Pinkerton et al., Portman et al.	Paroxetine 7.5 mg (n= 301) Placebo (n=305)	Multicenter, double-blind, randomized, placebo-controlled, phase 3 studies	Postmenopausal women 40 years or older who had an average of more than 7-8 moderate to severe hot flashes per day, or 50-60 moderate to severe hot flashes per week	Hypersensitivity to paroxetine; known nonresponse to previous SSRI or SNRI treatment of VMS; untreated hypertension; impaired liver or kidney function; unstable cardiac disease; pregnancy; a history of psychiatric disorder; and any other medical condition.	12 weeks
Role of paroxetine in the management of hot flashes in gynecological cancer survivors: Results of the first randomized single-center controlled trial[85]	Capriglione et al.	Paroxetine 7.5 mg (n=42) Placebo (n=38)	Randomized, single-center, double-blind, placebo-controlled study	Women aged between 18-80 years, completed all active cancer treatment including surgery, radiation, chemotherapy, and hormonal therapy at least	Metastatic cancer; history of prior malignancies; hypersensitivity to paroxetine; known nonresponse to previous SSRI or SNRI treatment of VMS; untreated hypertension; impaired liver or kidney function;	16 weeks

				60 days prior to enrollment	unstable cardiac disease; pregnancy; a history of self-injurious behavior; a history of clinical diagnosis or treatment of any psychiatric disorder, use of anti-epileptic medication or concomitant use of beta-blockers; warfarin; menopausal hormone therapy; black cohosh/flaxseed/soy supplements; or regular nightly use of sleep aids.	
Gabapentin[52,53,71,72,126]						
Gabapentin's effects on hot flashes in postmenopausal women: A randomized controlled trial[53,71]	Guttuso et al., Yurcheshen et al.	Gabapentin 900 mg (n=30) Placebo (n=29)	Randomized, double-blind, placebo-controlled trial	Postmenopausal women with an average of 7 or more hot flashes per day accompanied by sweating	Estrogen, progestin, leuprolide, or tamoxifen therapy within the past 2 months. More than 50% of a patient's hot flashes were associated with occurrence of migraine headaches or ingestion of particular foods or beverages.	12 weeks

Gabapentin, estrogen, and placebo for treating hot flushes: A randomized controlled trial[52]	Reddy et al.	Estrogen only (n= 20) Gabapentin 2400 mg (n=20) Placebo (n= 20)	Randomized, double-blind, placebo-controlled trial	Menopausal women, aged 35 - 60 years, experiencing at least 50 moderate to severe hot flashes per week >2 months	History of DVT, history of MI, stroke, and/or functional decline, history of malignancies or undiagnosed vaginal bleeding, history of chronic liver, gallbladder, chronic renal, cardiac, or endocrine diseases	12 weeks
Gabapentin for the treatment of menopausal hot flashes: A randomized controlled trial[126]	Butt et al.	Gabapentin 900 mg (n= 99) Placebo (n= 98)	Double-blind, placebo-controlled, randomized trial	Postmenopausal women, defined as those who had experienced natural cessation of menses for 1 year, who were between the ages of 45 and 65 years and who experienced at least 14 hot flashes per week	Use of HRTs, tamoxifen, raloxifene, SSRIs, SNRIs, or antiseizure medications; present or planned antineoplastic or radiation therapy; bilateral oophorectomy; neurologic conditions; hypothalamic dysfunction.	4 weeks
Breeze 3[72]	Pinkerton et al.	Gabapentin gastroretentive 1800 mg (n=299) Placebo (n=294)	Prospective, randomized, double-blind, placebo-controlled, multicenter trial	Healthy postmenopausal women who experienced 7 or more moderate-to-severe hot flashes per day during a 14-day baseline	Current treatment with MHT; history of gastric reduction; substance abuse within the past year; or any serious medical condition. Concomitant treatment of hot flashes except antidepressant with	12 weeks

					unchanged dosages were permitted.	

BMI: body mass index, DSRP: drospirenone, DVT: deep vein thrombosis, E2: estradiol, FSH: follicle stimulating hormone, HF: hot flashes, HRT: hormone replacement therapy, MHT: menopausal hormonal therapy, mg: milligram, MI: myocardial infarction, N: number, NETA: norethindrone acetate, NR: not reported, SERM: selective estrogen receptor modulators, SNRI: serotonin and norepinephrine reuptake inhibitors, SSRI: selective serotonin reuptake inhibitors; US: United States, VMS: vasomotor symptoms.

*Reddy et al. included a trial arm of 0.625 mg conjugated estrogens (N=20). This was excluded from this review as women were randomly allocated to the doses and no information was provided on the number of women without a uterus.

Table D3.2. Baseline Characteristics I

Title	Intervention	Arm Size	Age (Mean, SD)	BMI (Mean, SD)	Natural menopause % (n/N)	GCS (Mean, SD)	HFRDIS (Mean, SD)	MENQoL (Mean, SD)
Fezolinetant[56,57,99,108,115]								
Phase 2A	Placebo	44	53.7, 4.25	26.5, 6.15	NR	NR	NR	NR
	Fezolinetant 90 mg	43	53.3, 4.03	25.1, 4.71	NR	NR	NR	NR
VESTA	Fezolinetant 30mg	43	52.7, 3.8	28.8, 4	27 / 43 (62.8%)	20.5, 9.1	6, 2.1	4.4, 1.5
	Placebo	43	54.8, 5.5	27.3, 4.8	25 / 43 (58.1%)	21.7, 10.3	6, 2.3	4.3, 1.6
SKYLIGHT 1	Fezolinetant 30mg	173	NR	NR	NR	NR	NR	NR
	Fezolinetant 45mg	174	NR	NR	NR	NR	NR	NR
	Placebo	175	NR	NR	NR	NR	NR	NR
SKYLIGHT 2	Fezolinetant 30mg	166	53.9, 4.9	27.94	NR	NR	NR	NR
	Fezolinetant 45mg	167	54.3, 5.4	27.91	NR	NR	NR	NR
	Placebo/fezolinetant 30mg	76	54.3, 4.2		NR	NR	NR	NR
	Placebo/fezolinetant 45mg	75	55.4, 4.9		NR	NR	NR	NR
MHT: standard dose (1 mg estradiol) [33,34,36,37,39,62]								
Schurmann et al. 2004	Estradiol 1 mg /DRSP 1 mg	55	54.3, 5	26, 4.2	NR	NR	NR	NR

	Estradiol 1 mg/DRSP 2 mg	52	53.1, 4.4	26.2, 3.5	NR	NR	NR	NR
	Estradiol 1 mg/DRSP 3 mg	57	53.1, 4.9	26.2, 4	NR	NR	NR	NR
	Placebo	61	54, 4.7	26.5, 4.8	NR	NR	NR	NR
Endrikat et al. 2007	1 mg estradiol valerate/2 mg dienogest	162	56.3, 4.9	26.6, 3.7	108 / 162 (66.7%)	NR	NR	NR
	Placebo	162	56.2, 4.8	26.5, 3.9	109 / 162 (67.3%)	NR	NR	NR
Lin et al. 2011	Estradiol 1 mg/DRSP 2 mg	183	52, 3.81	23.8, 2.78	NR	NR	NR	NR
	Placebo	61	51.9, 3.56	22.4, 2.8	NR	NR	NR	NR
Lobo et al. 2018	Estradiol 1 mg and progesterone 100 mg	415	54.7, 4.4	26.8, 4.1	NR	NR	NR	NR
	Estradiol 0.5 mg and progesterone 100 mg	424	54.5, 4.5	26.7, 4.3	NR	NR	NR	NR
	Estradiol 0.5 mg and progesterone 50 mg	421	54.9, 4.3	26.7, 4	NR	NR	NR	NR
	Placebo	151	54.5, 4.3	26.6, 3.9	NR	NR	NR	NR
Simon et al. 2019 (substudy of Lobo et al. 2018)	Estradiol 1 mg and progesterone 100 mg	141	54.7, 4.8	26.5, 3.9	NR	NR	NR	4.5, 1.2

	Estradiol 0.5 mg and progesterone 100 mg	149	54.9, 4.5	27.1, 4.3	NR	NR	NR	4.3, 1.3
	Estradiol 0.5 mg and progesterone 50 mg	147	54.8, 4.6	26.6, 3.9	NR	NR	NR	4.7, 1.4
	Placebo	135	54.3, 4.3	26.6, 3.8	NR	NR	NR	4.6, 1.3
MHT: low dose (0.5 mg estradiol or equivalent) [35,38-41]								
Panay et al. 2007	Estradiol 0.5 mg and NETA 1 mg	194	55.2, 4.8	25, 3.6	NR	17.96	NR	NR
	Estradiol 0.5 mg and NETA 0.25 mg	181	55.3, 4.4	25.4, 3.5	NR	17.8	NR	NR
	Placebo	200	56.1, 4.7	25.3, 3.6	NR	17.64	NR	NR
Stevenson et al. 2010	Estradiol 1 mg/DYD 5 mg	62	54, 5	25.98, 3.46	NR	NR	NR	NR
	Estradiol 0.5 mg/DYD 2.5 mg	124	53.5, 4.6	26.51, 11.3	NR	NR	NR	NR
	Placebo	127	53.8, 4	26.58, 4.49	NR	NR	NR	NR
Archer et al. 2013	Estradiol 0.5 mg/DRSP 0.25 mg	177	53.5, 5.77	28.2, 5.7	NR	NR	NR	NR
	Estradiol 0.5 mg /DRSP 0.5 mg	178	53.8, 5.61	29.1, 6.1	NR	NR	NR	NR

	Placebo	176	53.4, 6.46	27.8, 5.79	NR	NR	NR	NR
MsFLASH 03 (Joffe et al., 2014)	Estrogen only	97	54.9, 4.1	28.5, 6.5	NR	NR	NR	NR
	Venlafaxine 75 mg	96	54.8, 3.7	29.3, 6.9	NR	NR	NR	NR
	Placebo	146	54.3, 3.8	27.6, 6.8	NR	NR	NR	NR
Speroff et al. 1996 (study 1 and study 2)	Estradiol transdermal system: 0.02 mg	54	NR	NR	NR	NR	NR	NR
	Estradiol transdermal system: 0.04 mg	53	NR	NR	NR	NR	NR	NR
	Placebo single dose	54	NR	NR	NR	NR	NR	NR
	Placebo double dose	52	NR	NR	NR	NR	NR	NR
	Estradiol transdermal system: 0.02 mg	37	NR	NR	NR	NR	NR	NR
	Estradiol transdermal system: 0.04 mg	37	NR	NR	NR	NR	NR	NR
	Placebo double dose	37	NR	NR	NR	NR	NR	NR

SNRIs[44,46-48,50,64-66,84]								
Evans et al. 2005	Venlafaxine 75 mg	40	52.7, 4.9	NR	30 / 37 (81.1%)	NR	NR	NR
	Placebo	40	51.6, 6.1	NR	31 / 40 (77.5%)	NR	NR	NR
Speroff et al. 2008	Desvenlafaxine 100 mg	157	53.5, 5.33	27.1, 4.64	115 / 145 (79.3%)	NR	NR	NR
	Desvenlafaxine 150 mg	163	53.3, 4.59	26.6, 4.47	106 / 137 (77.4%)	NR	NR	NR
	Desvenlafaxine 200 mg	155	53.5, 4.51	27.3, 4.57	94 / 120 (78.3%)	NR	NR	NR
	Placebo	78	54.2, 5.44	26.7, 4.72	59 / 77 (76.6%)	NR	NR	NR
Wyrwich et al. 2008 (Secondary analysis of Speroff et al. 2008)	Desvenlafaxine 100 mg	145	53.48, 5.33	27.06, 4.64	115 / 145 (79.3%)	NR	NR	NR
	Desvenlafaxine 150 mg	137	53.29, 4.59	26.63, 4.47	106 / 137 (77.4%)	NR	NR	NR
	Desvenlafaxine 200 mg	120	53.51, 4.51	27.33, 4.57	94 / 120 (78.3%)	NR	NR	NR
	Placebo	77	54.22, 5.44	26.72, 4.72	59 / 77 (76.6%)	NR	NR	NR
(12-week) Archer et al. 2009	Desvenlafaxine 150 mg	151	53.43, 4.64	NR	119 / 151 (78.8%)	NR	NR	NR

	Desvenlafaxine 100 mg	150	53.29, 4.7	NR	120 / 150 (80.0%)	NR	NR	NR
	Placebo	150	53.36, 5.05	NR	122 / 150 (81.3%)	NR	NR	NR
(26-weeks) Archer et al. 2009	Desvenlafaxine 100 mg	182	53.3, 5.2	27.5, 4.78	131 / 182 (72.0%)	NR	NR	NR
	Desvenlafaxine 150 mg	179	53.9, 5.14	27.5, 4.72	140 / 179 (78.2%)	NR	NR	NR
	Placebo	180	54, 4.74	26.3, 4.27	141 / 180 (78.3%)	NR	NR	NR
Boekhout et al. 2011	Venlafaxine 75 mg	41	median [IQR]: 48, [28, 69]	NR	13 / 41 (31.7%)	NR	NR	NR
	Placebo	20	median [IQR]: 50, [34, 62]	NR	3 / 20 (15.0%)	NR	NR	NR
Bouchard et al. 2012	Desvenlafaxine 100 mg	137	54, 4	26, 4	NR	NR	NR	NR
	Placebo	150	54, 5	26, 4	NR	NR	NR	NR
Pinkerton et al. 2013; Archer et al. 2013	Desvenlafaxine 100 mg	1066	54, 4.9	26.6, 4	872 / 1066 (81.8%)	NR	NR	NR
	Placebo	1052	53.6, 4.9	26.8, 4	829 / 1052 (78.8%)	NR	NR	NR

SSRIs[43,68,69,85,86,89,125]								
Stearns et al. 2003	Paroxetine 12.5 mg CR	51	53.6	NR	NR	NR	NR	NR
	Paroxetine 25 mg CR	58	55	NR	NR	NR	NR	NR
	Placebo	56	53.6	NR	NR	NR	NR	NR
Kalay et al. 2007	Citalopram 20 mg	25	53.5, 5.3	26.3, 4.2	19 / 25 (76.0%)	NR	NR	NR
	Placebo	25	51.7, 4.6	29.6, 4.1	19 / 25 (76.0%)	NR	NR	NR
FAST Trial (Grady et al., 2007)	Sertraline 100 mg	50	50.5, 5	NR	NR	NR	NR	NR
	Placebo	49	52.6, 4.2	NR	NR	NR	NR	NR
MsFLASH 01 (Freeman et al., 2011)	Escitalopram 10 mg	104	53.45, 4.2	28.58, 6.59	NR	NR	NR	NR
	Placebo	101	54.36, 3.86	29.7, 6.42	NR	NR	NR	NR
Simon et al. 2011 (Study 1)	Paroxetine 7.5 mg	301	median [IQR]: 54, [40, 73]	28.3	242 / 301 (80.4%)	NR	NR	NR
	Placebo	305	Median [IQR]: 53, [40, 79]	29	253 / 305 (83.0%)	NR	NR	NR

Simon et al. 2011 (Study 2)	Paroxetine 7.5 mg	284	Median [IQR]: 54, [41, 70]	27.4	227 / 284 (79.9%)	NR	NR	NR
	Placebo	284	54, [40, 74]	27.7	230 / 284 (81.0%)	NR	NR	NR
Simon et al. 2011 (Study 1 & 2) Sleep outcomes (Pinkerton 2015)	Paroxetine 7.5 mg	585	54.6, 5.73	28.62, 5.73	469 / 585 (80.2%)	56.1	7.6, 2.69	NR
	Placebo	589	54.5, 6.01	29.03, 5.51	483 / 589 (82.0%)	57.47	7.66, 2.47	NR
Capriglione et al. 2016	Paroxetine 7.5 mg	42	53.5, 5.71	26.7, 4.62	NR	55.6	7.4, 2.34	NR
	Placebo	38	53.6, 6.01	27.5, 4.71	NR	56.37	7.54, 2.16	NR
Gabapentin[52,53,72,126]								
Guttuso et al. 2003	Gabapentin 900 mg	30	52.7, 3.6	NR	NR	NR	NR	NR
	Placebo	29	53, 3.1	NR	NR	NR	NR	NR
Reddy et al. 2006	Gabapentin 2400 mg	20	51.25, 4.49	27.43, 5.01	17 / 20 (85.0%)	NR	NR	NR
	Placebo	20	52.25, 4.23	26.07, 6.73	19 / 20 (95.0%)	NR	NR	NR
Butt et al. 2008	Gabapentin 900 mg	99	55.9, 4.7	26, 4.5	NR	NR	NR	3.5, 1.3

	Placebo	98	56.5, 4.4	25.4, 4.5	NR	NR	NR	3.5, 1.3
Breeze 3 (Pinkerton et al., 2014)	Gabapentin 1800* mg	299	54, 6.1	NR	223 / 299 (74.6%)	NR	NR	NR
	Placebo	294	54, 6	NR	220 / 294 (74.8%)	NR	NR	NR

BMI: body mass index, DSRP: drospirenone, E2: estradiol, FEZ: fezolinetant, GCS: Greene climacteric scale, HFRDIS: hot flash related daily interference scale, IQR: interquartile range, MENQoL: the menopause-specific quality of life questionnaire, MHT: menopausal hormonal therapy, mg: milligrams, CR: controlled release, N: total number of participants, NETA: norethindrone acetate, NR: not reported, PBO: placebo, SD: standard deviation, SNRI: serotonin and norepinephrine reuptake inhibitors, SSRI: selective serotonin reuptake inhibitors

*gastroretentive gabapentin

Table D3.3. Baseline Characteristics II: Race/Ethnicity

Title	Intervention	Arm Size	American Indian/Alaska Native % (n/N)	Asian % (n/N)	Black/AA % (n/N)	White % (n/N)	Hispanic % (n/N)
Fezolinetant[54,56,57,99,108,115]							
Phase 2A	Placebo	44	NR	NR	NR	100.00%	NR
	Fezolinetant 90 mg	43	NR	NR	NR	97.70%	NR
VESTA	Fezolinetant 30 mg	43	NR	0.00%	25.60%	72.10%	39.50%
	Placebo	43	NR	4.70%	23.30%	69.80%	34.90%
SKYLIGHT 1	Fezolinetant 30 mg	173	NR	NR	NR	NR	NR
	Fezolinetant 45 mg	174	NR	NR	NR	NR	NR
	Placebo	175	NR	NR	NR	NR	NR
SKYLIGHT 2	Fezolinetant 30 mg	166	0.00%		21.10%	78.90%	20.50%
	Fezolinetant 45 mg	167	1.20%		19.80%	79.00%	24.60%
	Placebo/fezolinetant 30mg	76	2.6%		14.5%	82.9%	18.4%

	Placebo/fezolinetant 45mg	75	0%		24.0%	76.0%	21.6%
MHT: standard dose (1 mg estradiol)[33,34,36,37,39,62]							
Schurmann et al. 2004	Estradiol 1 mg/DSRP 1 mg	55	NR	NR	NR	NR	NR
	Estradiol 1 mg/DSRP 2 mg	52	NR	NR	NR	NR	NR
	Estradiol 1 mg/DSRP 3 mg	57	NR	NR	NR	NR	NR
	Placebo	61	NR	NR	NR	NR	NR
Endrikat et al. 2007	1 mg estradiol valerate/2 mg dienogest	162	NR	NR	NR	NR	NR
	Placebo	162	NR	NR	NR	NR	NR
Lin et al. 2011	Estradiol 1 mg/DSRP 2 mg	183	NR	NR	NR	NR	NR
	Placebo	61	NR	NR	NR	NR	NR
Lobo et al. 2018	Estradiol 1 mg and progesterone 100 mg	415	NR	NR	32.30%	65.30%	NR
	Estradiol 0.5 mg and progesterone 100 mg	424	NR	NR	32.10%	66.30%	NR
	Estradiol 0.5 mg and progesterone 50 mg	421	NR	NR	31.60%	65.60%	NR
	Placebo	151	NR	NR	30.50%	66.20%	NR
Simon et al. 2019 (substudy of Lobo et al. 2018)	Estradiol 1 mg and progesterone 100 mg	141	NR	NR	31.90%	67.40%	NR
	Estradiol 0.5 mg and progesterone 100 mg	149	NR	NR	32.20%	66.40%	NR
	Estradiol 0.5 mg and progesterone 50 mg	147	NR	NR	29.30%	67.30%	NR

