Comparative Effectiveness Review
Number 147

Effective Health Care Program

Menopausal Symptoms: Comparative Effectiveness of Therapies

Executive Summary

Background

Menopause is defined as the permanent cessation of menstruation and ovulation due to ovarian failure. "Spontaneous" menopause occurs after 12 months of amenorrhea as ovarian hormone secretion diminishes, on average around the age of 51 years. Menopause may be induced prematurely (before age 40 years) or early (before age 45 years) through medical interventions such as surgery (e.g., bilateral oophorectomy with or without hysterectomy), chemotherapy, or radiation. In the United States, the number of women entering menopause each year is estimated to be approximately 2 million.[1]

Current terminology describing the stages of menopause was updated in 2012 at the Stages of Reproductive Aging Workshop + 10 (STRAW + 10).[2] The STRAW + 10 stages describe early and late phases of menopausal transition and early and late phases of postmenopause. Menopausal transition is defined by variability in menstrual cycle length, followed by periods of amenorrhea lasting 60 days or longer. Perimenopause is defined as the entire menopausal transition phase, extending into the first 12 months of the early postmenopause stage. Early postmenopause lasts from 5 to 8 years, from final menstrual period to stabilization of low estradiol levels.[2]

Effective Health Care Program

The Effective Health Care Program was initiated in 2005 to provide valid evidence about the comparative effectiveness of different medical interventions. The object is to help consumers, health care providers, and others in making informed choices among treatment alternatives. Through its Comparative Effectiveness Reviews, the program supports systematic appraisals of existing scientific evidence regarding treatments for high-priority health conditions. It also promotes and generates new scientific evidence by identifying gaps in existing scientific evidence and supporting new research. The program puts special emphasis on translating findings into a variety of useful formats for different stakeholders, including consumers.

The full report and this summary are available at **www.effectivehealthcare.ahrq.gov/reports/final.cfm**.

Approximately 85 percent of women report experiencing symptoms of varying type and severity during menopause.[3] Types of symptoms experienced may include[1]—

Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov

Effective Health Care

- Vasomotor symptoms: Hot flushes are recurrent, transient episodes of intense heat in the face and upper body, sometimes followed by chills. These symptoms can occur while sleeping, producing intense perspiration. Individual hot flushes may last from 1 to 5 minutes. After irregular menses, vasomotor symptoms are the second most frequently reported perimenopausal symptoms.
- Sleep disturbances: Lengthy times to fall asleep, inability to sleep through the night, or inability to resume sleeping when waked prematurely are signs of insomnia. Sleep apnea symptoms range from slight airflow reductions causing snoring to periodic cessation of breathing.
- Psychological symptoms: Depressive symptoms, anxiety, and mood disturbances may occur. Depressive symptoms can range from a depressed mood to clinical depression. A depressed mood may not require treatment, but if clinical depression is suspected, assessment and treatment are recommended. Symptoms of anxiety may include tension, nervousness, panic, and worry.
- Urogenital problems: Urinary incontinence and vaginal atrophy may occur. Vaginal atrophy involves vaginal walls that are thin, pale, dry, and sometimes inflamed. These changes cause discomfort and potential trauma during intercourse and pelvic examinations.
- Sexual function effects: Dyspareunia (pain during intercourse) and decreased libido are also reported by perimenopausal and postmenopausal women.

Longitudinal studies have shown that during early postmenopause, the prevalence of vasomotor symptoms among women ranges from 30 to 80 percent, depressed mood occurs in approximately one-third, and sleep disturbance occurs in more than 40 percent.[4-6] Vasomotor symptoms generally begin 2 years before the final menstrual period, peak during the 1 year after the final menstrual period, and then diminish.[7] Urogenital atrophy symptoms increase during the late postmenopause stage.[2] Differences in symptoms have been found among subpopulations of women. In the Penn Ovarian Aging Trial[8] and the Study of Women's Health Across the Nation,[9] researchers report differences in prevalence and duration of vasomotor symptoms among women depending on ethnicity and body mass index (BMI).