	Placebo	135	NR	NR	30.40%	67.40%	NR
MHT: low dose (0.5 mg estradiol or equivalent) [35,38-41]							
Panay et al. 2007	Estradiol 0.5 mg and NETA 1 mg	194	NR	1.00%	0.00%	93.80%	NR
	Estradiol 0.5 mg and NETA 0.25 mg	181	NR	0.60%	1.10%	95.00%	NR
	Placebo	200	NR	0.50%	0.00%	95.50%	NR
Stevenson et al. 2010	Estradiol 1 mg/DYD 5 mg	59	NR	NR	NR	NR	NR
	Estradiol 0.5 mg/DYD 2.5 mg	122	NR	NR	NR	NR	NR
	Placebo	124	NR	NR	NR	NR	NR
Archer et al. 2013	Estradiol 0.5 mg/DRSP 0.25 mg	177	NR	0.60%	21.50%	70.60%	7.30%
	Estradiol 0.5 mg /DRSP 0.5 mg	178	NR	0.00%	23.00%	69.10%	6.70%
	Placebo	176	NR	1.10%	24.40%	67.60%	6.30%
MsFLASH 03 (Joffe et al., 2014)	Estrogen only	97	NR	NR	33.00%	61.90%	NR
	Venlafaxine 75 mg	96	NR	NR	39.60%	55.20%	NR
	Placebo	146	NR	NR	31.50%	61.60%	NR
Speroff et al. 1996	Estradiol transdermal system: 0.02 mg	54	NR	NR	NR	NR	NR
	Estradiol transdermal system: 0.04 mg	53	NR	NR	NR	NR	NR
	Placebo single dose	54	NR	NR	NR	NR	NR

	Placebo double dose	52	NR	NR	NR	NR	NR
	Estradiol transdermal system: 0.02 mg	37	NR	NR	NR	NR	NR
	Estradiol transdermal system: 0.04 mg	37	NR	NR	NR	NR	NR
	Placebo double dose	37	NR	NR	NR	NR	NR
SNRIs[44,46-48,50,64-66,84]							
Evans et al. 2005	Venlafaxine 75 mg	40	NR	5.00%	5.00%	85.00%	5.00%
	Placebo	40	NR	12.00%	12.00%	68.00%	8.00%
Speroff et al. 2008	Desvenlafaxine 100 mg	157	NR	NR	9.70%	86.20%	NR
	Desvenlafaxine 150 mg	163	NR	NR	8.80%	85.40%	NR
	Desvenlafaxine 200 mg	155	NR	NR	8.30%	87.50%	NR
	Placebo	78	NR	NR	13.00%	76.60%	NR
Wyrwich et al. 2008 (Secondary analysis of Speroff et al. 2008)	Desvenlafaxine 100 mg	145	NR	NR	9.70%	86.20%	9.00%
	Desvenlafaxine 150 mg	137	NR	NR	8.80%	85.40%	7.30%
	Desvenlafaxine 200 mg	120	NR	NR	8.30%	87.50%	6.70%
	Placebo	77	NR	NR	13.00%	76.60%	9.10%
(12-week) Archer et al. 2009	Desvenlafaxine 150 mg	151	NR	NR	18.50%	79.50%	NR
	Desvenlafaxine 100 mg	150	NR	NR	13.30%	84.70%	NR

	Placebo	150	NR	NR	14.70%	84.70%	NR
(26-weeks) Archer et al. 2009	Desvenlafaxine 100 mg	182	NR	NR	10.40%	88.50%	NR
	Desvenlafaxine 150 mg	179	NR	NR	11.20%	86.60%	NR
	Placebo	180	NR	NR	11.10%	86.70%	NR
Boekhout et al. 2011	Venlafaxine 75 mg	41	NR	NR	NR	NR	NR
	Placebo	20	NR	NR	NR	NR	NR
Bouchard et al. 2012	Desvenlafaxine 100 mg	137	NR	NR	0.70%	92.00%	7.30%
	Placebo	150	NR	NR	0.00%	93.30%	6.00%
Archer et al. 2013 Secondary data analysis of Pinkerton et al. (2013)	Desvenlafaxine 100 mg	1066	0.50%	0.70%	13.70%	83.80%	4.80%
	Placebo	1052	0.60%	1.00%	14.30%	83.40%	4.70%
SSRIs[43,68,69,85,86,89,125]							
Stearns et al. 2013	Paroxetine 12.5 mg	51	NR	0.00%	11.80%	86.30%	NR
	Paroxetine 25 mg	58	NR	0.00%	12.10%	87.90%	NR
	Placebo	56	NR	1.80%	10.70%	87.50%	NR
Kalay et al. 2007	Citalopram 20 mg	25	NR	NR	NR	NR	NR
	Placebo	25	NR	NR	NR	NR	NR
FAST Trial (Grady et al., 2007)	Sertraline 100 mg	50	NR	NR	38.00%	46.00%	NR

	Placebo	49	NR	NR	14.30%	67.30%	NR
MsFLASH 01 (Freeman et al., 2011)	Escitalopram 10 mg	104	NR	NR	45.20%	51.00%	NR
	Placebo	101	NR	NR	47.50%	48.50%	NR
Simon et al. 2011 (Study 1)	Paroxetine 7.5 mg	301	NR	0.30%	35.20%	63.10%	9.00%
	Placebo	305	NR	0.30%	30.50%	66.20%	12.10%
Simon et al. 2011 (Study 2)	Paroxetine 7.5 mg	284	NR	1.10%	24.30%	72.20%	5.60%
	Placebo	284	NR	2.10%	18.70%	78.90%	7.40%
Simon et al. 2011 (Study 1 & 2) Sleep outcomes (Pinkerton 2015)	Paroxetine 7.5 mg	585	NR	0.70%	29.90%	67.50%	NR
	Placebo	589	NR	1.20%	24.80%	72.30%	NR
Capriglione et al. 2016	Paroxetine 7.5 mg	42	NR	NR	NR	97.60%	NR
	Placebo	38	NR	NR	NR	100.00%	NR
Gabapentin[52,53,72,126]							
Guttuso et al. 2003	Gabapentin 900 mg	30	NR	NR	6.70%	93.30%	NR
	Placebo	29	NR	NR	6.90%	93.10%	NR
Reddy et al. 2006	Gabapentin 2400 mg	20	NR	NR	NR	NR	NR
	Placebo	20	NR	NR	NR	NR	NR
Butt et al. 2008	Gabapentin 900 mg	99	NR	NR	NR	79.80%	NR

	Placebo	98	NR	NR	NR	73.50%	NR
Breeze 3 (Pinkerton et al., 2014)	Gabapentin 1800* mg	299	NR	NR	24.70%	71.90%	3.00%
	Placebo	294	NR	NR	27.90%	67.00%	3.10%

AA: African American, BMI: body mass index, DSRP: drospirenone, E2: estradiol, N: total number of participants, NETA: norethindrone acetate, NR: not reported, MHT: menopausal hormone therapy, PBO: placebo, SNRI: serotonin and norepinephrine reuptake inhibitors, SSRI: selective serotonin reuptake inhibitors, mg: milligrams, CR: controlled release.
*gastroretentive gabapentin

Table D3.4. VMS Frequency

Title	Intervention	Arm Size	VMS Frequency Daily		Change / Difference from Placebo (Daily or Weekly)		Reduction in VMS			VMS Frequency Weekly	
			Daily Score Baseline (Mean, SD)	Daily Score (Mean, SD)	Change from baseline: Mean (SE)	Diff from placebo: Mean Change (SE/95% CI), P Value	Overall %	50%	75%	Weekly Score Baseline (Mean, SD)	Weekly Score (Mean, SD)
Fezolinetant[56,57,99,108,115]											
Phase 2A	Fezolinetant 90 mg	43	NR	NR	-76.1† (95% CI: -87.2, -65.0)	-35.2 (95% CI: -47.6, -22.8), p=0.001	93%	NR	NR	80.7 (NR)	5.7 (NR) (95% CI, 2.4, 9.1)
	Placebo	44	NR	NR	-35.3† (95% CI -46.9, -23.6)	REF	46%	NR	NR	72 (NR)	39 (NR) (95% CI, 26.6,515)
VESTA	Fezolinetant 30 mg	43	NR	NR	-7.4 (0.58)	-2.1 (0.75) (95% CI: -3.52 to -0.58), p=0.0064	NR	NR	NR	NR	NR
	Placebo	43	NR	NR	-5.3 (0.58)	REF	NR	NR	NR	NR	NR

SKYLIGHT 1	Fezolinetant 30 mg	173	NR	NR	NR	–2.39 (0.44), P<0.001	NR	NR	NR	NR	NR
	Fezolinetant 45 mg	174	NR	NR	NR	–2.55 (0.43), P<0.001	NR	NR	NR	NR	NR
	Placebo	175	NR	NR	NR	REF	NR	NR	NR	NR	NR
SKYLIGHT 2	Fezolinetant 30 mg	166	11.23, 4.88	4.8, 5.59	NR	-1.86 (0.55), [<0.001]	NR	NR	NR	NR	NR
	Fezolinetant 45 mg	167	11.79, 8.26	4.49, 5.39	NR	-2.53 (0.55), [<0.001]	NR	NR	NR	NR	NR
	Placebo	167	11.59, 5.02	6.73, 7.58	NR	REF	NR	NR	NR	NR	NR
MHT: standard dose (1 mg estradiol) [33,34,36,37,39]											
Schurman n et al. 2004	Estradiol 1 mg/DSRP 1 mg	55	NR	NR	-85.6% (3.0%)*‡	-38.6% (95% CI: -51.1, -26.1), p<.0001	NR	NR	NR	61.99	4.23
	Estradiol 1 mg/DSRP 2 mg	52	NR	NR	-88.0% (2.5%)*‡	-41.0% 995% CI: -53.7, -28.3), p<.0001	NR	NR	NR	67.19	7.2
	Estradiol 1 mg/DSRP 3 mg	57	NR	NR	-84.5% (3.0%)*‡	-37.5% (95% CI: -49.9, -25.1), p<.0001	NR	NR	NR	59.92	4.97
	Placebo	61	NR	NR	-47.0% (5.5%)*‡	REF	NR	NR	NR	62.44	34.97
Endrikat et al. 2007	1 mg estradiol valerate/2 mg dienogest	162	7.54	1.05	-80.8% (30.9%)‡	p<0.0001	80.8% (30.9%) p<0.0001	NR	NR	NR	NR
	Placebo	162	7.12	3.81	-41.5% (39.4%)‡	REF	41.5% (39.4%)	NR	NR	NR	NR
Lin et al. 2011	Estradiol 1 mg/DSRP 2 mg	183	NR	NR	-80.4% (21.7%)*‡	-28.5% (NR), p=0.0001	NR	NR	NR	57.8, 36.9	11.1, 15.1
	Placebo	61	NR	NR	-51.9% (32.4%)*‡	REF	NR	NR	NR	50.3, 31.1	22.4, 17.3
Lobo et al. 2018	Estradiol 1 mg and progesterone 100 mg	415	NR	NR	-55.1† (NR)	p<0.05	NR	NR	NR	74.4, 35.3	NR

	Estradiol 0.5 mg and progesterone 100 mg	424	NR	NR	-53.7† (NR)	p<0.05	NR	NR	NR	72.1, 27.8	NR
	Estradiol 0.5 mg and progesterone 50 mg	421	NR	NR	-49.85† (NR)	p<0.05	NR	NR	NR	75.9, 28	NR
	Placebo	151	NR	NR	-55.1† (NR)	p<0.05	NR	NR	NR	72.4, 23.3	NR
MHT: low dose (0.5 mg estradiol or equivalent) [35,38-41]											
Panay et al. 2007	Estradiol 0.5 mg and NETA 1 mg	194	NR	NR	-57.7† (actual values reported: 70.9 to 13.2)	p<0.001	NR	NR	NR	70.9	7.73
	Estradiol 0.5 mg and NETA 0.25 mg	181	NR	NR	-59.7† (actual values reported: 69.2 to 9.5)	p<0.001	NR	NR	NR	69.2	11.36
	Placebo	200	NR	NR	-36.7† (actual values reported: 70.0 to 33.3)	REF	NR	NR	NR	70	31.82
Stevenson et al. 2010	Estradiol 1 mg/DYD 5 mg	59	7.5,2.1	NR	-6.2 (2.6)	NR	NR	NR	NR	NR	NR
	Estradiol 0.5 mg/DYD 2.5 mg	122	8.0, 2.9	NR	-6.3 (3.4)	1.19 (95% CI 0.53, 1.86) p<0.001	NR	NR	NR	NR	NR
	Placebo	124	7.7, 2.7	NR	-4.9 (3.5)	REF	NR	NR	NR	NR	NR
Archer et al. 2013	Estradiol 0.5 mg/DRSP 0.25 mg	177	NR	NR	-55.32 (30.23)†	-22.2 (95% CI:-27.8 to -16.6), p<0.0001	NR	NR	NR	74.64	-55.32, 30.23
	Estradiol 0.5 mg /DRSP 0.5 mg	178	NR	NR	-60.33 (37.52)†	-27.6 (95% CI:-33.2 to -22.0), p<0.001	NR	NR	NR	NR	-60.33, 37.52
	Placebo	176	NR	NR	-31.92 (44.45)†	REF	NR	NR	NR	NR	-31.92, 44.45

MsFLASH 03 (Joffe et al., 2014)	Estrogen only	97	8.5, 5.7	NR	-4.5* (95% CI: -5.4, -3.6)	-2.3 (95% CI: -3.4, -1.3), p<0.001	NR	NR	NR	NR	NR
	Venlafaxine 75 mg	96	8.2, 5.5	NR	-3.9* (95% CI: -4.7, -3.1)	-1.8 (95% CI: -2.7, -0.8), p=0.005	NR	NR	NR	NR	NR
	Placebo	146	7.7, 4.9	NR	-4.5* (95% CI: -5.4, -3.6)	-2.3 (95% CI: -3.4, -1.3), p<0.001	NR	NR	NR	NR	NR
Speroff et al. 1996	Estradiol transdermal system: 0.02 mg	45	NR	NR	-47† (actual values reported 60.4 to 13.4)	p=0.088	NR	NR	NR	60.4	13.4
	Estradiol transdermal system: 0.04 mg	44	NR	NR	-50.7†	p<0.001	NR	NR	NR	60.1	9.4
	Placebo (single dose)	47	NR	NR	-41.9†	REF	NR	NR	NR	63.4¤	21.5
	Placebo (double dose)	44	NR	NR	-24.2†	REF	NR	NR	NR	60.1¤	35.9
	Estradiol transdermal system: 0.02 mg	31	NR	NR	-46.2†	p=0.004	NR	NR	NR	59.6¤	13.4
	Estradiol transdermal system: 0.04 mg	26	NR	NR	-48.4†	p=0.006	NR	NR	NR	60.8¤	12.4
	Placebo (double dose)	28	NR	NR	-31.3†	REF	NR	NR	NR	61.1¤	29.8

SNRIs[44,46-48,50,64,65,84]											
Evans et al. 2005	Venlafaxine 75 mg	40	6.9	4.7	NR*	p=0.20	1.4 episodes (95% CI 0.7 to 3.6,P .20)	NR	NR	NR	NR
	Placebo	40	9	6	NR*	REF	NR	NR	NR	NR	NR
Speroff et al. 2008	Desvenlafaxine 100 mg	157	10.5, 4.1	–7.23 (0.38)	-1.76, p=0.003[#]	NR	NR	NR	72 / 145 (49.7 %)	NR	NR
	Desvenlafaxine 150 mg	163	11.2, 6.4	–6.94 (0.38)	-0.96, p=0.11[#]	NR	NR	NR	56 / 137 (40.9 %)	NR	NR
	Desvenlafaxine 200 mg	155	11.1, 4.3	–6.46 (0.41)	-0.88, p=0.15[#]	NR	NR	NR	54 / 120 (45.0 %)	NR	NR
	Placebo	78	11, 4.6	–5.50 (0.46)	REF	NR	NR	NR	22 / 77 (28.6 %)	NR	NR
Wyrwich et al. 2008 (Secondary analysis of Speroff et al. 2008)	Desvenlafaxine 100 mg	145	11.21, 6.39	NR	-7.00 (0.39), N=109	0.96 (NR), p=0.111	NR	NR	NR	NR	NR
	Desvenlafaxine 150 mg	137	11.02, 4.62	NR	-6.04 (0.49), N=67	REF	NR	NR	NR	NR	NR
	Desvenlafaxine 200 mg	120	10.51, 4.06	NR	-7.80 (0.36), N=121	1.76 (NR), P=0.003;	NR	NR	NR	NR	NR
	Placebo	77	11.13, 4.31	NR	-6.92 (0.41), N=97	0.88 (NR), P=0.153	NR	NR	NR	NR	NR
(12-week) Archer et al. 2009	Desvenlafaxine 150 mg	151	10.5, 3.4	3.8, 2.4	-7.0 (0.35)	p=0.012	66.6% (P =.012)	111 / 143 (77.6%)	76 / 151 (50.3 %)	NR	NR

	Desvenlafaxine 100 mg	150	11.1, 4.5	3.8, 2.3	-7.1 (0.34)	p=0.005	65.4% (P =.005)	107 / 143 (74.8%)	71 / 150 (47.3 %)	NR	NR
	Placebo	150	10.9, 4.6	5.07, 1.9	-5.8 (0.34)	REF	50.80%	77 / 150 (51.3%)	44 / 150 (29.3 %)	NR	NR
(26-weeks) Archer et al. 2009	Desvenlafaxine 100 mg	182	10.8, 4.2	4.3, 2.3	-6.3 (0.34)	p=0.002	60%	110 / 162 (67.9%)	67 / 162 (41.4 %)	NR	NR
	Desvenlafaxine 150 mg	179	10.3, 4.1	3.5, 2.1	-7.0 (0.35)	p<0.001	66%	108 / 144 (75.0%)	65 / 144 (45.1 %)	NR	NR
	Placebo	180	10.6, 4	5.6, 2.3	-4.9 (0.31)	REF	47%	85 / 178 (47.8%)	47 / 178 (26.4 %)	NR	NR
Boekhout et al. 2011	Venlafaxine 75 mg	41	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	20	NR	NR	NR	NR	NR	NR	NR	NR	NR
Bouchard et al. 2012	Desvenlafaxine 100 mg	137	10.1, 4.2	4.51	-5.78 (0.33)	p=0.921	57.70%	90 / 137 (65.7%)	55 / 137 (40.1 %)	NR	NR
	Placebo	150	9.6, 2.9	4.36	-5.82 (0.31)	REF	57.50%	92 / 150 (61.3%)	57 / 150 (38.0 %)	NR	NR
Pinkerton et al. 2013	Desvenlafaxine 100 mg	1066	NR	NR	-7.5 (0.35)	-2.48 (95% CI:-3.47, -1.50), p<0.001	3 months: 64% 12 months: 66%	NR	NR	NR	NR

	Placebo	1052	NR	NR	-5 (0.35)	REF	3 months; 43; % 12 months: 41%	NR	NR	NR	NR
SSRIs[43,68,85,86,89,125]											
Stearns et al. 2013	Paroxetine 12.5 mg CR	51	7.1	3.8	-3.3* (NR)	-1.55* (95% CI: -2.76 to -0.34), p=0.01	NR	30 / 51 (58.8%) (odds ratio: 1.95; 95% CI, 0.86-4.40; p=0.11).	NR	NR	NR
	Paroxetine 25 mg CR	58	6.4	3.2	-3.2* (NR)	-1.50* (95% CI: -2.66 to -0.34); p= 0.01	NR	37 / 58 (63.8%)(OR, 2.56; 95% CI, 1.15-5.68; p=0.02).	NR	NR	NR
	Placebo	56	6.6	4.8	-1.8* (NR)	REF	NR	24 / 56 (42.9%)	NR	NR	NR
Kalay et al. 2007	Citalopram 20 mg	25	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	25	NR	NR	NR	NR	NR	NR	NR	NR	NR

FAST Trial (Grady et al., 2007)	Sertraline 100 mg	50	8.6, 4.4	5.1, 4.7	-39%*‡ (44.8)	-0.7% (95% CI:–15.9 to 17.2%), p=0.94	39%; DIFF: 0.7%, 95% CI: –15.9 to 17.2%;P=0.94	NR	NR	NR	NR
	Placebo	49	9.3, 7.2	5.8, 5.3	-38%*‡ (32.8)	REF	38%	NR	NR	NR	NR
MsFLASH 01 (Freeman et al., 2011)	Escitalopram 10 mg	104	9.88, 6.32	5.26, 5.9	-4.60* (95% CI:−5.47 to −3.74)	−1.41* (95% CI:−2.69 to −0.13), p<.001	NR	NR	NR	NR	NR
	Placebo	101	9.66, 5.02	6.43, 6.56	−3.20* (95% CI:−4.15 to −2.24)	REF	NR	NR	NR	NR	NR
Simon et al. 2011 (Study 1)	Paroxetine 7.5 mg	301	11.79, 4.87	NR	-43.5*† (NR)	p=0.009	NR	NR	NR	NR	-43.5 (p=0.0090)
	Placebo	305	11.65, 4.39	NR	-37.3*† (NR)	REF	NR	NR	NR	NR	-37.3
Simon et al. 2011 (Study 2)	Paroxetine 7.5 mg	284	10.83, 3.86	NR	-43.5*† (NR)	p=0.009	NR	NR	NR	NR	-37.2, p=0.0001
	Placebo	284	10.9, 3.96	NR	-37.3*† (NR)	REF	NR	NR	NR	NR	-27.6
Capriglione et al. 2016	Paroxetine 7.5 mg	42	12.21, 3.43	NR	-46.5† (NR)	p=0.009	NR	NR	NR	NR	NR
	Placebo	38	12.15, 3.23	NR	-39.3† (NR)	REF	NR	NR	NR	NR	NR
Gabapentin[52,53,72,126]											
Guttuso et al. 2003	Gabapentin 900 mg	30	10.8, 4.1	NR	-45% (SD=31.5)*§	-20.9 (95% CI 2.7, 34.0), p=0.02	44.6 (31.5)	NR	NR	NR	NR
	Placebo	29	10.3, 3.7	NR	-29% (SD=32.1)*§	REF	28.9 (32.1)	NR	NR	NR	NR
Reddy et al. 2006	Gabapentin 2400 mg	20	NR	NR	NR	NR	NR	NR	NR	94.78, 61.32	27.05
	Placebo	20	NR	NR	NR	NR	NR	NR	NR	82.18, 28.92	42.56

Butt et al. 2008	Gabapentin 900 mg	99	8.5, 4.6	4.5, 3.2	-46%*†‡ (95% CI: 38.7, 52.7	P<0.001	45.7 (38.7-52.7)	NR	NR	NR	NR
	Placebo	98	8.5, 5.1	6.5, 4.5	-25%*†‡ (95% CI: 17.3, 32.1)	REF	24.7 (17.3-32.1)	NR	NR	NR	NR
Breeze 3 (Pinkerton et al., 2014)	Gabapentin 1800* mg	299	11.8, 4.7	NR	-7.64* (NR)	-1.14 (95% CI: 1.8 to -0.48), p=0.0007	NR	72.90 %	NR	NR	NR
	Placebo	294	12, 5.5	NR	-6.50* (NR)	REF	NR	59.90 %	NR	NR	NR

CI: confidence interval, DSRP: drospirenone, E2: estradiol, LS: least squares, LSMD: least squares mean difference, MHT: menopausal hormonal therapy, mg: milligrams, CR: controlled release, N: total number of participants, NETA: norethindrone acetate, NR: not reported, PBO: placebo, REF: reference, SD: standard deviation, SE: standard error, SNRI: serotonin and norepinephrine reuptake inhibitors, SSRI: selective serotonin reuptake inhibitors, VMS: vasomotor symptoms.

*All VMS (mild, moderate, and severe)

†Weekly score

‡Percentage change

§VMS severity was rated from 1 (mild) to 4 (very severe) and multiplied by the number of VMS events at that level.

#Wyrwich et al. (2008) reported mean difference from placebo for efficacy data from Speroff et al. (2008) and adjusted for site and baseline values. N=121 desvenlafaxine 100 mg, 109 desvenlafaxine 150 mg, 97 desvenlafaxine 200 mg, and 67 placebo.

¤Data reported from week 1

Table D3.5. VMS Severity

Title	Intervention	Arm Size	VMS Severity Daily		Change / Difference from Placebo (Daily or Weekly)		Reduction in VMS			VMS Severity Weekly	
			Daily Score Baseline (Mean, SD)	Daily Score (Mean, SD)	Change from baseline: LS Mean (SE)	Diff from placebo: Mean Change (SE/95% CI), P Value	Overall %	50 %	75 %	Weekly Score Baseline (Mean, SD)	Weekly Score (Mean, SD)
Fezolinetant[55-57,99,108,115]											
Phase 2A	Fezolinetant 90 mg	43	NR	NR	-26.6† (95% CI: -31.1, -22.2)	-12.4 (95% CI: -17.0, -7.8), p=0.001	93%	NR	NR	27.8	1.7
	Placebo	44	NR	NR	-12.1† (95% CI: -16.6, -7.7)	REF	46%	NR	NR	24.9	13.5
VESTA	Fezolinetant 30 mg	43	NR	NR	-0.9 (0.16)	-0.2 (0.21) (95% CI: -0.58 to 0.26), p=0.4647	NR	NR	NR	NR	NR
	Placebo	43	NR	NR	-0.8 (0.16)	REF	NR	NR	NR	NR	NR
SKYLIGHT 1	Fezolinetant 30 mg	173	NR	NR	NR	–0.2 (0.08), p<0.007	NR	NR	NR	NR	NR
	Fezolinetant 45 mg	174	NR	NR	NR	–0.24 (0.08), p<0.002	NR	NR	NR	NR	NR
	Placebo	175	NR	NR	NR	REF	NR	NR	NR	NR	NR
SKYLIGHT 2	Fezolinetant 30 mg	166	2.44, 0.33	1.84, 0.79	NR	-0.16 (0.08), p=0.049	NR	NR	NR	NR	NR
	Fezolinetant 45 mg	167	2.41, 0.34	1.66, 0.79	NR	-0.29 (0.08), p<0.001	NR	NR	NR	NR	NR

	Placebo	167	2.41, 0.32	1.95, 0.68	NR	REF	NR	NR	NR	NR	NR
MHT: standard dose (1 mg estradiol) [33,34,36,37,39,62]											
Schurman n et al. 2004	Estradiol 1 mg/DSRP 1 mg	55	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Estradiol 1 mg/DSRP 2 mg	52	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Estradiol 1 mg/DSRP 3 mg	57	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	61	NR	NR	NR	NR	NR	NR	NR	NR	NR
Endrikat et al. 2007	1 mg estradiol valerate/2 mg dienogest	162	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	162	NR	NR	NR	NR	NR	NR	NR	NR	NR
Lin et al. 2011	Estradiol 1 mg/DSRP 2 mg	183	NR	NR	-0.57 (0.85)	-0.30, p=0.103	NR	NR	NR	2.27, 0.25	NR
	Placebo	61	NR	NR	-0.28 (0.58)	REF	NR	NR	NR	2.34, 0.3	NR
Lobo et al. 2018	Estradiol 1 mg and progesterone 100 mg	415	NR	NR	-1.12 (NR)	p<0.05	NR	NR	NR	NR	NR
	Estradiol 0.5 mg and progesterone 100 mg	424	NR	NR	-0.90 (NR)	p<0.05	NR	NR	NR	NR	NR
	Estradiol 0.5 mg and progesterone 50 mg	421	NR	NR	-0.76 (NR)	p<0.05	NR	NR	NR	NR	NR
	Placebo	151	NR	NR	-1.12 (NR)	p<0.05	NR	NR	NR	NR	NR
Simon et al. 2019 (Substudy of Lobo et al. 2018)	Estradiol 1 mg and progesterone 100 mg	141	NR	NR	NR	NR	NR	NR	NR	2.54, 0.32	NR
	Estradiol 0.5 mg and progesterone 100 mg	149	NR	NR	NR	NR	NR	NR	NR	2.51, 0.25	NR
	Estradiol 0.5 mg and progesterone 50 mg	147	NR	NR	NR	NR	NR	NR	NR	2.5, 0.23	NR
	Placebo	135	NR	NR	NR	NR	NR	NR	NR	2.52, 0.25	NR

MHT: low dose (0.5 mg estradiol or equivalent)[35,38-41]											
Panay et al. 2007	Estradiol 0.5 mg and NETA 1 mg	194	NR	NR	NR	NR	NR	NR	NR	185.8	32.4
	Estradiol 0.5 mg and NETA 0.25 mg	181	NR	NR	NR	NR	NR	NR	NR	180.5	23.6
	Placebo	200	NR	NR	NR	NR	NR	NR	NR	183.5	87.2
Stevenson et al. 2010	Estradiol 1 mg/DYD 5 mg	60	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Estradiol 0.5 mg/DYD 2.5 mg	122	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	125	NR	NR	NR	NR	NR	NR	NR	NR	NR
Archer et al. 2013	Estradiol 0.5 mg/DRSP 0.25 mg	177	NR	NR	NR	-1.21 (1.08)	-0.80 (95% CI:-1.01 to -0.59), p<0.0001	NR	NR	2.58	-1.21, 1.08
	Estradiol 0.5 mg /DRSP 0.5 mg	178	NR	NR	NR	- 1.45 (1.12)	-1.07 (95% CI:-1.28 to -0.86), p<0.0001	NR	NR	NR	-1.45, 1.12
	Placebo	176	NR	NR	NR	-0.39 (0.77)	REF	NR	NR	2.52	-0.39, 0.77
MsFLASH 03 (Joffe et al., 2014)	Estrogen only	97	NR	NR	NR	-0.3* (95% CI: -0.4, 0.1), p=0.02	NR	NR	NR	NR	NR
	Venlafaxine 75 mg	96	NR	NR	NR	-0.2* (95% CI: -0.3, 0.0), p=0.02	NR	NR	NR	NR	NR
	Placebo	146	NR	NR	NR	REF	NR	NR	NR	NR	NR
Speroff et al. 1996	Estradiol transdermal system: 0.02 mg	54	NR	NR	NR	NR	NR	NR	NR	NR	NR

(study 1 and study 2)	Estradiol transdermal system: 0.04 mg	53	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo single dose	54	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo double dose	52	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Estradiol transdermal system: 0.02 mg	37	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Estradiol transdermal system: 0.04 mg	37	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo double dose	37	NR	NR	NR	NR	NR	NR	NR	NR	NR
SNRIs[44,46-48,50,64,65,84]											
Evans et al. 2005	Venlafaxine 75 mg	40	NR	NR	NR*	p=0.30	2.6 points, (95% CI -2.3 to 7.5)	NR	NR	12.6	7.1
	Placebo	40	NR	NR	NR*	REF	NR	NR	NR	16.1	9.9
Speroff et al. 2008	Desvenlafaxine 100 mg	157	2.4, 0.3	NR	–0.80* (0.06)	-0.33, p=0.006[#]	NR	NR	NR	NR	NR
	Desvenlafaxine 150 mg	163	2.4, 0.3	NR	–0.59* (0.07)	-0.09, p=0.47[#]	NR	NR	NR	NR	NR
	Desvenlafaxine 200 mg	155	2.4, 0.3	NR	–0.74* (0.07)	-0.25, p=0.04[#]	NR	NR	NR	NR	NR
	Placebo	78	2.5, 0.3	NR	–0.47* (0.09)	REF	NR	NR	NR	NR	NR
Wyrwich et al. 2008 (Secondary analysis of Speroff et al. 2008)	Desvenlafaxine 100 mg	145	2.38, 0.27	NR	-0.64 (0.08), N=109	0.09, P=0.466	NR	NR	NR	NR	NR
	Desvenlafaxine 150 mg	137	2.47, 0.32	NR	-0.55 (0.10), P=67	REF	NR	NR	NR	NR	NR
	Desvenlafaxine 200 mg	120	2.39, 0.29	NR	-0.88 (0.07), N=121	0.33, P=0.006;	NR	NR	NR	NR	NR

	Placebo	77	2.36, 0.32	NR	-0.80 (0.08), N=97	0.25, P=0.040	NR	NR	NR	NR	NR
(12-week) Archer et al. 2009	Desvenlafaxine 150 mg	151	2.4, 0.3	NR	-0.66* (0.07)	p<0.001	73.0%; P =.025	NR	NR	NR	NR
	Desvenlafaxine 100 mg	150	2.4, 0.3	NR	-0.65* (0.07)	p<0.001	73.0%; P =.025	NR	NR	NR	NR
	Placebo	150	2.4, 0.3	NR	-0.33* (0.07)	REF	60.40 %	NR	NR	NR	NR
(26-weeks) Archer et al. 2009	Desvenlafaxine 100 mg	182	2.4, 0.3	NR	-0.54* (0.07)	p=0.002	24%	NR	NR	NR	NR
	Desvenlafaxine 150 mg	179	2.4, 0.3	NR	-0.71* (0.07)	p<0.001	29%	NR	NR	NR	NR
	Placebo	180	2.4, 0.3	NR	-0.28* (0.06)	REF	13%	NR	NR	NR	NR
Boekhout et al. 2011	Venlafaxine 75 mg	41	13.3	NR	-41%‡	P=0.07	NR	NR	NR	NR	NR
	Placebo	20	14.4	NR	-29%‡	REF	29% ‡ (P<0.001)	NR	NR	NR	NR
Bouchard et al. 2012	Desvenlafaxine 100 mg	137	2.2, 0.3	NR	-0.61* (0.07)	REF	26.80 %	NR	NR	NR	NR
	Placebo	150	2.2, 0.3	NR	-0.61* (0.07)	p=0.943	26.50 %	NR	NR	NR	NR
Pinkerton et al. 2013	Desvenlafaxine 100 mg	1066	2.4, 0.3	NR	-0.63* (0.05)	-0.33 (95% CI:-0.48, -0.18), p<0.001	6months: 27% 12 months: 32%	NR	NR	NR	NR
	Placebo	1052	2.4, 0.4	NR	-0.3* (0.05)	REF	6months: 13; % 12	NR	NR	NR	NR

							months: 19%				
SSRIs[43,68,85,86,89,125]											
Stearns et al. 2003	Paroxetine 12.5 mg CR	51	16.5	8.14	−8.52§ (1.27)	−4.7 (95% CI:−8.1 to −1.3), p=0.007	62.20 %	NR	NR	NR	NR
	Paroxetine 25 mg CR	58	13.6	6.28	−7.43§ (1.18)	−3.6 (95% CI:−6.8 to −0.4), p=0.03	64.60 %	NR	NR	NR	NR
	Placebo	56	14.2	10.35	−3.82§ (1.17)	REF	37.80 %	NR	NR	NR	NR
Kalay et al. 2007	Citalopram 20 mg	25	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	25	NR	NR	NR	NR	NR	NR	NR	NR	NR
FAST Trial (Grady et al., 2007)	Sertraline 100 mg	50	16.4, 10.6	8.8, 9.6	-42%*‡ (48.0)	-1.6% (95% CI: −16.4 to 19.6%), p=0.86	NR	NR	NR	NR	NR
	Placebo	49	18.4, 17.9	10.3, 10.4	-41%*‡ (36.5)	REF	NR	NR	NR	NR	NR
MsFLASH 01 (Freeman et al., 2011)	Escitalopram 10 mg	104	2.16, 0.44	1.63, 0.62	−0.52* (95% CI:−0.64 to −0.40)	−0.22 (95% CI:−0.40 to −0.05), p<0.001	NR	NR	NR	NR	NR
	Placebo	101	2.19, 0.47	1.89, 0.62	−0.30* (95% CI:−0.42 to −0.17)	REF	NR	NR	NR	NR	NR
Simon et al. 2011 (Study 1)	Paroxetine 7.5 mg	301	2.53, 0.3	NR	-0.10* (NR)	p=0.29	NR	NR	NR	NR	-0.10, p= 0.2893
	Placebo	305	2.53, 0.31	NR	-0.09* (NR)	REF	NR	NR	NR	NR	-0.09
Simon et al. 2011 (Study 2)	Paroxetine 7.5 mg	284	2.53, 0.3	NR	-0.12* (NR)	p=0.01	NR	NR	NR	NR	NR
	Placebo	284	2.53, 0.32	NR	-0.07* (NR)	REF	NR	NR	NR	NR	NR

Capriglione et al. 2016	Paroxetine 7.5 mg	42	NR	NR	-0.09 (NR)	P=0.005	−0.09; p =0.0048	NR	NR	NR	NR
	Placebo	38	NR	NR	-0.05 (NR)	REF	−0.05	NR	NR	NR	NR
Gabapentin[52,53,72,126]											
Guttuso et al. 2003	Gabapentin 900 mg	30	44.5, 19	NR	-54%* (SD=35.9)	-25.5% (95% CI 6.8, 42.3), p=0.01	53.5 (35.9)	NR	NR	NR	NR
	Placebo	29	39.5, 19.1	NR	-31%* (SD=38.7)	REF	31.4 (38.7)	NR	NR	NR	NR
Reddy et al. 2006	Gabapentin 2400 mg	20	NR	NR	NR*†	t=3.03, p=.004	NR	NR	NR	233.88, 159.14	27.05
	Placebo	20	NR	NR	NR*†	NR	NR	NR	NR	193.73, 82.7	42.56
Butt et al. 2008	Gabapentin 900 mg	99	19.6, 13.5	9.5, 9.6	-51.0%*‡ (95% CI: 43.3, 58.5)	p<0.001	NR	NR	NR	NR	NR
	Placebo	98	18.3, 16.9	13.5, 11.6	-26.5%*‡ (95% CI: 18.3, 34.7)	REF	NR	NR	NR	NR	NR
Breeze 3 (Pinkerton et al., 2014)	Gabapentin 1800* mg	299	2.55, 0.29	NR	-0.65* (NR)	-0.19 (95% CI: -0.33 to -0.04), p=0.012	NR	NR	NR	NR	NR
	Placebo	294	2.54, 0.28	NR	-0.46* (NR)	REF	NR	NR	NR	NR	NR

CI: confidence interval, DSRP: drospirenone, E2: estradiol, LS: least squares, LSMD: least squares mean difference, MHT: menopausal hormonal therapy, mg: milligrams, CR: controlled release, N: total number of participants, NETA: norethindrone acetate, NR: not reported, PBO: placebo, REF: reference, SD: standard deviation, SE: standard error, SNRI: serotonin and norepinephrine reuptake inhibitors, SSRI: selective serotonin reuptake inhibitors, VMS: vasomotor symptoms

*gastroretentive gabapentin

*All VMS (mild, moderate, and severe)

†Weekly score

‡Percentage change

§VMS severity was rated from 1 (mild) to 4 (very severe) and multiplied by the number of VMS events at that level.