Objectives and Key Questions

The objective of this review is to systematically review and synthesize evidence evaluating the comparative effectiveness of treatments for menopausal symptoms, along with potential long-term benefits and harms.

The Key Questions we considered are—

Key Question 1. What is the comparative effectiveness of different treatments for reducing symptoms of menopause (vasomotor symptoms, sleep disturbance, psychological symptoms, urogenital atrophy, and sexual function) and for improving quality of life? Individual agents will be compared to the extent permitted by the evidence.

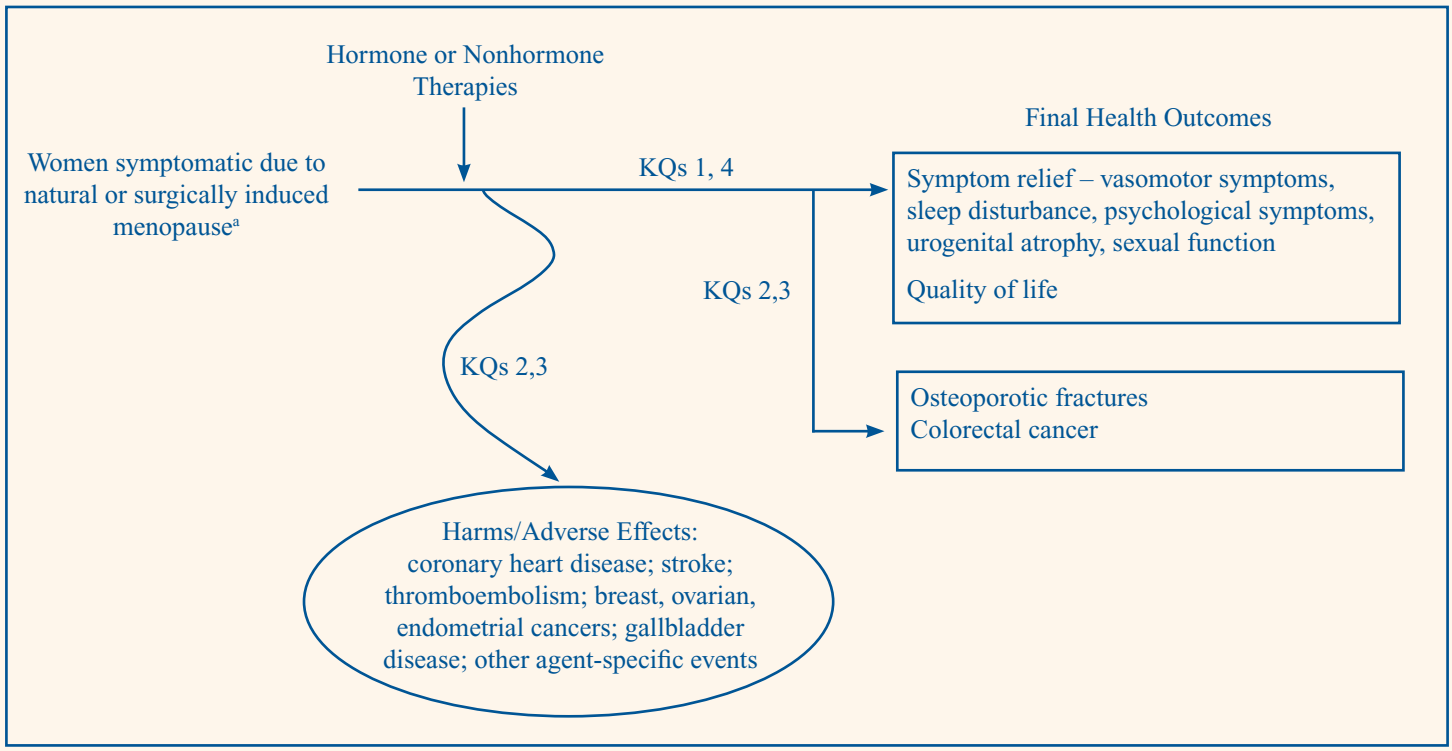
Key Question 2. What are the effects of menopausal hormone therapy preparations on coronary heart disease, stroke, or venous thromboembolism; gallbladder disease; osteoporotic fractures; or endometrial, breast, colorectal, or ovarian cancer? Exposure will be examined according to duration of use and initiation relative to age and onset of menopause. (For women desiring contraception, combined estrogen-progestogen and progesterone-only contraceptives are included.)