#Wyrwich et al. (2008) reported mean difference from placebo for efficacy data from Speroff et al. (2008) and adjusted for site and baseline values. N=121 desvenlafaxine 100 mg, 109 desvenlafaxine 150 mg, 97 desvenlafaxine 200 mg, and 67 placebo.

Table D3.6. Sleep Outcomes

Title	Intervention	Arm Size	Sleep Problems	LSEQ Change from Baseline (95% CI)	PROMIS SD SF 8b		PSQI		ISI		MOS	
					(Mean, SD)	Change from Baseline (SD)	(Mean, SD)	Change from Baseline (SD)	(Mean, SD)	Change from Baseline (SD)	(Mean, SD)	Change from Baseline (SD)
Fezolinetant[56,57,108,115]												
Phase 2A	Fezolinetant 90 mg	43	NR	4.4 (3.3, 5.5); LSMD: 2.43; 95% CI: (1.33, 3.53); p<0.001	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	44	NR	1.9 (1.0, 2.8)	NR	NR	NR	NR	NR	NR	NR	NR
VESTA	Fezolinetant 30 mg	43	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	43	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
SKYLIGHT 1	Fezolinetant 30 mg	173	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Fezolinetant 45 mg	174	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	175	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
SKYLIGHT 2	Fezolinetant 30 mg	166	NR	NR	27.4 (6.7)	-4.1 (0.5)	NR	NR	NR	NR	NR	NR
	Fezolinetant 45 mg	167	NR	NR	26.2 (6.6)	-5.5 (0.5)	NR	NR	NR	NR	NR	NR
	Placebo	167	NR	NR	27.4 (7.0)	NR	NR	NR	NR	NR	NR	NR

MHT: standard dose (1 mg estradiol)[34,36,37,59]												
Schurmann et al. 2004	Estradiol 1 mg/DSRP 1 mg	55	35.1%	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Estradiol 1 mg/DSRP 2 mg	52	23.7%	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Estradiol 1 mg/DSRP 3 mg	57	21.4%	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	61	52.2%	NR	NR	NR	NR	NR	NR	NR	NR	NR
Endrikat et al. 2007	1 mg estradiol valerate/2 mg dienogest	162	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	162	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Lin et al. 2011	Estradiol 1 mg/DSRP 2 mg	183	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	61	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Kagan et al. 2019 (Secondary analyses from Lobo et al. 2018)	Estradiol 1 mg and progesterone 100 mg	415	NR	NR	NR	NR	NR	NR	NR	NR	44, 18.7	Week 12: -4.88 (1.6) Month 6: -5.39 (1.7) Month 12: -6.54 (1.9)
	Estradiol 0.5 mg and progesterone 100 mg	424	NR	NR	NR	NR	NR	NR	NR	NR	43.2, 18.3	Week 12: -3.61 (1.6)

												Month 6: -5.39 (1.7) Month 12: -7.61 (1.8)
	Estradiol 0.5 mg and progesterone 50 mg	421	NR	NR	NR	NR	NR	NR	NR	NR	44.2, 19	Week 12: -3.44 (1.6) Month 6: -4.88 (1.7) Month 12: -7.44 (1.8)
	Placebo	151	NR	NR	NR	NR	NR	NR	NR	NR	48.1, 19	NR
MHT: low dose (0.5 mg estradiol or equivalent)[35,38-41,78]												
Panay et al. 2007	Estradiol 0.5 mg and NETA 1 mg	194	Week 24: +40%	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Estradiol 0.5 mg and NETA 0.25 mg	181	Week 24: +49%	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	200	Week 24: +16%	NR	NR	NR	NR	NR	NR	NR	NR	NR
Stevenson et al. 2010	Estradiol 1 mg/DYD 5 mg	60	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Estradiol 0.5 mg/DYD 2.5 mg	122	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

	Placebo	125	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Archer et al. 2013	Estradiol 0.5 mg/DRSP 0.25 mg	177	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Estradiol 0.5 mg /DRSP 0.5 mg	178	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	176	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
MsFLASH 03 Ensrud et al. 2015 (secondary analyses of Joffe et al 2014)	Estrogen only	97	NR	NR	NR	NR	NR	−2.2; 95% CI: (−2.8, −1.6), DIFF FROM PBO: −1.0; 95% CI: (−1.8, −0.2), p=0.04	10.9	−4.1; 95% CI: (−5.3, −3.0) DIFF FROM PBO: −1.1; 95% CI: (−2.4, 0.2), p=0.09.	NR	NR
	Venlafaxine 75 mg	96	NR	NR	NR	NR	NR	−2.3; 95% CI: (−2.9, −1.6) DIFF FROM PBO: −1.0 95% CI: (−1.8, −0.2), p=0.06	11.7	−5.0; 95% CI: (−6.1, −3.9) DIFF FROM PBO: −2.0 95% CI: (−3.3, −0.7), p=0.007	NR	NR
	Placebo	146	NR	NR	NR	NR	NR	−1.2; 95% CI: (−1.7, −0.8)	10.4	−3.0; 95% CI: (−3.8, −2.3)	NR	NR

Speroff et al. 1996	Estradiol transdermal system: 0.02 mg	54	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Estradiol transdermal system: 0.04 mg	53	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo single dose	54	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo double dose	52	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Estradiol transdermal system: 0.02 mg	37	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Estradiol transdermal system: 0.04 mg	37	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo double dose	37	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
SNRIs[44,46-48,50,64,65,84]												
Evans et al. 2005	Venlafaxine 75 mg	40	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	40	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Speroff et al. 2008	Desvenlafaxine 100 mg	157	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Desvenlafaxine 150 mg	163	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Desvenlafaxine 200 mg	155	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	78	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Wyrwich et al. 2008 (Secondary analysis of	Desvenlafaxine 100 mg	145	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Desvenlafaxine 150 mg	137	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

Speroff et al. 2008)	Desvenlafaxine 200 mg	120	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	77	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
(12-week) Archer et al. 2009	Desvenlafaxine 150 mg	151	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Desvenlafaxine 100 mg	150	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	150	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
(26-weeks) Archer et al. 2009	Desvenlafaxine 100 mg	182	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Desvenlafaxine 150 mg	179	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	180	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Boekhout et al. 2011	Venlafaxine 75 mg	41	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	20	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Bouchard et al. 2012	Desvenlafaxine 100 mg	137	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	150	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Pinkerton et al. 2013	Desvenlafaxine 100 mg	1066	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	1052	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

SSRIs[43,68,69,79,85,86,89,125]												
Stearns et al. 2003	Paroxetine 12.5 mg CR	51	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Paroxetine 25 mg CR	58	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	56	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Kalay et al. 2007	Citalopram 20 mg	25	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	25	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
FAST Trial (Grady et al., 2007)	Sertraline 100 mg	50	NR	NR	NR	NR	NR	–1.2 (4.5)	NR	NR	NR	NR
	Placebo	49	NR	NR	NR	NR	NR	–1.3 (2.4)	NR	NR	NR	NR
MsFLASH 01 (Freeman et al., 2011; Guthrie et al., 2018)	Escitalopram 10 mg	104	NR	NR	NR	NR	10.4, 3.4	-1.50	16.7, 3.8	-1.84	NR	NR
	Placebo	101	NR	NR	NR	NR		REF		REF	NR	NR
Simon et al. 2011 (Study 1)	Paroxetine 7.5 mg	301	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	305	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Simon et al. 2011 (Study 2)	Paroxetine 7.5 mg	284	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	284	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Simon et al. 2011 (Study 1 & 2) Sleep	Paroxetine 7.5 mg	585	Nighttime awakenings: -54%, p< 0.0001.	NR	NR	NR	NR	NR	NR	NR	NR	NR

substudy (Pinkerton et al., 2015)			Duration of sleep: +9% (35 mins)									
	Placebo	589	Nighttime awakenings: -43%, p< 0.0001. Duration of sleep: +6% (23 mins)	NR	NR	NR	NR	NR	NR	NR	NR	NR
Capriglione et al. 2016	Paroxetine 7.5 mg	42	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	38	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Gabapentin[52,53,71,72,126]												
Guttuso et al. 2003	Gabapentin 900 mg	30	NR	NR	NR	NR	-2.9 (3.3)	NR	NR	NR	NR	NR
	Placebo	29	NR	NR	NR	NR	-1.2 (3)	NR	NR	NR	NR	NR
Yurcheshen et al. 2009 (secondary analysis of Guttuso et al. 2003)	Gabapentin 900 mg	30	NR	NR	NR	NR	NR	-2.78 DIFF FROM PBO: -1.63; 95% CI: (-3.37, -0.12), p=0.07	NR	NR	NR	NR
	Placebo	29	NR	NR	NR	NR	NR	-1.16	NR	NR	NR	NR
Reddy et al. 2006	Gabapentin 2400 mg	20	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	20	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

Butt et al. 2008	Gabapentin 900 mg	99	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	98	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Breeze 3 (Pinkerton et al., 2014)	Gabapentin 1800* mg	299	-3.1, p=0.0001	NR	NR	NR	NR	NR	NR	NR	NR	NR
	Placebo	294	-2.2	NR	NR	NR	NR	NR	NR	NR	NR	NR

CI: confidence interval, DSRP: drospirenone, E2: Estradiol, ISI: insomnia severity index, LSEQ: Leeds sleep evaluation questionnaire, LSM: least square mean, MHT: menopausal hormonal therapy, mg: milligrams per deciliter, CR: controlled release, MOS: medical outcomes study, N: total number of participants, NETA: norethindrone acetate, NR: not reported, DIFF FROM PBO: difference from placebo, DIF FROM PBO: difference from placebo, PBO: placebo, PROMIS SB SF 8b: Patient reported outcomes measurement information system sleep disturbance and sleep related impairment item banks, PSQI: , SD: standard deviation, SIS: Sleep Interference Scale, SNRI: serotonin and norepinephrine reuptake inhibitors, SSRI: selective serotonin reuptake inhibitors
*gastroretentive gabapentin

Table D3.7. Patient Reported Outcomes

Title	Intervention	Arm Size	Patient Global Impression of Change	HFRDIS Change from baseline, LSM (95% CI)	GCS Change from baseline LSM (95% CI)	SDS change form baseline, LSM (95% CI)	MENQoL	POMS change from baseline, LSM (95% CI)
Fezolinetant[54-57,99,108,115]								
Phase 2A	Placebo	44	NR	-1.98; 95% CI (-2.83, -1.13); p<0.001	-6.3; 95% CI (-9.3, -3.3); p<0.0001	-5.3; 95% CI: (-7.8, -2.8); p<0.001	NR	NR
	Fezolinetant 90 mg	43	NR	-4.1; 95% CI (-4.8, -3.5)	-13; 95% CI: (-16.4, -9.6)	-11.7 (-13.9, -9.6)	NR	NR
VESTA	Fezolinetant 30 mg	43	NR	-3.3 (0.4)	-2.9 (0.3)	NR	-0.2 (-0.7, 0.3)	NR
	Placebo	43	NR	-2.9 (0.3)	-2.1 (0.3)	NR	REF	NR

SKYLIGHT 1	Fezolinetant 30 mg	173	NR	NR	NR	NR	NR	NR
	Fezolinetant 45 mg	174	NR	NR	NR	NR	NR	NR
	Placebo	175	NR	NR	NR	NR	NR	NR
SKYLIGHT 2	Fezolinetant 30 mg	166	NR	NR	NR	NR	NR	NR
	Fezolinetant 45 mg	167	NR	NR	NR	NR	NR	NR
	Placebo	167	NR	NR	NR	NR	NR	NR
MHT: standard dose (1 mg estradiol) [33,34,36,37,62]								
Schurmann et al. 2004	Placebo	61	NR	NR	NR	NR	NR	NR
	Estradiol 1 mg/DSRP 1 mg	55	NR	NR	NR	NR	NR	NR
	Estradiol 1 mg/DSRP 2 mg	52	NR	NR	NR	NR	NR	NR
	Estradiol 1 mg/DSRP 3 mg	57	NR	NR	NR	NR	NR	NR
Endrikat et al. 2007	1 mg estradiol valerate/2 mg dienogest	162	NR	NR	NR	NR	NR	NR
	Placebo	162	NR	NR	NR	NR	NR	NR
Lin et al. 2011	Estradiol 1 mg/DSRP 2 mg	183	87.9% 'much improved' or 'very much improved'	NR	NR	NR	NR	NR
	Placebo	61	47.3% 'much improved' or 'very much improved'	NR	NR	NR	NR	NR
Lobo et al. 2018	Estradiol 1 mg and progesterone 100 mg	415	68–73%	NR	NR	NR	NR	NR
	Estradiol 0.5 mg and progesterone 100 mg	424		NR	NR	NR	NR	NR
	Estradiol 0.5 mg and progesterone 50 mg	421		NR	NR	NR	NR	NR
	Placebo	151	52%	NR	NR	NR	NR	NR
	Estradiol 1 mg and progesterone 100 mg	141	NR	NR	NR	NR	-1.92	NR

Simon, J.A. 2019 (Substudy of Lobo et al. 2018)	Estradiol 0.5 mg and progesterone 100 mg	149	NR	NR	NR	NR	-1.62	NR
	Estradiol 0.5 mg and progesterone 50 mg	147	NR	NR	NR	NR	-1.9	NR
	Placebo	135	NR	NR	NR	NR	-1.39	NR
MHT: low dose (0.5 mg estradiol or equivalent)[35,38-41]								
Panay et al. 2007	Estradiol 0.5 mg and NETA 1 mg	194	NR	NR	-10.87, p vs. placebo<0.001	NR	NR	NR
	Estradiol 0.5 mg and NETA 0.25 mg	181	NR	NR	-10.5, p vs. placebo<0.001	NR	NR	NR
	Placebo	200	NR	NR	-6.27	NR	NR	NR
Stevenson et al. 2010	Estradiol 1 mg/DYD 5 mg	60	NR	NR	NR	NR	NR	NR
	Estradiol 0.5 mg/DYD 2.5 mg	122	NR	NR	NR	NR	NR	NR
	Placebo	125	NR	NR	NR	NR	NR	NR
Archer et al. 2013	Estradiol 0.5 mg/DRSP 0.25 mg	177	NR	NR	NR	NR	NR	NR
	Estradiol 0.5 mg /DRSP 0.5 mg	178	NR	NR	NR	NR	NR	NR
	Placebo	176	NR	NR	NR	NR	NR	NR
MsFLASH 03 (Joffe et al., 2014)	Estrogen only	97	NR	DIFF from PBO: −9.3; 95% CI: (−15.3 to −3.4), p<0.001	NR	NR	NR	NR
	Venlafaxine 75 mg	96	NR	DIFF from PBO: −6.4; 95% CI: (−12.7 to −0.1), p=0.03	NR	NR	NR	NR
	Placebo	146	NR	REF	NR	NR	NR	NR
Speroff et al. 1996	Estradiol transdermal system: 0.02 mg	54	NR	NR	NR	NR	NR	NR
	Estradiol transdermal system: 0.04 mg	53	NR	NR	NR	NR	NR	NR
	Placebo single dose	54	NR	NR	NR	NR	NR	NR
	Placebo double dose	52	NR	NR	NR	NR	NR	NR

	Estradiol transdermal system: 0.02 mg	37	NR	NR	NR	NR	NR	NR
	Estradiol transdermal system: 0.04 mg	37	NR	NR	NR	NR	NR	NR
	Placebo double dose	37	NR	NR	NR	NR	NR	NR
SNRIs[44,46-48,50,64,65,84]								
Evans et al. 2005	Venlafaxine 75 mg	40	NR	NR	NR	NR	NR	NR
	Placebo	40	NR	NR	NR	NR	NR	NR
Speroff et al. 2008	Desvenlafaxine 100 mg	157	NR	NR	NR	NR	NR	NR
	Desvenlafaxine 150 mg	163	NR	NR	NR	NR	NR	NR
	Desvenlafaxine 200 mg	155	NR	NR	NR	NR	NR	NR
	Placebo	78	NR	NR	NR	NR	NR	NR
Wyrwich et al. 2008 (Secondary analysis of Speroff et al. 2008)	Desvenlafaxine 100 mg	145	NR	NR	NR	NR	NR	NR
	Desvenlafaxine 150 mg	137	NR	NR	NR	NR	NR	NR
	Desvenlafaxine 200 mg	120	NR	NR	NR	NR	NR	NR
	Placebo	77	NR	NR	NR	NR	NR	NR
(12-week) Archer et al. 2009	Desvenlafaxine 150 mg	151	NR	NR	NR	NR	NR	NR
	Desvenlafaxine 100 mg	150	NR	NR	NR	NR	NR	NR
	Placebo	150	NR	NR	NR	NR	NR	NR
(26-weeks) Archer et al. 2009	Desvenlafaxine 100 mg	182	NR	NR	NR	NR	NR	NR
	Desvenlafaxine 150 mg	179	NR	NR	NR	NR	NR	NR
	Placebo	180	NR	NR	NR	NR	NR	NR
Boekhout, A.H. 2011	Venlafaxine 75 mg	41	NR	NR	NR	NR	NR	NR
	Placebo	20	NR	NR	NR	NR	NR	NR
Bouchard et al. 2012	Desvenlafaxine 100 mg	137	NR	NR	-8.03 (0.80), p value vs. placebo=0.169	NR	NR	-17.75 (2.84), p value vs. placebo=0.216
	Placebo	150	NR	NR	-6.90 (0.75)	NR	NR	-13.51 (2.60)
Pinkerton et al. 2013	Desvenlafaxine 100 mg	1066	58.1% "much improved or very much improved"	NR	-10.64 (0.23)	NR	NR	NR
	Placebo	1052	38.6%; "much improved or very much	NR		NR	NR	NR

			improved" P < 0.001)					
SSRIs[43,68,69,85,86,89,125]								
Stearns et al. 2003	Paroxetine 12.5 mg CR	51	NR	NR	-1.75 (0.24), Diff from PBO: -0.9; 95% CI: (-1.6 to -0.3); p=0.005	−0.83 (0.54); DIFF from PBO: −0.9; 95% CI: (-2.3 to 0.5), p=0.22	NR	NR
	Paroxetine 25 mg CR	58	NR	NR	-1.55 (0.23), Diff from PBO: -0.7 95% CI: (-1.3 to -0.1); p=0.02	0.01 (0.52); DIFF from PBO: -0.1; 95% CI: (-1.5 to 1.4), p=0.94	NR	NR
	Placebo	56	NR	NR	REF	0.06 (0.50)	NR	NR
Kalay et al. 2007	Citalopram 20 mg	25	NR	NR	NR	NR	Reported by subdomain only	NR
	Placebo	25	NR	NR	NR	NR	Reported by subdomain only	NR
FAST Trial (Grady et al., 2007)	Sertraline 100 mg	50	NR	NR	0.7 (10.4)	NR	NR	NR
	Placebo	49	NR	NR	–1.0 (5)	NR	NR	NR
MsFLASH 01 (Freeman et al., 2011)	Escitalopram 10 mg	104	NR	NR	NR	NR	NR	NR
	Placebo	101	NR	NR	NR	NR	NR	NR

Simon et al. 2011 (Study 1)	Paroxetine 7.5 mg	301	NR	NR	NR	NR	NR	NR
	Placebo	305	NR	NR	NR	NR	NR	NR
Simon et al. 2011 (Study 2)	Paroxetine 7.5 mg	284	NR	NR	NR	NR	NR	NR
	Placebo	284	NR	NR	NR	NR	NR	NR
Simon et al. 2011 (Study 1 & 2) Sleep substudy (Pinkerton et al. 2014)	Paroxetine 7.5 mg	585	NR	-3.17 (NR)	Difficulty sleeping decreased from baseline in both treatment arms.	NR	NR	NR
	Placebo	589	NR	-2.66 (NR)	NR	NR	NR	NR
Capriglione et al. 2016	Paroxetine 7.5 mg	42	NR	-42%, p=0.066	Difficulty sleeping item: proportion of participants reporting moderate to severe difficulty sleeping decreased from baseline in both treatment arms.	NR	NR	NR
	Placebo	38	NR	-35%	NR	NR	NR	NR