Key Question 3. What are the effects of nonhormone therapy preparations on coronary heart disease, stroke, or venous thromboembolism; gallbladder disease; osteoporotic fractures; or endometrial, breast, colorectal, or ovarian cancer? Exposure will be examined according to duration of use and initiation relative to age and onset of menopause. What are the significant agent-specific harms/adverse effects of nonhormone therapies?

Key Question 4. Do effectiveness and adverse effects vary among subgroups of participants defined by demographics, symptom severity, other medications, and comorbidities or according to agent, preparation, or dose?

Figure A shows the analytic framework for our review.

Figure A. Analytic framework



[a]Excludes women with breast cancer or receiving tamoxifen.

KQ = Key Question.

Methods

Input From Stakeholders

During topic refinement, input was sought from Key Informants representing clinicians (internal medicine, family practice, and gynecology), academicians, researchers, and patients. Key Questions were subsequently posted and public comment obtained. A Technical Expert Panel was assembled, including content and clinical experts. Comments were reviewed and appropriate changes to Key Questions made.

Data Sources and Selection

The final literature search, including articles through January 2014, was run on MEDLINE®, Embase®, Cochrane Controlled Trials Register, and AMED Allied and Complementary Medicine. The reference lists for systematic reviews and meta-analyses were also screened to identify additional references. The gray literature search included extensive reviews of clinicaltrials.gov, the U.S. Food and Drug Administration (FDA) Web site, and relevant conference abstracts. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) diagram (Figure B) depicts the flow of search, title/abstract screening, full-text screening, and study selection.

For Key Question 1 (symptom relief from any therapy), we included randomized controlled trials (RCTs) with 25 or more participants per arm and a followup of 4 weeks or longer for centrally acting agents and 12 weeks for all other therapies. Trials enrolling women with preexisting conditions (e.g., heart disease, lupus, fibromyalgia, breast cancer) were excluded. For Key Question 2 (long-term effects of hormone therapies), systematic reviews and meta-analyses were included. Studies with intermediate outcomes and studies with both pre- and postmenopausal women combined were excluded. Key Question 3 was a two-part question examining adverse events and long-term effects of nonhormone therapies. For the adverse events question, trials included in Key Question 1 that also reported adverse events were included. For the long-term effects question, RCTs and observational studies were

included. Exclusions for Key Question 3 included dietary population studies, studies with intermediate outcomes, and studies with both pre- and postmenopausal women combined. For Key Question 4, subgroup analyses of symptom relief from any therapy, trials from Key Question 1 that reported subgroup analyses were included.

A total of 8,372 records were excluded in the first round of screening because, from the title and abstract, the screeners could discern that the articles did not meet one or more of the inclusion criteria relating to study design, outcome, population, or comparator.

Data Abstraction and Quality Assessment

Data Abstraction

Key Question 1 and Key Question 4

Data were abstracted into collection forms created in DistillerSR. Two training sets of three articles each were abstracted by all team members. Results of each training set were reviewed to discuss any discrepancies in abstraction. Final data abstraction was performed by one team member and verified by a different team member, with any identified inconsistencies resolved by consensus. The following data were abstracted:

- Trial characteristics: Author, year, country, number of trial sites, trial design, total number randomized, length of followup, intervention, uterine status, disclosures and conflicts of interest, funding, primary and secondary outcomes
- Trial arm characteristics: Participant information such as number of participants, age, ethnicity, BMI, time since menopause, tobacco use; treatment specifics such as type of treatment, dosage, dosage category, and route of administration
- Outcomes: Scale; results from baseline, 12-week, and final assessments; mean scores, mean changes, percent reductions, standard deviations, 95% confidence intervals, pre/post intervention comparisons, and between-group comparisons

When only graphical outcomes were presented, figures were digitized. For Key Question 1, standardized mean differences were calculated from reported estimates of treatment effects, standard deviations, and p-values.