Gabapentin[52,53,72,126]								
Guttuso et al. 2003	Gabapentin 900 mg	30	NR	NR	NR	NR	NR	NR
	Placebo	29	NR	NR	NR	NR	NR	NR
Reddy et al. 2006	Gabapentin 2400 mg	20	NR	NR	NR for total score	NR	NR	NR
	Placebo	20	NR	NR	REF	NR	NR	NR
Butt et al. 2008	Gabapentin 900 mg	99	NR	NR	NR	NR	-0.8; 95% CI: (-1.0 to -0.6), p=0.004	NR
	Placebo	98	NR	NR	NR	NR	-0.4 95% CI: (-0.6 to -0.2)	NR
Breeze 3 (Pinkerton et al., 2014)	Gabapentin 1800* mg	299	NR	NR	NR	NR	NR	NR
	Placebo	294	NR	NR	NR	NR	NR	NR

CI: confidence interval, DIFF: difference, DSRP: drospirenone, E2: estradiol, GCS: Greene climacteric scale, HFRDIS: hot flash related daily interference scale, LSM: least square mean, LSMD: least square mean difference, MENQoL: the menopause-specific quality of life questionnaire, MHT: menopausal hormonal therapy, mg: milligrams, CR: controlled release, N: total number of participants, NETA: norethindrone acetate, NR: not reported, DIFF FROM PBO: Difference From Placebo, PBO: placebo, POMS: profile of mood states, SD: standard deviation, SDS: sheehan disability scale, SNRI: serotonin and norepinephrine reuptake inhibitors, SSRI: selective serotonin reuptake inhibitors

*gastroretentive gabapentin

Table D3.8. Safety I

Title	Intervention	Arm Size	Any AE % (n/N)	Serious Adverse events % (n/N)	All-cause mortality % (n/N)	Discontinuation due to AE % (n/N)	Fractures % (n/N)	Stroke % (n/N)
Fezolinetant[54-57,99,108,115]								
Phase 2A	Fezolinetant 90 mg	43	29 / 43 (67.4%)	0 / 43 (0.0%)	0 / 43 (0.0%)	2 / 43 (4.7%)	NR	NR
	Placebo	44	35 / 44 (79.5%)	1 / 44 (2.3%)	0 / 44 (0.0%)	0 / 44 (0.0%)	NR	NR
VESTA	Fezolinetant 30 mg	43	23 / 43 (53.5%)	0 / 43 (0.0%)	0 / 43 (0.0%)	2 / 43 (4.7%)	NR	NR

	Placebo	43	21 / 43 (48.8%)	0 / 43 (0.0%)	0 / 43 (0.0%)	1 / 43 (2.3%)	NR	NR
SKYLIGHT 1 abstract	Fezolinetant 30 mg	173	65 / 173 (37.6%)	NR	NR	NR	NR	NR
	Fezolinetant 45 mg	174	76 / 174 (43.7%)	NR	NR	NR	NR	NR
	Placebo	175	78 / 175 (44.6%)	NR	NR	NR	NR	NR
SKYLIGHT 2	Fezolinetant 30 mg	166	67/166 (40.4%)	3/166 (1.8%)	NR	2/166 (1.2%)	NR	NR
	Fezolinetant 45 mg	167	60/167 (35.9%)	2/167 (1.2%)	NR	5/167 (3.0%)	NR	NR
	Placebo	167	54/167 (32.3%)	0/167 (0%)	NR	1/167 (0.6%)	NR	NR
	Fezolinetant 30 mg	166	107 / 166 (64.5%)†	9 / 166 (5.4%)†	0 / 166 (0.0%)†	4 / 166 (2.4%)†	NR	NR
	Fezolinetant 45 mg	167	106 / 167 (63.5%)†	8 / 167 (4.8%)†	0 / 167 (0.0%)†	7 / 167 (4.2%)†	NR	NR
	PBO/FEZ 30 mg	76	43 / 76 (56.6%)†	2 / 76 (2.6%)†	0 / 76 (0.0%)†	2 / 76 (2.6%)†	NR	NR
	PBO/FEZ 45 mg	75	45 / 75 (60.0%)†	4 / 75 (5.3%)†	1 / 75 (1.3%)†	3 / 75 (4.0%)†	NR	NR
MHT: standard dose (1 mg estradiol) [33,34,36,37]								
Schurmann et al. 2004	Estradiol 1 mg/DSRP 1 mg	55	23 / 54 (42.6%)	0 / 55 (0.0%)	NR	3 / 55 (5.5%)	NR	NR
	Estradiol 1 mg/DSRP 2 mg	52	14 / 52 (26.9%)	1 / 52 (1.9%)	NR	3 / 52 (5.8%)	NR	NR
	Estradiol 1 mg/DSRP 3 mg	57	15 / 56 (26.8%)	3 / 57 (5.3%)	NR	2 / 57 (3.5%)	NR	NR
	Placebo	61	10 / 61 (16.4%)	1 / 61 (1.6%)	NR	3 / 61 (4.9%)	NR	NR
Endrikat et al. 2007	1 mg estradiol valerate/2 mg dienogest	162	39 / 162 (24.1%)	0 / 162 (0.0%)	0 / 162 (0.0%)	1 / 162 (0.6%)	NR	NR
	Placebo	162	43 / 162 (26.5%)	2 / 162 (1.2%)	0 / 162 (0.0%)	4 / 162 (2.5%)	NR	NR

Lin et al. 2011	Estradiol 1 mg/DSRP 2 mg	183	55 / 183 (30.1%)	NR	NR	7 / 183 (3.8%)	NR	NR
	Placebo	61	16 / 61 (26.2%)	NR	NR	5 / 61 (8.2%)	NR	NR
Lobo et al. 2018	Estradiol 1 mg and progesterone 100 mg	415	297 / 415 (71.6%)	9 / 415 (2.2%)	0 / 415 (0.0%)	45 / 415 (10.8%)	NR	NR
	Estradiol 0.5 mg and progesterone 100 mg	424	302 / 424 (71.2%)	13 / 424 (3.1%)	0 / 424 (0.0%)	31 / 424 (7.3%)	NR	NR
	Estradiol 0.5 mg and progesterone 50 mg	421	293 / 421 (69.6%)	8 / 421 (1.9%)	1 / 421 (0.2%)	34 / 421 (8.1%)	NR	NR
	Placebo	151	78 / 151 (51.7%)	2 / 151 (1.3%)	0 / 151 (0.0%)	10 / 151 (6.6%)	NR	NR
MHT: low dose (0.5 mg estradiol or equivalent) [35,38-41]								
Panay et al. 2007	Estradiol 0.5 mg and NETA 1 mg	194	NR	NR	NR	11 / 194 (5.7%)	NR	NR
	Estradiol 0.5 mg and NETA 0.25 mg	181	NR	NR	NR	4 / 181 (2.2%)	NR	NR
	Placebo	200	NR	NR	NR	16 / 200 (8.0%)	NR	NR
Stevenson et al. 2010	Estradiol 1 mg/DYD 5 mg	60	22 / 60 (36.7%)	1 / 60 (1.7%)	NR	2 / 60 (3.3%)	NR	NR
	Estradiol 0.5 mg/DYD 2.5 mg	122	78 / 122 (63.9%)	5 / 122 (4.1%)	NR	6 / 122 (4.9%)	NR	NR
	Placebo	125	25 / 125 (20%)	1 / 125 (0.8%)	NR	4 / 125 (3.2%)	NR	NR
Archer et al. 2013	Estradiol 0.5 mg/DRSP 0.25 mg	177	101 / 183 (55.2%)	1 / 183 (0.5%)	NR	4 / 183 (2.2%)	NR	NR
	Estradiol 0.5 mg /DRSP 0.5 mg	178	112 / 181 (61.9%)	0 / 181 (0.0%)	NR	7 / 181 (3.9%)	NR	NR
	Placebo	176	101 / 180 (56.1%)	1 / 180 (0.6%)	NR	4 / 180 (2.2%)	NR	NR
MsFLASH 03 (Joffe et al., 2014)	Estrogen only	97	53 / 97 (54.6%)	NR	NR	4 / 97 (4.1%)	NR	NR
	Venlafaxine 75 mg	96	65 / 96 (67.7%)	NR	NR	5 / 96 (5.2%)	NR	NR

	Placebo	146	88 / 146 (60.3%)	NR	NR	2 / 146 (1.4%)	NR	NR
Speroff et al. 1996	Estradiol transdermal system: 0.02 mg	54	NR	0 / 54 (0.0%)	NR	0 / 54 (0%)	NR	NR
	Estradiol transdermal system: 0.04 mg	53	NR	0 / 53 (0.0%)	NR	1 / 53 (2%)	NR	NR
	Placebo single dose	54	NR	0 / 54 (0.0%)	NR	1 / 54 (2%)	NR	NR
	Placebo double dose	52	NR	0 / 52 (0.0%)	NR	3 / 52 (6%)	NR	NR
	Estradiol transdermal system: 0.02 mg	37	NR	0 / 37 (0.0%)	NR	2 / 37 (5%)	NR	NR
	Estradiol transdermal system: 0.04 mg	37	NR	0 / 37 (0.0%)	NR	2 / 37 (5%)	NR	NR
	Placebo double dose	37	NR	0 / 37 (0.0%)	NR	0 / 37 (0%)	NR	NR
SNRIs[44,46-48,50,64-66,84]								
Evans et al. 2005	Venlafaxine 75 mg	40	NR	NR	NR	NR	NR	NR
	Placebo	40	NR	NR	NR	NR	NR	NR
Speroff et al. 2008	Desvenlafaxine 100 mg	157	146 / 155 (94.2%)	NR	NR	33 / 157 (21.0%)	NR	NR
	Desvenlafaxine 150 mg	163	149 / 157 (94.9%)	NR	NR	58 / 163 (35.6%)	NR	NR
	Desvenlafaxine 200 mg	155	147 / 151 (97.4%)	NR	NR	63 / 155 (40.6%)	NR	NR
	Placebo	78	67 / 77 (87.0%)	NR	NR	12 / 78 (15.4%)	NR	NR
Wyrwich et al. 2008 (Secondary analysis of Speroff et al. 2008)	Desvenlafaxine 100 mg	145	NR	NR	NR	NR	NR	NR
	Desvenlafaxine 150 mg	137	NR	NR	NR	91 / 137 (66.4%)	NR	NR
	Desvenlafaxine 200 mg	120	NR	NR	NR	NR	NR	NR

	Placebo	77	NR	NR	NR	10 / 77 (13.0%)	NR	NR
(12-week) Archer et al. 2009	Desvenlafaxine 150 mg	151	128 / 151 (84.8%)	1 / 151 (0.7%)	NR	NR	NR	NR
	Desvenlafaxine 100 mg	150	125 / 150 (83.3%)	0 / 150 (0.0%)	NR	NR	NR	NR
	Placebo	150	105 / 150 (70.0%)	1 / 150 (0.7%)	NR	NR	NR	NR
(26-weeks) Archer et al. 2009	Desvenlafaxine 100 mg	182	173 / 182 (95.1%)	NR	NR	49 / 182 (26.9%)	NR	NR
	Desvenlafaxine 150 mg	179	162 / 179 (90.5%)	NR	NR	54 / 179 (30.2%)	NR	NR
	Placebo	180	159 / 180 (88.3%)	NR	NR	16 / 180 (8.9%)	NR	NR
Boekhout et al. 2011	Venlafaxine 75 mg	41	NR	NR	NR	2 / 41 (4.9%)	NR	NR
	Placebo	20	NR	NR	NR	1 / 20 (5.0%)	NR	NR
Bouchard et al. 2012	Desvenlafaxine 100 mg	137	116 / 158 (73.4%)	2 / 137 (1.5%)	0 / 137 (0.0%)	35 / 137 (25.5%)	NR	NR
	Placebo	150	85 / 152 (55.9%)	4 / 150 (2.7%)	0 / 150 (0.0%)	14 / 150 (9.3%)	NR	NR
Pinkerton et al. 2013	Desvenlafaxine 100 mg	1066	893 / 1066 (84.0%)	43 / 1066 (4.0%)	0 / 1066 (0.0%)	195 / 1066 (18.3%)	NR	1 / 1066 (0.1%)
	Placebo	1052	832 / 1052 (79.0%)	36 / 1052 (3.4%)	1 / 1052 (0.1%)	102 / 1052 (9.7%)	NR	0 / 1052 (0.0%)‡
SSRIs[43,68,69,85,86,125]								
Stearns et al. 2003	Paroxetine 12.5 mg CR	51	NR	NR	NR	NR	NR	NR
	Paroxetine 25 mg CR	58	NR	NR	NR	NR	NR	NR
	Placebo	56	NR	NR	NR	NR	NR	NR
Kalay et al. 2007	Citalopram 20 mg	25	8 / 25 (32.0%)	NR	NR	NR	NR	NR

	Placebo	25	NR	NR	NR	NR	NR	NR
FAST Trial (Grady et al., 2007)	Sertraline 100 mg	50	NR	NR	NR	NR	NR	NR
	Placebo	49	NR	NR	NR	NR	NR	NR
MsFLASH 01 (Freeman et al., 2011)	Escitalopram 10 mg	104	55 / 104 (52.9%)	0 / 104 (0.0%)	NR	7 / 104 (6.7%)	NR	NR
	Placebo	101	64 / 101 (63.4%)	0 / 101 (0.0%)	NR	2 / 101 (2.0%)	NR	NR
Simon et al. 2011 (Study 1)	Paroxetine 7.5 mg	301	295 / 586 (50.3%)	14 / 301 (4.7%)	1 / 301 (0.3%)	8 / 306 (2.6%)	1 / 301 (0.3%)	NR
	Placebo	305	275 / 589 (46.7%)	8 / 305 (2.6%)	0 / 305 (0.0%)	4 / 308 (1.3%)	3 / 305 (1.0%)	NR
Simon et al. 2011 (Study 2)	Paroxetine 7.5 mg	284	295 / 586 (50.3%)	14 / 586 (2.4%)	1 / 586 (0.2%)	15 / 285 (5.3%)	1 / 284 (0.4%)	NR
	Placebo	284	275 / 589 (46.7%)	8 / 589 (1.4%)	0 / 589 (0.0%)	15 / 284 (5.3%)	3 / 284 (1.1%)	NR
Simon et al. 2011 (Study 1 & 2) Sleep substudy (Pinkerton et al., 2015)	Paroxetine 7.5 mg	585	18 / 586 (3.1%)	NR	NR	23 / 591 (3.9%)	NR	NR
	Placebo	589	18 / 589 (3.1%)	NR	NR	19 / 593 (3.2%)	NR	NR
Capriglione et al. 2016	Paroxetine 7.5 mg	42	NR	0 / 42 (0.0%)	NR	NR	NR	NR
	Placebo	38	NR	0 / 38 (0.0%)	NR	NR	NR	NR
Gabapentin[52,53,71,72,126]								
Guttuso et al. 2003	Gabapentin 900 mg	30	15 / 30 (50.0%)	NR	NR	NR	NR	NR
	Placebo	29	8 / 29 (27.6%)	NR	NR	NR	NR	NR
Reddy et al. 2006	Gabapentin 2400 mg	20	NR	NR	NR	1 / 20 (5.0%)	NR	NR
	Placebo	20	NR	NR	NR	0 / 20 (0.0%)	NR	NR

Butt et al. 2008	Gabapentin 900 mg	99	2 / 95 (2.1%)	0 / 95 (0.0%)	NR	13 / 99 (13.1%)	NR	NR
	Placebo	98	7 / 98 (7.1%)	0 / 98 (0.0%)	NR	6 / 98 (6.1%)	NR	NR
Yurcheshen et al. 2009 (Secondary data analysis of Guttuso et al., 2003)	Gabapentin 900 mg	30	NR	NR	NR	4 / 30 (13.3%)	NR	NR
	Placebo	29	NR	NR	NR	1 / 29 (3.4%)	NR	NR
Breeze 3 (Pinkerton et al., 2014)	Gabapentin 1800* mg	299	197 / 300 (65.7%)	4 / 300 (1.3%)	0 / 299 (0.0%)	50 / 300 (16.7%)	NR	NR
	Placebo	294	165 / 295 (55.9%)	7 / 295 (2.4%)	1 / 294 (0.3%)	34 / 295 (11.5%)	NR	NR

Data on colorectal cancer was not reported by any trial and does not appear in the table. AE: adverse event, DSRP: drospirenone, E2: estradiol, MHT: menopausal hormonal therapy, mg: milligrams, CR: controlled release, N: total number of participants, NETA: norethindrone acetate, NR: not reported, PBO: placebo, SNRI: serotonin and norepinephrine reuptake inhibitors, SSRI: selective serotonin reuptake inhibitors

*gastroretentive gabapentin

†52-week data

‡Data from secondary analysis of safety outcomes in Archer et al. (2013)[66]

Table D3.9. Safety II

Title	Intervention	Arm Size	Active liver disease % (n/N)	Breast cancer % (n/N)	Cardiovascular events	Elevated ALT % 3X ULN (n/N)	Elevated AST % 3x ULN (n/N)	Endometrial hyperplasia % (n/N)
Fezolinetant[54-57,99,108,115]								
Phase 2A	Fezolinetant 90 mg	43	NR	NR	NR	5 / 43 (11.6%)	2 / 43 (4.7%)	NR
	Placebo	44	NR	NR	NR	1 / 44 (2.3%)	4 / 44 (9.1%)	NR
VESTA	Fezolinetant 30 mg	43	0 / 43 (0.0%)	NR	NR	0 / 43 (0.0%)	0 / 43 (0.0%)	NR
	Placebo	43	NR	NR	NR	0 / 43 (0.0%)	0 / 43 (0.0%)	NR

SKYLIGHT 1	Fezolinetant 30 mg	173	NR	NR	NR	NR	NR	NR
	Fezolinetant 45 mg	174	NR	NR	NR	NR	NR	NR
	Placebo	175	NR	NR	NR	NR	NR	NR
SKYLIGHT 2	Fezolinetant 30 mg	166	NR	NR	NR	3 / 164 (1.8%)†	1 / 164 (0.6%)†	NR
	Fezolinetant 45 mg	167	NR	NR	NR	7 / 164 (4.3%)†	1 / 164 (0.6%)†	NR
	PBO/FEZ 30 mg	76	NR	NR	NR	0 / 76 (0.0%)†	0 / 76 (0.0%)†	NR
	PBO/FEZ 45 mg	75	NR	NR	NR	2 / 74 (2.7%)†	0 / 74 (0.0%)†	NR
MHT: standard dose (1 mg estradiol)[33,34,36,37,61]								
Schurmann et al. 2004	Estradiol 1 mg/DSRP 1 mg	55	NR	NR	NR	NR	NR	1 / 54 (1.9%)
	Estradiol 1 mg/DSRP 2 mg	52	NR	NR	NR	NR	NR	0 / 52 (0.0%)
	Estradiol 1 mg/DSRP 3 mg	57	NR	NR	NR	NR	NR	1 / 56 (1.8%)
	Placebo	61	NR	NR	NR	NR	NR	0 / 61 (0.0%)
Endrikat et al. 2007	1 mg estradiol valerate/2 mg dienogest	162	NR	NR	NR	2 / 162 (1.2%)	NR	NR
	Placebo	162	NR	NR	NR	0 / 162 (0.0%)	NR	NR
Lin et al. 2011	Estradiol 1 mg/DSRP 2 mg	183	NR	NR	NR	NR	NR	0 / 183 (0.0%)
	Placebo	61	NR	NR	NR	NR	NR	0 / 61 (0.0%)
Lobo et al. 2018	Estradiol 1 mg and progesterone 100 mg	415	NR	2 / 415 (0.5%)	4 / 1411 (0.3%)	NR	NR	0 / 415 (0.0%)
	Estradiol 0.5 mg and progesterone 100 mg	424	NR	2 / 424 (0.5%)		NR	NR	0 / 424 (0.0%)
	Estradiol 0.5 mg and progesterone 50 mg	421	NR	1 / 421 (0.2%)		NR	NR	0 / 421 (0.0%)