Key Question 2

Data abstracted from the systematic reviews and meta-analyses include the following: included trials, treatment type, treatment dose, length of followup, and results.

Key Question 3

Summary tables of long-term effects of nonhormone therapies contained the following information: condition, treatment, study design, study descriptions, and results.

Agent-specific adverse events for nonhormone therapies were categorized using a system recommended by the International Federation of Pharmaceutical Manufacturers and Associations.[10] The following data were abstracted for each category: author, year, country, treatment, dose, trial size, total adverse events, and percentage of events.

Quality Assessment

In adherence with the "Methods Guide for Effectiveness and Comparative Effectiveness Reviews" (Methods Guide),[11] the general approach to grading trials was performed by applying the criteria of the U.S. Preventive Services Task Force (USPSTF).[12] Discordant assessments were resolved with input from a third reviewer.

Study quality of RCTs was assessed by assembly of comparable groups, blinding of researchers and subjects, concealment of group assignment, maintenance of comparable groups, differential loss to followup, equal and reliable measurements, clearly defined interventions, important outcomes considered and defined, and intention-to-treat analysis.

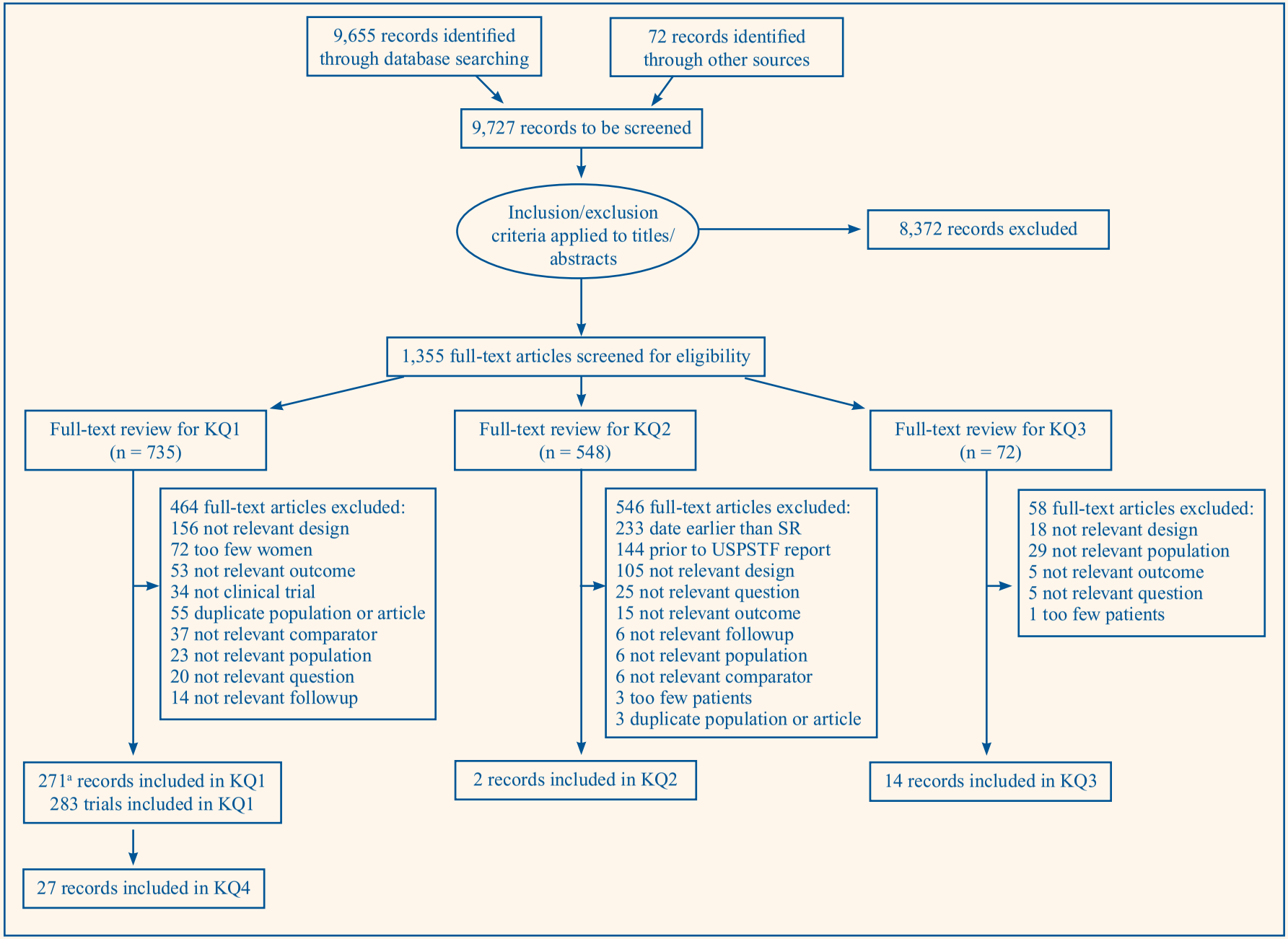
Study quality of cohort studies was assessed by assembly of comparable groups, maintenance of comparable groups, differential loss to followup, equal and reliable measurements, important outcomes considered and defined, and statistical adjustment for potential confounders.

Study quality of case-control studies was assessed by accurate ascertainment of cases, nonbiased selection of cases and controls, response rate, equal application of diagnostic tests, accurate and equal measure of exposure, and attention to potential confounders.

Data Synthesis and Analysis

For Key Question 1, trials employed a variety of outcome instruments. Standardized mean differences were calculated and pooled according to the Methods Guide.[11,13] Calculating the standardized mean difference (SMD), which is (effect treatment – effect comparator)/ standard deviation, allows comparison of results across trials using different measures. Clinical heterogeneity and appropriateness for pooling were judged by the review team on the basis of study characteristics together with clinical context. Because the goal of any pooling is to estimate unconditional effects,[14] random-effects models

Figure B. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram



KQ = Key Question; SR = systematic review; USPSTF = U.S. Preventive Services Task Force.

[a]12 records presented results from 2 distinct patient populations and were divided into 2 trials each.

were used. The magnitude of statistical heterogeneity was examined by using tau^2 owing to limitations of the I^2 metric and because between-trial variances are more intuitively interpreted on the effect-estimate scale.[15] Evidence of possible publication bias were explored using funnel plots and Egger test when results from at least 10 studies were pooled.

For vasomotor symptoms and quality-of-life outcomes, network meta-analyses formed the primary analyses, including the most relevant comparisons with sufficient data. Network meta-analysis formally allows quantitative indirect and mixed-treatment comparisons. The random-effects network meta-analysis was performed by pooling standardized mean differences in a Bayesian model described by Chaimani (www.mtm.uoi.gr/). Models were fitted in OpenBUGS using noninformative priors and convergence assessed using the Brooks-Gelman-Rubin plot and statistic, autocorrelation, and history plots. A burn-in of 20,000 samples was discarded and the subsequent 40,000 analyzed. Rankings were estimated for the probability a treatment was most effective, next most effective, and so on. Effect estimates and accompanying 95% credible intervals were obtained from the samples. To evaluate consistency, we compared available pairwise estimates with the network results[16] and explored them graphically (www.mtm.uoi.gr). We examined pairwise comparisons individually in random-effects models and graphically using forest plots.

Evidence for the remaining Key Questions consisted of systematic reviews, observational studies, and a few RCTs. Quantitative analyses were not possible, and therefore a qualitative discussion of the evidence was conducted.

Strength of the Body of Evidence

Strength-of-evidence (SOE) assessments were based on the Evidence-based Practice Center approach,[11] which is conceptually similar to the GRADE (Grading of Recommendations Assessment, Development and Evaluation) system.[17] Two reviewers graded the strength of evidence, resolving disagreements by consensus.