	Placebo	151	NR	0 / 151 (0.0%)		NR	NR	0 / 151 (0.0%)
Mirkin et al. 2020 (Secondary data analysis of Lobo et al., 2018)	Estradiol 1 mg and progesterone 100 mg	280	NR	NR	NR	NR	NR	1 / 280 (0.4%)
	Estradiol 0.5 mg and progesterone 100 mg	303	NR	NR	NR	NR	NR	0 / 303 (0.0%)
	Estradiol 0.5 mg and progesterone 50 mg	306	NR	NR	NR	NR	NR	0 / 306 (0.0%)
	Placebo	274	NR	NR	NR	NR	NR	0 / 274 (0.0%)
MHT: low dose (0.5 mg estradiol or equivalent)[35,38-41]								
Panay et al. 2007	Estradiol 0.5 mg and NETA 1 mg	194	NR	NR	NR	NR	NR	NR
	Estradiol 0.5 mg and NETA 0.25 mg	181	NR	NR	NR	NR	NR	NR
	Placebo	200	NR	NR	NR	NR	NR	NR
Stevenson et al. 2010	Estradiol 1 mg/DYD 5 mg	60	NR	NR	NR	NR	NR	NR
	Estradiol 0.5 mg/DYD 2.5 mg	122	NR	NR	NR	NR	NR	NR
	Placebo	125	NR	NR	NR	NR	NR	NR
Archer et al. 2013	Estradiol 0.5 mg/DRSP 0.25 mg	177	NR	NR	NR	NR	NR	NR
	Estradiol 0.5 mg /DRSP 0.5 mg	178	NR	NR	NR	NR	NR	NR
	Placebo	176	NR	NR	NR	NR	NR	NR
MsFLASH 03 (Joffe et al., 2014)	Estrogen only	97	NR	NR	NR	NR	NR	NR
	Venlafaxine 75 mg	96	NR	NR	NR	NR	NR	NR
	Placebo	146	NR	NR	NR	NR	NR	NR

Speroff et al. 1996	Estradiol transdermal system: 0.02 mg	54	NR	NR	NR	NR	NR	NR
	Estradiol transdermal system: 0.04 mg	53	NR	NR	NR	NR	NR	NR
	Placebo single dose	54	NR	NR	NR	NR	NR	NR
	Placebo double dose	52	NR	NR	NR	NR	NR	NR
	Estradiol transdermal system: 0.02 mg	37	NR	NR	NR	NR	NR	NR
	Estradiol transdermal system: 0.04 mg	37	NR	NR	NR	NR	NR	NR
	Placebo double dose	37	NR	NR	NR	NR	NR	NR
SNRIs[44,46-48,50,64-66,84]								
Evans et al. 2005	Venlafaxine 75 mg	40	NR	NR	NR	NR	NR	NR
	Placebo	40	NR	NR	NR	NR	NR	NR
Speroff et al. 2008	Desvenlafaxine 100 mg	157	NR	NR	5 / 475 (1%)	NR	NR	NR
	Desvenlafaxine 150 mg	163	NR	NR		NR	NR	NR
	Desvenlafaxine 200 mg	155	NR	NR		NR	NR	NR
	Placebo	78	NR	NR	0 / 78 (0%)	NR	NR	NR
Wyrwich et al. 2008 (Secondary analysis of Speroff et al. 2008)	Desvenlafaxine 100 mg	145	NR	NR	5 / 402 (1.2%)	NR	NR	NR
	Desvenlafaxine 150 mg	137	NR	NR		NR	NR	NR
	Desvenlafaxine 200 mg	120	NR	NR		NR	NR	NR
	Placebo	77	NR	NR	0 / 77 (0%)	NR	NR	NR

(12-week) Archer et al. 2009	Desvenlafaxine 150 mg	151	NR	NR	NR	NR	NR	NR
	Desvenlafaxine 100 mg	150	NR	NR	NR	NR	NR	NR
	Placebo	150	NR	NR	NR	NR	NR	NR
(26-weeks) Archer et al. 2009	Desvenlafaxine 100 mg	182	NR	NR	NR	NR	NR	NR
	Desvenlafaxine 150 mg	179	NR	NR	NR	NR	NR	NR
	Placebo	180	NR	NR	NR	NR	NR	NR
Boekhout et al. 2011	Venlafaxine 75 mg	41	NR	NR	NR	NR	NR	NR
	Placebo	20	NR	NR	NR	NR	NR	NR
Bouchard et al. 2012	Desvenlafaxine 100 mg	137	NR	NR	2 / 137 (1.5%)	NR	NR	NR
	Placebo	150	NR	NR	4 / 150 (2.7%)	NR	NR	NR
Archer et al. 2013	Desvenlafaxine 100 mg	1066	NR	NR	0 / 1066 (0%)	2 / 1066 (0.2%)	NR	NR
	Placebo	1052	NR	NR	1 / 1052 (0.1%)	2 / 1052 (0.2%)	NR	NR
Pinkerton et al. 2013	Desvenlafaxine 100 mg	1066	NR	NR	NR	NR	NR	NR
	Placebo	1052	NR	NR	NR	NR	NR	NR
SSRIs[43,68,85,86,89,125]								
Stearns et al. 2003	Paroxetine 12.5 mg CR	51	NR	NR	NR	NR	NR	NR
	Paroxetine 25 mg CR	58	NR	NR	NR	NR	NR	NR
	Placebo	56	NR	NR	NR	NR	NR	NR

Kalay et al. 2007	Citalopram 20 mg	25	NR	NR	NR	NR	NR	NR
	Placebo	25	NR	NR	NR	NR	NR	NR
FAST Trial (Grady et al., 2007)	Sertraline 100 mg	50	NR	NR	NR	NR	NR	NR
	Placebo	49	NR	NR	NR	NR	NR	NR
MsFLASH 01 (Freeman et al., 2011)	Escitalopram 10 mg	104	NR	NR	NR	NR	NR	NR
	Placebo	101	NR	NR	NR	NR	NR	NR
Simon et al. 2011 (Study 1)	Paroxetine 7.5 mg	301	NR	NR	NR	NR	NR	NR
	Placebo	305	NR	NR	NR	NR	NR	NR
Simon et al. 2011 (Study 2)	Paroxetine 7.5 mg	284	NR	NR	NR	NR	NR	NR
	Placebo	284	NR	NR	NR	NR	NR	NR
Capriglione et al. 2016	Paroxetine 7.5 mg	42	NR	NR	NR	NR	NR	NR
	Placebo	38	NR	NR	NR	NR	NR	NR

Gabapentin[52,53,72,126]								
Guttuso et al. 2003	Gabapentin 900 mg	30	NR	NR	NR	NR	NR	NR
	Placebo	29	NR	NR	NR	NR	NR	NR
Reddy et al. 2006	Gabapentin 2400 mg	20	NR	NR	NR	NR	NR	NR
	Placebo	20	NR	NR	NR	NR	NR	NR
Butt et al. 2008	Gabapentin 900 mg	99	NR	NR	NR	NR	NR	NR
	Placebo	98	NR	NR	NR	NR	NR	NR
Breeze 3 (Pinkerton et al., 2014)	Gabapentin 1800* mg	299	NR	NR	NR	NR	NR	NR
	Placebo	294	NR	NR	NR	NR	NR	NR

AE: adverse event, ALT: alanine transaminase, AST: aspartate transaminase, DSRP: drospirenone, E2: estradiol, MHT: menopausal hormonal therapy, mg: milligrams, mg: milligrams, CR: controlled release, N: total number of participants, NETA: norethindrone acetate, NR: not reported, PBO: placebo, SNRI: serotonin and norepinephrine reuptake inhibitors, SSRI: selective serotonin reuptake inhibitors, ULN: upper limit of the normal range

*gastroretentive gabapentin

†52-week data

Table D3.10. Long Term Outcomes (MHT)[80-82]

Main Study	Outcomes by drug class	Mean follow-up	Risk Ratio / Relative effect (95%CI)	N of participants
Marjoribanks et al. 2017 (Meta-analysis) **N of RCTs: 22** **N of participants: 43,637** **Age: 48-76 years (26-91 years)**	**Estrogen-only**			
	Coronary event (MI or cardiac death)	7.1 years	0.94 (0.78 to 1.13)	10739
	Stroke	7.1 years	1.33 (1.06, 1.67)	10739
	Venous thromboembolism (DVT or PE)	7.1 years	1.32 (1.00 per 1.74)	10739
	Breast cancer	7.1 years	0.79 (0.61 to 1.01)	10739
	Gallbladder disease	7.1 years	1.78 (1.42 to 2.24)	8376
	All clinical fractures	7.1 years	0.73 (0.65 to 0.80)	10739
	Death all-cause (low dose)	2 years	0.33 (0.01 to 8.10)	222
	Death all-cause (moderate dose)	7.1 years	1.03 (0.88 to 1.20)	10739
	Death all-cause (moderate dose)	10.7 years	1.02 (0.91 to 1.13)	10739
	Combined estrogen/progesterone			
	Coronary event (MI or cardiac death)	1 year	1.89 (1.15 to 3.10)	20993
	Stroke	3 years	1.46 (1.02 to 2.09)	17585
	Venous thromboembolism (DVT or PE)	5.6 years	2.03 (1.55, 2.64)	16608
	Breast cancer	5.6 years*	1.27 (1.03 to 1.56)	16608
	Death from lung cancer	8 years*	1.74 (1.18 to 2.55)	16608
	Gallbladder disease	5.6 years	1.64 (1.30 to 2.06)	14203
	All clinical fractures	5.6 years	0.78 (0.71 to 0.86)	16608
	Death all-cause (low dose)	4 years	3.58 (0.15 to 87.57)	505
	Death all-cause (moderate dose)	7.9 years	1.06 (0.93 to 1.20)	16608
	Death all-cause (moderate dose)	13.2 years	1.00 (0.92 to 1.08)	16608
	Death from any cancer (moderate dose)	5.2 years	1.16 (0.87 to 1.53)	16608

Main Study	Outcomes by study design	Median follow-up	Summary estimates (95% CI)†	N of trials
Kim et al. 2020 (Meta-analysis) **N of RCTs: 26** **N of observational studies: 47** **N of total studies: 73** **Median age for RCTs: 63.6 (49.7-75.0)** **Median age for observational studies: 60.6 (48.8-77.0)**	**RCTs**			
	All-cause death	3.4 (0.7-1.8)	1.00 (0.96 to 1.04)	17
	Cardiovascular death		0.96 (0.83 to 1.12)	11
	Stroke		1.14 (1.04 to 1.25)	13
	Venous thromboembolism		1.70 (1.33 to 2.16)	15
	Pulmonary embolism		1.26 (1.06 to 1.50)	8
	Myocardial infarction		1.04 (0.94 to 1.14)	17
	Coronary heart disease		1.02 (0.94 to 1.10)	5
	Angina		0.95 (0.84 to 1.08)	8
	Revascularization		0.96 (0.87 to 1.06)	7
	Observational studies			
	All-cause death	6.8 (1-21.5)	0.90 (0.79 to 1.02)	15
	Cardiovascular death		0.81 (0.61 to 1.07)	6
	Stroke		0.98 (0.85 to 1.13)	13
	Venous thromboembolism		1.32 (1.13 to 1.54)	12
	Pulmonary embolism		1.44 (1.17 to 1.76)	4
	Myocardial infarction		0.79 (0.75 to 0.84)	10
	Coronary heart disease		0.91 (0.72 to 1.15)	7
	Angina		1.11 (0.86 to 1.43)	1
	Revascularization	NR	NR	NR
	Subgroup for observational studies: Estrogen only			
	All-cause death	6.8 (1-21.5)	0.85 (0.77 to 0.95)	7
	Stroke		1.02 (0.90 to 1.16)	9
	Venous thromboembolism		0.93 (0.79 to 1.08)	8
	Myocardial infarction		0.85 (0.79 to 0.91)	9
	Subgroup for observational studies: Combined estrogen/progesterone			
	All-cause death	6.8 (1-21.5)	0.61 (0.34 to 1.09)	7
	Stroke		1.05 (0.81 to 1.35)	6
	Venous thromboembolism		2.21 (1.51 to 3.22)	6

	Myocardial infarction		0.77 (0.71 to 0.84)	8
Main Study	**Outcomes by drug class**	**Median follow-up**	**HR (95% CI)**	**N of trials**
Prentice et al. 2020 **N of RCTs: 2** **N of participants: 27,347** **Mean age (SD) for estrogen only trials: 54.9 (2.9)** **Mean age (SD) for combined estrogen/progesterone trial: 55.3 (2.6)**	**Estrogen-only trial**	**Median intervention phase**	**Parsimonious model**	NA
	Coronary heart disease	7.2 years	0.75 (0.55 to 1.02)	
	Invasive breast cancer		0.63 (0.45 to 0.88)	
	Stroke		1.08 (0.78 to 1.50)	
	Pulmonary embolism		1.09 (0.68 to 1.74)	
	Colorectal cancer		0.90 (0.59 to 1.36)	
	Hip fracture		0.48 (0.30 to 0.76)	
	Death (all-cause)		0.81 (0.61 to 1.08)	
	Estrogen-only trial	**Median cumulative follow-up**	**Parsimonious model**	
	Coronary heart disease	18 years	0.83 (0.70 to 1.00)	
	Invasive breast cancer		0.70 (0.57 to 0.86)	
	Stroke		0.89 (0.73 to 1.09)	
	Pulmonary embolism		0.92 (0.71 to 1.19)	
	Colorectal cancer		0.87 (0.66 to 1.16)	
	Hip fracture		0.76 (0.60 to 0.97)	
	Death (all-cause)		0.85 (0.73 to 1.00)	
	Combined estrogen/progesterone trial	**Median intervention phase**	**Parsimonious model**	
	Coronary heart disease	5.6 years	1.11 (0.83 to 1.49)	
	Invasive breast cancer		1.18 (0.89 to 1.56)	
	Stroke		1.29 (0.93 to 1.78)	
	Pulmonary embolism		1.91 (1.25 to 2.90)	
	Colorectal cancer		0.58 (0.38 to 0.89)	
	Hip fracture		0.61 (0.40 to 0.94)	
	Death (all-cause)		0.93 (0.70 to 1.22)	

	Combined estrogen/progesterone trial	**Median cumulative follow-up**	**Parsimonious model**	
	Coronary heart disease	18 years	1.07 (0.92 to 1.24)	
	Invasive breast cancer		1.29 (1.11 to 1.50)	
	Stroke		1.13 (0.96 to 1.33)	
	Pulmonary embolism		1.14 (0.92 to 1.41)	
	Colorectal cancer		0.86 (0.69 to 1.08)	
	Hip fracture		0.90 (0.75 to 1.07)	
	Death (all-cause)		1.00 (0.89 to 1.14)	
Rossouw et al. 2007 **N of participants: 27,347**	**Combined trials**	**Median intervention phase**	**HR (95% CI)**	**N of participants**
	Coronary heart disease (50-59 years old)	5.6 years	0.93 (0.65-1.33)	8832
	Coronary heart disease (60-69 years old)		0.98 (0.79-1.21)	12362
	Coronary heart disease (70-79 years old)		1.26 (1.00-1.59)	6152

CI: confidence interval, DVT: deep venous thromboembolism, HR: hazard ratio, MHT: menopausal hormone therapy, MI: myocardial infarction, N: number, NA: not applicable, NR: not reported, PE: pulmonary embolism, RCT: randomized controlled trial

* Median follow-up

† Via fixed effects

Table D3.11. MENQoL in MsFlash

		Diem 2020[76]							Caan 2015[77]		
MENQoL change from baseline, Mean (95% CI)		**MsFLASH01**	**MsFLASH03**				**MENQoL change from baseline, Mean (95% CI)**	**MsFLASH 01**	**MsFLASH 03**		
		Escitalopram vs. placebo	**Estradiol vs. placebo**	**P-value (vs. placebo)**	**Venlafaxine vs. placebo**	**P-value (vs. placebo)**		**Placebo**	**Placebo**	**Estradiol**	**Venlafaxine**
MENQOL Total							**MENQoL total**				
	Baseline, mean (SD)	3.8 (1.3)					Baseline	--	3.5 (3.3, 3.7)	3.5 (3.2, 3.7)	3.7 (3.4, 3.9)
	Week 4	-0.4 (-0.7, -0.2)	-0.3 (-0.6, -0.1)	<.001	-0.2 (-0.5, -0.1)	0.042	Week 4 - baseline	-0.5 (-0.7, -0.3)	-0.6 (-0.7, -0.4)	-0.9 (-1.1, -0.7)	-0.9 (-1.1, -0.7)
	Week 8	-0.4 (-0.6, -0.1)	-0.5 (-0.7, -0.2)		-0.2 (-0.4, 0.0)		Week 8 - baseline	-0.7 (-0.9, -0.5)	-0.7 (-0.9, -0.5)	-1.1 (-1.3, -0.9)	-0.9 (-1.1, -0.7)
MENQOL Vasomotor							**MENQoL Vasomotor**				
	Baseline, mean (SD)	5.9 (1.7)					Baseline	--	5.6 (5.4, 5.9)	5.7 (5.4, 6.0)	5.9 (5.5, 6.2)
	Week 4	-0.4 (-0.9, 0.0)	-0.5 (-1.0, -0.1)	<.001	-0.3 (-0.7, 0.2)	0.211	Week 4 - baseline	-1.0 (-1.3, -0.7)	-0.8 (-1.1, -0.6)	0.7 (0.3, 1.1)	0.8 (0.3, 1.2)
	Week 8	-0.6 (-1.1, -0.2)	-1.2 (-1.7, -0.7)		-0.2 (-0.7, 0.2)		Week 8 - baseline	-1.0 (-1.4, -0.6)	-1.1 (-1.4, -0.8)	-0.1 (-0.5, 0.4)	0.7 (0.3, 1.1)
MENQOL Psychosocial							**MENQoL Psychosocial**				
	Baseline, mean (SD)	2.9 (1.5)					Baseline	--	2.7 (2.4, 2.9)	2.8 (2.4, 3.1)	2.9 (2.5, 3.2)
	Week 4	-0.4 (-0.7, -0.2)	-0.2 (-0.4, 0.0)	0.12	-0.3 (-0.6, -0.1)	0.008	Week 4 - baseline	-0.3 (-0.5, -0.1)	-0.3 (-0.5, -0.2)	-1.3 (-1.6, -1.1)	-1.5 (-1.8, -1.3)

	Week 8	-0.3 (-0.6, -0.1)	-0.1 (-0.4, 0.1)		-0.3 (-0.5, 0.0)		Week 8 - baseline	-0.6 (-0.88, -0.4)	-0.4 (-0.6, 0.3)	-1.4 (-1.7, -1.1)	-1.5 (-1.8, -1.3)
MENQOL Physical							**MENQoL Physical**				
	Baseline, mean (SD)	3.2 (1.3)					Baseline	--	3.0 (2.8, 3.2)	3.0 (2.7, 3.3)	3.2 (2.9, 3.5)
	Week 4	-0.5 (-0.7, -0.2)	-0.2 (-0.5, 0.0)	0.039	-0.2 (-0.5, 0.0)	0.082	Week 4 - baseline	-0.5 (-0.7, -0.3)	-0.5 (-0.6, -0.3)	-1.3 (-1.5, -1.0)	-1.3 (-1.5, -1.0)
	Week 8	-0.2 (-0.5, 0.0)	-0.2 (-0.3, 0.0)		-0.1 (-0.3, 0.1)		Week 8 - baseline	-0.7 (-0.9, -0.5)	-0.6 (-0.8, -0.5)	-1.3 (-1.6, -1.1)	-1.3 (-1.5, -1.1)
MENQOL Sexual							**MENQoL Sexual**				
	Baseline, mean (SD)	3.3 (2.5)					Baseline	--	3.0 (2.6, 3.4)	2.8 (2.3, 3.2)	3.0 (2.5, 3.4)
	Week 4	-0.3 (-0.7, 0.1)	-0.2 (-0.6, 0.1)	0.047	-0.1 (-0.4, 0.3)	0.447	Week 4 - baseline	-0.5 (-0.8, -0.1)	-0.5 (-0.8, -0.2)	-1.5 (-1.8, -1.1)	-1.5 (-1.8, -1.1)
	Week 8	-0.2 (-0.6 0.2)	-0.4 (-0.8, -0.1)		-0.2 (-0.5, 0.2)		Week 8 - baseline	-0.7 (-1.1, -0.4)	-0.6 (-0.9, -0.4)	-1.7 (-2.1, -1.4)	-1.6 (-1.9, -1.3)

CI: confidence interval, MENQoL: the menopause-specific quality of life questionnaire, SD: standard deviation

Table D3.12. SKYLIGHT Subgroups

Outcome	**VMS Frequency, LSM (SE)**				**VMS Severity, LSM (SE)**		
	Study	**SKYLIGHT Pooled (Nappi 2022)[55]**					
		FEZ 30 mg	**FEZ 45 mg**	**PBO**	**FEZ 30 mg**	**FEZ 45 mg**	**PBO**
	N	186	179	180	186	179	180
Subgroup	**Age**	< 55 years			< 55 years		
	n	169	169	166	169	169	166
Timepoint	**Week 4**	-5.45 (0.31)	-5.95 (0.31)	-3.56 (0.31)	-0.45 (0.04)	-0.54 (0.04)	-0.34 (0.04)
	DIFF vs. PBO	-1.89 (0.44)	-2.38 (0.44)	--	-0.11 (0.06)	-0.20 (0.06)	--

	P-value	<0.001	<0.001	--	0.061	0.001	--
	n	142	151	139	142	151	139
	Week 12	-6.62 (0.34)	-7.20 (0.34)	-4.42 (0.35)	-0.65 (0.05)	-0.67 (0.05)	-0.47 (0.06)
	DIFF vs. PBO	-2.20 (0.49)	-2.78 (0.49)	--	-0.18 (0.08)	-0.19 (0.08)	--
	P-value	<0.001	<0.001	--	0.023	0.013	--
	N	153	162	162	153	162	162
		>55 years			>55 years		
	n	143	150	151	143	150	151
	Week 4	-5.34 (0.34)	-5.63 (0.33)	-3.45 (0.33)	-0.43 (0.05)	-0.53 (0.04)	-0.25 (0.04)
	DIFF vs. PBO	-1.89 (0.47)	-2.18 (0.46)	--	-0.19 (0.06)	-0.28 (0.06)	--
	P-value	<0.001	<0.001	--	0.004	<0.001	--
	n	122	140	140	122	140	140
	Week 12	-6.53 (0.38)	-6.67 (0.36)	-4.43 (0.36)	-0.59 (0.06)	-0.67 (0.06)	-0.37 (0.06)
	DIFF vs. PBO	-2.10 (0.52)	-2.24 (0.51)	--	-0.21 (0.08)	-0.30 (0.08)	--
	P-value	<0.001	<0.001	--	0.009	<0.001	--
Subgroup	**VMS Severity**	<2.36			<2.36		
	N	163	176	172	163	176	172
	n	155	164	160	155	164	160
Timepoint	**Week 4**	-5.33 (0.29)	-6.07 (0.28)	-3.71 (0.28)	-0.40 (0.04)	-0.50 (0.04)	-0.25 (0.04)
	DIFF vs. PBO	-1.62 (0.41)	-2.36 (0.40)	--	-0.15 (0.06)	-0.25 (0.06)	--
	P-value	<0.001	<0.001	--	0.007	0.001	--
	n	135	152	144	135	152	144
	Week 12	-6.46 (0.32)	-6.90 (0.30)	-4.55 (0.31)	-0.53 (0.05)	-0.60 (0.05)	-0.31 (0.05)
	DIFF vs. PBO	-1.91 (0.44)	-2.36 (0.43)	--	-0.22 (0.07)	-0.30 (0.07)	--
	P-value	<0.001	<0.001	--	<0.001	<0.001	--
	N	176	165	170	176	165	170
		>2.36			>2.36		
	n	157	155	157	157	157	155
	Week 4	-5.53 (0.34)	-5.48 (0.35)	-3.25 (0.34)	-0.49 (0.05)	-0.57 (0.05)	-0.34 (0.05)
	DIFF vs. PBO	-2.28 (0.49)	-2.23 (0.49)	--	-0.15 (0.07)	-0.23 (0.07)	--

	P-value	<0.001	<0.001	--	0.028	<0.001	--
	n	131	140	136	131	140	136
	Week 12	-6.77 (0.39)	-6.91 (0.39)	-4.35 (0.39)	-0.71 (0.06)	-0.74 (0.06)	-0.54 (0.06)
	DIFF vs. PBO	-2.42 (0.55)	-2.56 (0.55)	--	-0.17 (0.09)	-0.19 (0.09)	--
	P-value	<0.001	<0.001	--	0.066	0.037	--

DIFF: difference, FEZ: fezolinetant, LSM: least square mean, N: number, PBO: placebo, SE: standard error, VMS: vasomotor symptoms,

Table D3.14. MHT Subgroups

Outcome	**VMS Frequency**		
	Study	**Tsiligiannis et al. 2020[63]**	
		Estradiol (0.5 mg)/dydrogesterone (2.5 mg)	**PBO**
Subgroup	**Age**	**45 to < 55 years**	
	N	76	74
	Baseline mean (SD)	7.6 (2.8)	7.3 (2.1)
	Change from baseline at week 13	-6.0 (3.3)	-4.6 (3.1)
	Difference of LS mean	1.04 (95% CI: 0.272, 1.80), P=0.008	
	Irregular intermenstrual bleeding between normal menstrual periods and abnormal uterine bleeding	8.6% and 9.30%	
	Age	**> 55 years**	
	N	46	50
	Baseline mean (SD)	8.6 (2.9)	8.3 (3.3)
	Change from baseline at week 13	-7.0 (3.8)	-5.4 (3.9)
	Difference of LS mean	1.43 (95% CI: 0.10, 2.76), P=0.036	
	Irregular intermenstrual bleeding between normal menstrual periods and abnormal uterine bleeding	2.7% and 2.3%	
	BMI	**< 25 kg/m^2**	
	N	47	50
	Baseline mean (SD)	8.4 (2.5)	8.2 (3.0)
	Change from baseline at week 13	-6.3 (3.3)	-4.9 (3.3)
	Difference of LS mean	1.52, (95% CI: 0.21, 2.82), P=0.02	
	BMI	**25 to < 30 kg/m^2**	

	N	45	46
	Baseline mean (SD)	8.1 (3.4)	7.4 (2.5)
	Change from baseline at week 13	-6.9 (4.0)	-4.7 (3.8)
	Difference of LS mean	1.29 (95% CI: 0.46, 2.12), P=0.003	
	BMI	**30 kg/m²**	
	N	30	28
	Baseline mean (SD)	7.2 (2.3)	7.4 (2.1)
	Change from baseline at week 13	-5.7 (3.3)	-5.5 (3.2)
	Difference of LS mean	0.22 (95% CI: -1.01, 1.45), P=0.72	

BMI: body mass index, LS: Least Squares, MHT: menopausal hormonal therapy, mg: milligrams, N: total number of participants, NR: not reported, PBO: placebo, SD: standard deviation, VMS: vasomotor symptoms

Table D3.15. Subgroup data Black 2020

Outcome					
		Black 2020[58]			
VMS frequency		**1mg E2/100mg P4**	**0.5mg E2/100mg P4**	**0.5mg E2/50mg P4**	**PBO**
	BMI	**< 25 kg/m²**			
	N	54	49	52	46
	Baseline mean (SD)	80.4 (45.4)	69.0 (21.4)	75.9 (23.3)	73.0 (20.8)
	Mean change from baseline (SE) at week 12	-56.49	-47.19	-57.09	-36.41
	BMI	**25 to < 30 kg/m²**			
	N	56	55	65	61
	Baseline mean (SD)	73.5 (23.1)	70.3 (30.3)	76.5 (30.8)	70.5 (25.40)
	Mean change from baseline (SE) at week 12	-54.86	-49.26	-46.21	-42.88
	BMI	**30 kg/m²**			
	N	31	45	31	28
	Baseline mean (SD)	65.5 (32.3)	77.6 (30.3)	74.6 (30.3)	75.7 (22.7)
	Mean change from baseline (SE) at week 12	-52.54	-63.83	-46.66	-37.77

VMS severity		1mg E2/100mg P4	0.5mg E2/100mg P4	0.5mg E2/50mg P4	PBO
	BMI	**< 25 kg/m^2**			
	N	54	49	52	46
	Baseline mean (SD)	2.53 (0.25)	2.53 (0.27)	2.49 (0.23)	2.48 (0.27)
	Mean change from baseline (SE) at week 12	-1.08	-0.86	-0.89	-0.47
	BMI	**25 to < 30 kg/m^2**			
	N	56	55	64	61
	Baseline mean (SD)	2.61 (0.24)	2.46 (0.25)	2.48 (0.23)	2.54 (0.24)
	Mean change from baseline (SE) at week 12	-1.2	-0.69	-0.64	-0.64
	BMI	**30 kg/m^2**			
	N	31	45	31	28
	Baseline mean (SD)	2.44 (0.50)	2.56 (0.21)	2.56 (0.24)	2.55 (0.23)
	Mean change from baseline (SE) at week 12	-1.02	-1.13	-0.77	-0.5

BMI: body mass index, kg/ m2: kilogram per meters squared, E2: estradiol, P4: progesterone, MHT: menopausal hormonal therapy, mg: milligrams, N: total number of participants, NR: not reported, PBO: placebo, SD: standard deviation, SE: standard error.

D4. Ongoing Studies

Table D4.1. Ongoing Studies

Trial/ Trial Sponsor	Study Design	Comparators	Patient Population	Primary Outcomes	Estimated Completion Date
Fezolinetant					
A Study to Find the Best Dose of Fezolinetant to Treat Hot Flashes in Women Going Through Menopause (Starlight)	Phase 2, RCT, DB, PC	Fezolinetant 30 mg and 40 mg Placebo	Women aged 40 to 65 years seeking treatment for VMS associated with menopause	Mean change from baseline in the frequency of mild, moderate and severe vasomotor symptom (VMS) from baseline to 8 weeks	Estimated: November 30th, 2022 Primary: December 31st, 2022
A Study of Fezolinetant to Treat Hot Flashes in Women Going Through Menopause (Daylight)	Phase 3, RCT, DB, PC	Fezolinetant twice daily (dose NR) Placebo	Women aged 40 to 65 years seeking treatment for VMS associated with menopause	Mean change in the frequency of moderate to severe VMS from baseline to week 24	Estimated: May 30th, 2023 Primary: May 30th, 2023

Source: www.ClinicalTrials.gov (NOTE: studies listed on site include both clinical trials and observational studies)

D5. Previous Systematic Reviews and Technology Assessments

We identified 19 systematic literature reviews or meta-analyses evaluating the efficacy of therapies of interest for the treatment of VMS associated with menopause, two of which are summarized below.

Sarri, G., et al. (2017). "Vasomotor symptoms resulting from natural menopause: a systematic review and network meta-analysis of treatment effects from the National Institute for Health and Care Excellence guideline on menopause."[118]

This systematic review and network meta-analysis evaluated the comparative efficacy and safety of several medications for non-hysterectomized women in natural menopause. The interventions assessed included sham acupuncture, estrogen plus progestogen non-oral, estrogen plus progestogen oral, tibolone, raloxifene, SSRIs/SNRIs, isoflavones, Chinese herbal medicine, black cohosh, multibotanicals, and acupuncture. Inclusion criteria included randomized controlled trials that assessed pharmacological and/or non-pharmacological treatments for reducing the frequency of VMS for women aged 45 years or older with a diagnosis of natural menopause (defined as amenorrhea for at least 12 consecutive months). The overall NMA protocol stratified studies into three groups: women with a uterus, women without a uterus, women with a history or at risk of breast cancer, and the manuscript presented on the first network. The search was conducted in MEDLINE, Embase, and The Cochrane Library and were restricted to English written articles. There were 47 RCTs that matched the NMA protocol; a total of 32 RCTs (N=4165 women) were used in the analysis of VMS frequency; a total of 21 RCTs (N=4829) were used in the analysis on treatment discontinuation; and a total of 5 RCTs (N=1367) were used in the analysis of vaginal bleeding. Here, we focus on the VMS frequency results to be consistent with this report.

This review reported that estrogen combined with progesterone via patches was more effective than placebo at reducing VMS frequency (MR: 0.23 [95% CI: 0.09, 0.57]) and had the highest probability of being the best treatment of those evaluated (68.9%). Oral estrogen with progesterone was found to have good efficacy (MR: 0.52 [0.25, 1.06]), although the credible intervals were wide. There was no evidence of effects among the other interventions in the network. The review cautioned that there was high heterogeneity between the studies which reduced the precision of estimates. Investigators concluded that there is sufficient evidence that transdermal estrogen combined with progesterone reduced frequency of VMS in women with a uterus and there is some evidence of efficacy for oral estrogen and progesterone.

Shams, T., et al. (2013). "SSRIs for Hot Flashes: A Systematic Review and Meta-Analysis of Randomized Trials"[21]

This systematic review and network meta-analysis evaluated the comparative efficacy and safety of SSRIs for the treatment of VMS in healthy women. The SSRIs assessed included paroxetine, escitalopram, citalopram, fluoxetine, and sertraline. Inclusion criteria included randomized controlled trials that enrolled healthy women who received any SSRI medication and evaluated VMS. This review excluded studies that enrolled cancer patients and patients receiving selective estrange receptor modulators. The search was conducted in MEDLINE, EMBASE, the Cochrane Central Register of Controlled Trials CENTRAL, Web of Science and Scopus) using various combinations of controlled terms: "menopause," post-menopause," "peri-menopause", "hot flushes," "hot flashes," "SSRIs," "climacteric" and "vasomotor." There were 14 manuscripts of 11 RCTs (N=2,069) that met the inclusion criteria.

The review reported that, compared to placebo, SSRIs was associated with a significant decrease in VMS frequency at 4-8 weeks (MD: -0.93 (95% CI: -1.49, -0.37), although this improvement was modest and may not have clinical significance. In terms of adverse events, SSRIs had no significantly higher adverse events compared to placebo but there were more frequent reports for nausea (RR: 1.7 (95% CI: 0.81, 3.59), fatigue (RR: 1.07 (95% CI: 0.60, 1.92), drowsiness (RR: 1.50 (95% CI: 0.42, 5.35), among others. For the NMA, each treatment performed better than placebo and escitalopram had the highest probability to be ranked the best in terms of efficacy in reducing VMS frequency. Results were similar with a fixed effects analysis. The investigators concluded that SSRIs are associated with a modest improvement in VMS frequency and have typical profile of SSRIs; although they caution around the short duration of the RCTs and small sample sizes.

Berhan, Y., & Berham, A. (2014). "Is desvenlafaxine effective and safe in the treatment of menopausal vasomotor symptoms? A meta-analysis and meta-regression of randomized double-blind controlled studies"[128]

This meta-analysis evaluated the efficacy and safety of desvenlafaxine (100 mg or 150 mg daily) for treatment of menopausal related VMS in symptomatic women. Inclusion criteria included double-blind randomized control trials evaluating desvenlafaxine in postmenopausal women seeking treatment for VMS with at least 7 VMS per day, studies published in English, and had a duration of at least 12 weeks. The search was conducted in HINARI, Medline and Cochrane library. They also searched Google Scholar and searching articles from the references of retrieved articles. A total of 7 articles met inclusion criteria (N=3685).

The meta-analysis reported a significant reduction in daily moderate-severe VMS frequency from baseline (SMD = -0.3; 95% CI, -0.41 to -0.22)). There was moderate inconsistency with one trial

reporting no significant reduction in moderate to severe VMS frequency and two trials reporting effect of 50 mg and 200 mg reported no significant improvement. The authors also reported that a large number of women achieved a 50% (overall OR = 2.5; 95% CI, 1.84 to 3.30) and 75% (overall OR = 2.1; 95% CI, 1.65 to 2.53) reduction in moderate to severe VMS frequency from baseline. There was no subgroup effect of BMI, baseline VMS frequency/severity, but there was a significant effect of natural menopause in which desvenlafaxine appeared to be more effective ((regression coefficient =−0.01; 95% CI,- 0.263 to 0.002; P=0.053). In terms of VMS severity, the meta-analysis reported that there was a significant mean reduction in severity of VMS in the desvenlafaxine groups (overall SMD = -0.3; 95% CI, -0.38 to -0.17). There was also moderate variability in this effect but no subgroup effects. Rate of discontinuation was higher in those treated with desvenlafaxine 150mg and 200mg. Although, again, this effect was not consistent with one study reported no differences. There were more adverse events in desvenlafaxine groups, including asthenia, hypertension, anorexia, constipation, dry mouth, among others.

The authors concluded that desvenlafaxine was associated with a significant improvement in VMS frequency and appeared more effective in those with natural menopause. However, the rate of discontinuation was high due to adverse events.

However, this meta-analysis incorrectly included Wyrwich et al. (2008)[50] as a separate data source. Wyrwich et al. was a secondary data analysis of Speroff et al. (2008)[46] and this meta-analysis may have double-counted and miscalculated the magnitude of the effect. Thus, we did not use this meta-analysis in our review and instead abstracted the individual trials, plus additional trials, and synthesized these in our main report.

E. Long-Term Cost-Effectiveness: Supplemental Information

E1. Detailed Methods

Table E.1. Impact Inventory

Sector	Type of Impact (Add additional domains, as relevant)	Included in This Analysis from […] Perspective?*		Notes on Sources (if quantified), Likely Magnitude & Impact (if not)
		Health Care Sector	Societal	
Formal Health Care Sector				
Health Outcomes	Longevity effects	X	X	
	Health-related quality of life effects	X	X	
	Adverse events	X	X	
Medical Costs	Paid by third-party payers	X	X	
	Paid by patients out-of-pocket	☐	☐	
	Future related medical costs	X	X	
	Future unrelated medical costs	X	X	
Informal Health Care Sector				
Health-Related Costs	Patient time costs	NA	☐	
	Unpaid caregiver-time costs	NA	☐	
	Transportation costs	NA	☐	

Non-Health Care Sector				
Productivity	Labor market earnings lost	NA	X	
	Cost of unpaid lost productivity due to illness	NA	X	
	Cost of uncompensated household production	NA	☐	
Consumption	Future consumption unrelated to health	NA	☐	
Social services	Cost of social services as part of intervention	NA	☐	
Legal/Criminal Justice	Number of crimes related to intervention	NA	☐	
	Cost of crimes related to intervention	NA	☐	
Education	Impact of intervention on educational achievement of population	NA	☐	
Housing	Cost of home improvements, remediation	NA	☐	
Environment	Production of toxic waste pollution by intervention	NA	☐	
Other	Other impacts (if relevant)	NA	☐	

NA: not applicable

Adapted from Sanders et al.[129]

Description of evLY Calculations

The equal value life year (evLY) considers any extension of life at the same "weight" no matter what treatment is being evaluated or what population is being modeled. Below are the stepwise calculations used to calculate the evLY.

1. First, we attribute a utility of 0.851, the age- and sex-adjusted utility of the general population in the US that are considered healthy.[103]
2. We calculate the evLY for each model cycle.
3. Within a model cycle, if using the intervention results in additional life years versus the primary comparator, we multiply the general population utility of 0.851 with the additional life years gained (ΔLY gained) within the cycle.
4. The life years shared between the intervention and the comparator use the conventional utility estimate for those life years within the cycle.
5. The total evLY for a cycle is calculated by summing steps 3 and 4.
6. The evLY for the comparator arm is equivalent to the QALY for each model cycle.
7. The total evLYs are then calculated as the sum of evLYs across all model cycles over the time horizon.

Finally, the evLYs gained is the incremental difference in evLYs between the intervention and the comparator arm.

E2. Model Inputs and Assumptions

Model Inputs

Clinical and economic inputs are described in the main report. Other inputs for the model are described below.

Safety and Mortality

While there were no mortality endpoints collected during fezolinetant clinical studies, there may be downstream mortality risks associated with complications from taking MHT. A review of the evidence found the risks of complications such as myocardial infarction, stroke, breast cancer, and lung cancer may be higher for women taking MHT as compared to no MHT. Complications were modeled for all treatment arms and reflect the natural history of unrelated health complications based on age, among other factors. The non-intervention costs listed in the main report tables reflect the average lifetime treatment costs for these complications both in the event year as well as follow-up management costs. These complications also contributed to reductions in quality of life that are reflected in the final QALY estimates.

To build in increased risk of these complications, we apply a risk ratio based on evidence from meta-analyses that include both the risk ratios and the duration of follow-up or duration of exposure on MHT. Both of these inputs were used to increase the incidence, costs, and decrements to quality of life associated with MHT. These complications were not modeled as health states, rather, are tracked outside of health states.

Table E2.1. Safety and Mortality

Category	Parameters	Impact on model outcomes	Values	Sources
MHT-specific Increased Risk of Complications or Death	Coronary event Venous thromboembolism Breast cancer Fractures	Costs Quality of life Life years	Relative risks range by age, complication, and type of therapy (estrogen-only, estrogen and progesterone)	MHT specific: Marjoribanks et al. 2017 and Prentice et al. 2020; [81,130] Baseline incidence of complications [131-134]

MHT: menopausal hormone therapy

Table E2.2. Risk Ratios associated with MHT

Parameter	Value (95% CI)	Age Range	Source
Cardiovascular event	0.93 (0.65, 1.33)	50-59	Roussouw et al. JAMA. 2007;297(13):1465-1477
	0.98 (0.79, 1.21)	60-69	Roussouw et al. JAMA. 2007;297(13):1465-1477
	1.26 (1.00, 1.59)	70-79	Roussouw et al. JAMA. 2007;297(13):1465-1477
Venous thromboembolism	1.91 (1.25, 2.90)	50-59	[80]
	2.03 (1.55, 2.64)	60 and older	Marjoribanks et al. 2017[81]
Breast Cancer	1.18 (0.89, 1.56)	50-59	[80]
	1.27 (1.03, 1.56)	60 and older	Marjoribanks et al. 2017[81]
Fractures	0.61 (0.40, 0.94)	50-59	[80]
	0.78 (0.93, 1.20)	60 and older	Marjoribanks et al. 2017[81]

CI: confidence interval, MHT: menopausal hormone therapy

Health State Utilities

Table E2.3. Disutility Associated with Complications

Parameter	Value
Cardiovascular event	-0.049
Venous Thromboembolism	-0.038
Breast Cancer	-0.02
Fractures	-0.018
Disutility of >1 chronic condition	-0.09

Source: Sullivan et al. [135]

Cost Inputs

All costs used in the model are in 2022 dollars.

Table E2.4. Health State Costs for Complications associated with MHT

Parameter	Value	Source
Cardiovascular event	$60,500	Kazi et al.[136]
Annual cost for managing cardiovascular disease	$3,500	O'Sullivan et al. 2011[137]
Venous thromboembolism	$17,000	Grosse et al. 2016[138]
Venous thromboembolism annual follow-up costs	$6,000	Grosse et al. 2016[138]
Annual cost of managing breast cancer	$105,000	McGarvey et al. 2022[139]
Fracture event	$8,000	Blume and Curtis, 2011[134]

Drug Costs

As no publicly available list or net price exists for fezolinetant, we used a placeholder price of $6,000 per year for estimates of cost-effectiveness based on analyst market projections and uptake assumptions (Table E2.5). This price was used for base-case assessments in the absence of a list price being furnished by the manufacturer; however, this placeholder price was not used to estimate any potential discounts necessary to achieve cost-effectiveness. As we are using notions of generic utilization for MHT, the lowest available WAC prices with no additional rebates or discounts was used for the proxy product chosen to represent the therapeutic class.

Table E2.5. Drug Costs

Intervention	WAC/Placeholder Price per Dose	Discount from WAC	Net/Placeholder Price per Dose	Net/Placeholder Price per Year
Fezolinetant	$16.43*	N/A	$16.43*	$6,000.00*
Menopausal Hormone Therapy	$0.29	N/A	$0.29	$104.83

WAC: wholesale acquisition cost

*Placeholder price of $6,000 annually – interpret findings with caution. No recommendations will be made around discounts to achieve cost-effectiveness unless a price is announced by the manufacturer of fezolinetant.

Non-Drug Costs

Indirect Costs

Table E2.6. details indirect unit costs that were used in the model for both treated and untreated VMS for the modified societal perspective analysis.

Table E2.6. Direct and Indirect Health Care Utilization Cost Inputs

Category	Untreated VMS Annual Costs (per person per year)	Treated VMS Annual Costs (per person per year)	Source
VMS-related Indirect Costs Including Disability and Absenteeism Costs	$790	$350	Sarrel 2015 Menopause inflated to 2022 US dollars[21]

ED: emergency department, VMS: vasomotor symptoms, US: United States

E3. Sensitivity Analyses

As described in the main report, we conducted sensitivity analyses such as one-way and probabilistic sensitivity analyses. The probabilistic analyses are shown in the main report.

Tornado diagrams for fezolinetant are shown in Figures E3.1 and E3.2. It was decided to present fezolinetant one-way sensitivity analysis findings in terms of incremental costs and incremental health outcomes separately, as the reasonable low and high bounds for the utility associated with symptoms while on fezolinetant led to a difficult-to-interpret cost-effectiveness ratio. Also, the tornado diagrams for fezolinetant appear truncated particularly because the agent has no documented impact on the comorbidities that MHT has been shown to have. Fezolinetant's effect on quality of life has been redacted due to academic-in-confidence data and therefore is not included in the sensitivity analyses for this draft report. In terms of incremental QALYs, the model was also sensitive to the proportion of those who discontinue fezolinetant during the first year. On the costs side, the model was most sensitive to the cost of treated VMS per year and also to the proportion who discontinue fezolinetant. The cost of fezolinetant was not included in this tornado diagram in this version as its price is unknown; future versions of the report will include a price in the tornado, likely at a threshold price for fezolinetant.

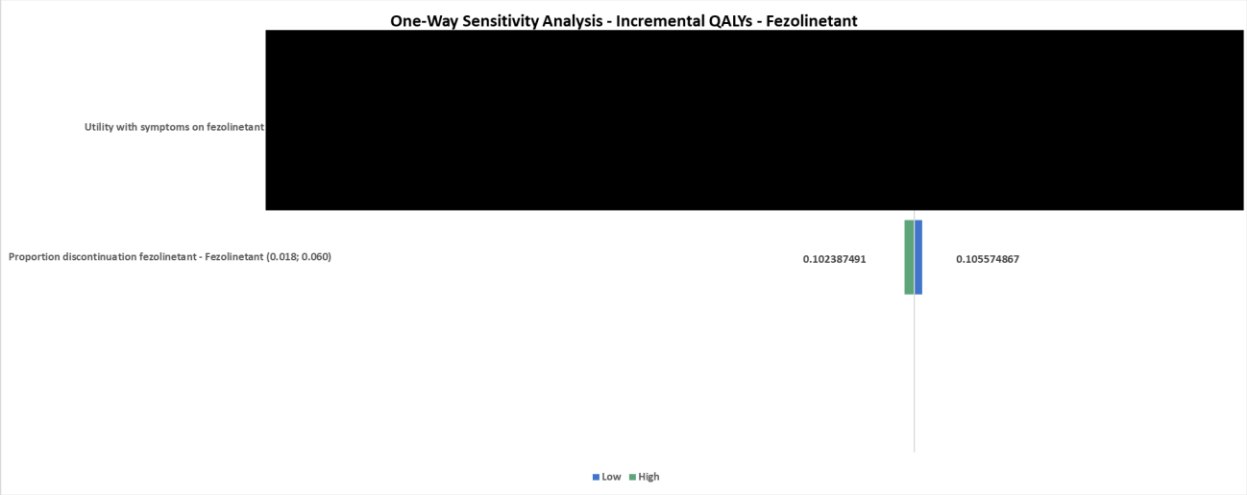
Figure E3.3 shows the tornado diagram for MHT compared to no pharmacologic treatment. The incremental cost per QALY results were most sensitive to the annual cost of managing breast cancer, the relative risk of breast cancer, and the cost of treated VMS per year. Interpretation of incremental cost-effectiveness shifted towards an estimate of less costly, more effective than placebo when favorable low input values were chosen for these parameters.

Figure E3.1. One-Way Sensitivity Analysis Results and Tornado Diagram for Fezolinetant Compared to No Pharmacologic Treatment: Incremental Costs

VMS: vasomotor symptoms

*This one-way sensitivity analysis uses a placeholder price for fezolinetant

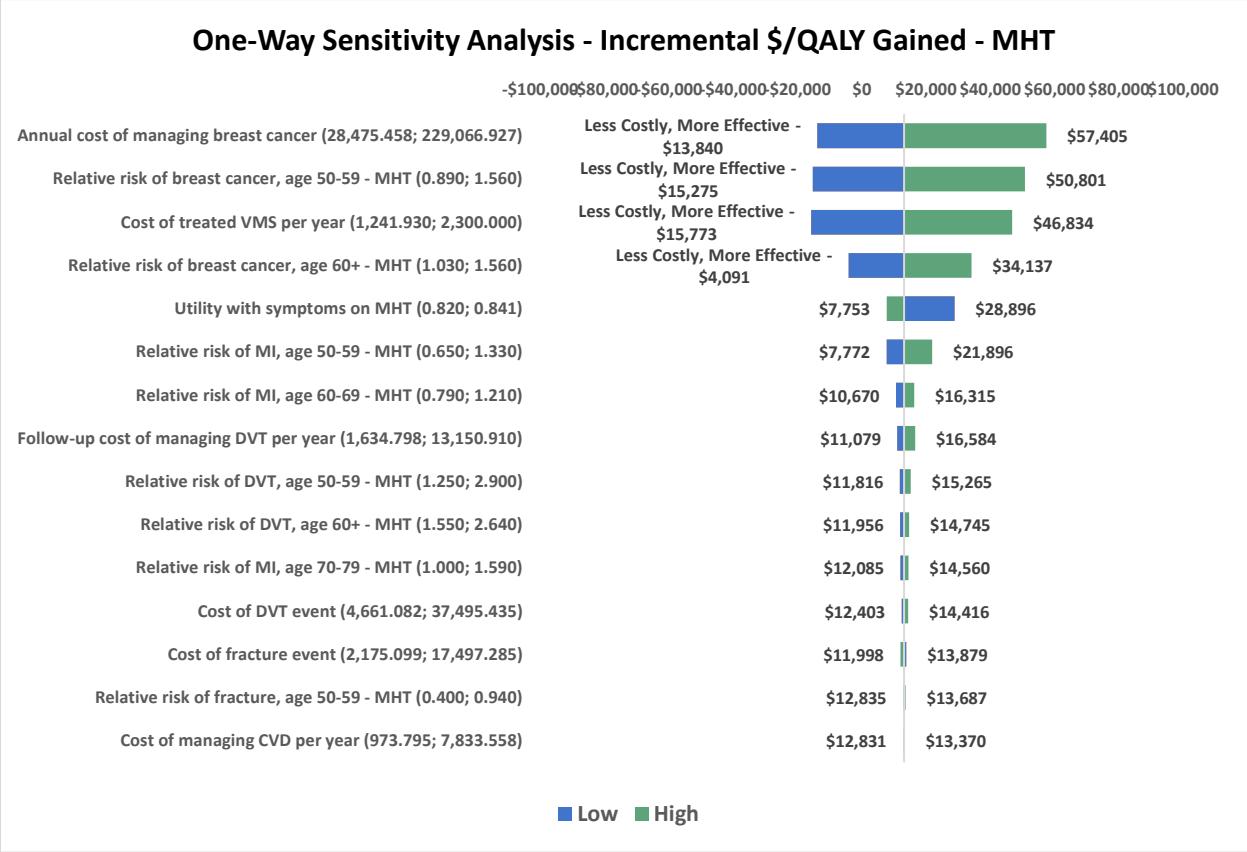
Figure E3.2. One-Way Sensitivity Analysis Results and Tornado Diagram for Fezolinetant Compared to No Pharmacologic Treatment: Incremental QALYs



QALY: quality-adjusted life-years

*This one-way sensitivity analysis uses a placeholder price for fezolinetant

Figure E3.3. One-Way Sensitivity Analysis Results and Tornado Diagram for MHT Compared to No Pharmacologic Treatment: Incremental Cost per QALY Gained



CVD: cardiovascular disease, DVT: deep vein thrombosis, MHT: menopausal hormone therapy, MI: myocardial infarction, VMS: vasomotor symptoms

Prior Economic Models

As expected, the majority of economic evaluation studies have been focused on MHT[140]. However, NICE conducted an analysis in 2015 that included both MHT and non-hormonal treatment options[98]. The model structure was similar to the one used for this analysis with a focus on continuing treatment or discontinuing treatment (i.e., labeled in this model as on or off treatment). Discontinuation was limited to meta-analyses or assumptions instead of the more dynamic approach taken in this analysis that uses both short-term discontinuation from trial evidence and long-term discontinuation from resolution of symptoms. Health state utilities were also derived using a mapping algorithm based on changes in frequency of VMS. This analysis also uses a mapping algorithm based on the MENQoL, which includes more health domains related to VMS and symptoms associated with VMS. Risks associated with MHT included VTE and breast cancer whereas this analysis included both risks in addition to cardiovascular disease and fractures. Finally,

it's important not to directly compare outcomes and costs between these analyses as NICE used a 5-year time horizon and the analysis was specific to an NHS perspective.

While the focus of this review was not MHT, we did include a scenario analysis comparing MHT to placebo to aid in our understanding and interpretation of the value of fezolinetant. Therefore, we modeled both the quality of life and cost implications of long-run risks associated with MHT. The systematic literature published in 2017 identified five evaluations specific to MHT since 2002 and noted considerable variation in modeled complications associated with MHT[140]. Key recommendations from this review were 1) to consider the full range of complications associated with MHT; and 2) ensure a long-run time horizon so modeled costs and health outcomes reflect the impact of MHT over both short-run and long-run use. As stated above, we include more complications than a previous analysis that included both MHT and non-hormonal therapies. Moreover, the model time horizon was lifetime as opposed to a shorter time horizon reflective of the duration of symptoms only.

F. Potential Budget Impact: Supplemental Information

Methods

We used results from the same model employed for the cost-effectiveness analysis of fezolinetant compared to placebo, which approximates no pharmacologic treatment – neither prescription nor non-prescription, to estimate total potential budget impact of fezolinetant. Potential budget impact was defined as the total differential cost of using fezolinetant rather than using no prescription treatments for the hypothetical treated population, calculated as differential health care costs (including drug costs) minus any offsets in these costs from averted health care events. All costs were undiscounted and estimated over one- and five-year time horizons. The five-year timeframe was of primary interest, given the potential for cost offsets to accrue over time and to allow a more realistic impact on the number of patients treated with the new therapy.

The potential budget impact analysis included women with VMS associated with menopause who were ultimately for treatment. To estimate the size of the potential candidate population for treatment, we use inputs for US population size (~342,000,000) [141], percent women (50.2%) [142], percent within menopausal age (40-65 years) (34.1%) [142], proportion experiencing menopause (63.5%) [113], proportion with VMS (80%) [105], and percent seeking care from a health care provider (72%) [143]. In the absence of more recent and more robust data, we assume 22% of patients are receiving MHT at baseline and thus would be excluded from the analysis.[27] Applying these sources results in estimates of 16,700,000 eligible patients in the US. For the purposes of this analysis, we assume that 20% of these patients would initiate treatment in each of the five years, or 3,340,000 patients per year. Of note, we plan to include MHT as an optional comparator for evaluation in the budget impact model published within ICER's Interactive Modeler tool for VMS associated with menopause upon posting of the corresponding Final Evidence Report.

ICER's methods for estimating potential budget impact are described in detail elsewhere and were last updated in 2020.[144,145] The intent of our revised approach to budgetary impact is to document the percentage of patients that could be treated at selected prices without crossing a budget impact threshold that is aligned with overall growth in the US economy.

Briefly, we evaluate a new drug that would take market share from one or more existing therapies, and calculate the blended budget impact associated with displacing use of existing therapies with the new intervention. As the cost-effectiveness analysis for fezolinetant as part of this review included patients who cannot or will not take MHT for VMS associated with menopause, we excluded MHT patients from the present budget impact analysis. And as there were insufficient

data to model SSRIs/SNRIs in the cost-effectiveness analyses, we also did not include that treatment modality bucket here. As such, all patients were assigned to the no pharmacologic treatment arm at baseline, and all would initiate fezolinetant over 5 years in order to estimate maximum feasible uptake without crossing ICER's potential budget impact threshold.

Using this approach to estimate potential budget impact, we then compared our estimates to an updated budget impact threshold that represents a potential trigger for policy mechanisms to improve affordability, such as changes to pricing, payment, or patient eligibility. As described in ICER's methods presentation, this threshold is based on an underlying assumption that health care costs should not grow much faster than growth in the overall national economy. From this foundational assumption, our potential budget impact threshold is derived using an estimate of growth in US gross domestic product (GDP)+1%, the average number of new drug approvals by the FDA over the most recent two-year period, and the contribution of spending on retail and facility-based drugs to total health care spending.

For 2022-2023, therefore, the five-year annualized potential budget impact threshold that should trigger policy actions to manage access and affordability is calculated to total approximately $777 million per year for new drugs.

ICER's methods for estimating potential budget impact are described in detail in the Value Assessment Framework. The intent of our approach to budgetary impact is to document the percentage of patients that could be treated at selected prices without crossing a budget impact threshold that represents a potential trigger for policy mechanisms to improve affordability, such as changes to pricing, payment, or patient eligibility.

Using this approach to estimate potential budget impact, we then compared our estimates to our updated budget impact threshold that represents a potential trigger for policy mechanisms to improve affordability, such as changes to pricing, payment, or patient eligibility.

Results

Table F1. illustrates the average annual per-patient budget impact calculations across fezolinetant's placeholder price ($6,000 per year) and the prices to reach $150,000, $100,000, and $50,000 per QALY compared to no pharmacologic treatment ($2,500, $2,000, and $1,500 per year, respectively).

Table F1. Average Annual Per-Patient Budget Impact Calculations Over a 5-Year Time Horizon

	Placeholder Price*	$150,000/QALY	$100,000/QALY	$50,000/QALY
Fezolinetant Compared to No Pharmacologic Treatment	$5,000	$2,000	$1,000	$500

QALY: quality-adjusted life year, WAC: wholesale acquisition cost

*Based on placeholder price of $6,000 annually. Interpret findings with caution.

In contrast, Figure F1. below illustrates the cumulative annual net budget impact of fezolinetant compared to no pharmacologic treatment. Fezolinetant's budget impact at its placeholder price is largely aligned with its placeholder acquisition cost minus a ~$500 annual resource utilization and cost offset benefit conferred with active treatment for VMS associated with menopause.

Figure F1. Estimated Cumulative Annual Net Budget Impact Per Treated Patient Per Year for Fezolinetant Compared to No Pharmacologic Treatment for VMS at Placeholder Price