We adopted a point-based approach to SOE ratings. Each rating started at high (3 points) and was downgraded by 1 point each for high risk of bias, inconsistent or unknown consistency, imprecise or unknown precision, indirect body of evidence, and suspected reporting bias. Domain ratings were entered into a spreadsheet that provided a summary SOE. If the summary SOE remained 3 with no downgrades, it was rated high; if the summary SOE equaled 2, it was rated moderate; if the summary SOE equaled 1, it was rated low; if the summary SOE was 0 or lower, it was rated insufficient. Following Agency for Healthcare Research and Quality guidance for assessing evidence on equivalence and noninferiority, studies can be appropriately considered individually in the presence of clinical heterogeneity; as stated by Treadwell and colleagues, "the lack of meta-analysis does not necessarily preclude a conclusion of EQ-NI [Equivalence-noninferiority], just as it does not preclude an evaluation of the strength of evidence in relation to a particular outcome."[18]

Results

Results are presented below for symptom relief (Key Question 1), other benefits and harms (Key Questions 2 and 3), and symptom relief among subgroups (Key Question 4).

Symptom Relief

Summary results are presented by outcome (vasomotor symptoms, quality of life, psychological symptoms, sexual function, urogenital atrophy, and sleep disturbances), followed by a brief discussion of compounded hormone therapies and limitations of the evidence base for symptom relief. Investigators used many different measurement rating scales to evaluate treatment effects. Pooling across scales can be accomplished only by using SMDs. Although they enable pooling, SMDs pose challenges for clinical interpretation. To place their magnitudes into context, with control-group event rates of 20 to 60 percent, SMDs can be expressed as approximate odds ratios (ORs). For example, SMDs and corresponding ORs (in parentheses) are as follows: SMD, -0.2 (OR, 0.7); -0.3 (0.6); -0.4 (0.5); -0.5 (0.4); 0.3 (2); 0.6 (3); and 0.75 (4). Although the ORs exceed relative risks when placebo group event rates exceed 10 percent, they provide a rough guide to the relative effect. For example, the placebo response rate of women with vasomotor symptoms can vary between approximately 20 and 40 percent.

For analytical purposes, estrogen doses were classified as low/ultralow, standard, and high. For oral treatment, which was the most common route of administration, the dosing categories were based on the 2009 Cochrane Review on hormone replacement therapy and endometrial hyperplasia.[19] For example, dose categories for oral conjugated equine estrogens were ultralow (0.15 to 0.3 mg), low (0.4 mg), standard (0.625 mg), and high (1.25 mg). For other routes of administration, such as transdermal and spray, dosing categorizations were established in consultation with the clinical content expert.

Vasomotor Symptoms

A large body of evidence was identified comparing the efficacy of agents versus placebo and other active treatments for the relief of vasomotor symptoms (Table A). One quarter of trials were rated good or fair quality and the remainder poor. Trials were most numerous for estrogens, isoflavones, selective serotonin reuptake inhibitors (SSRIs) and serotonin-norepinephrine reuptake inhibitors (SNRIs), gabapentin, black cohosh, and ginseng. Estrogens of any dose appeared more effective than other comparators, without apparent meaningful differences according to dose or route of administration. Small differences in effect magnitudes among SSRIs/SNRIs, isoflavones, gabapentin, black cohosh, and ginseng were apparent in network meta-analysis. Mean rankings of treatment effectiveness (1 being best, 9 worst; placebo ranked 8.9) were as follows: high-dose estrogens (1.9), standard-dose estrogens (1.3), low-dose estrogens (2.9), SSRI/SNRI (4.9), gabapentin (5.6), isoflavones (5.9), black cohosh (6.7), and ginseng (7.0). A host of other agents have been studied, but evidence is limited to single trials.

The efficacy of estrogens in treating vasomotor symptoms is well established. The comparative effectiveness of other agents relative to estrogens has been less clear. Albeit limited by the trial quality, the findings show that other agents can ameliorate vasomotor symptoms, but none have estrogen's effectiveness.

Table A. Magnitude and strength of evidence of treatments for vasomotor symptoms: standardized mean differences from pairwise comparisons

Number of Comparisons	Comparators	Effect Size (SMD) (95% CI)	Strength of Evidence
9	Estrogen (high) vs. placebo	-0.50 (-0.61 to -0.39)	High
39	Estrogen (standard) vs. placebo	-0.64 (-0.74 to -0.53)	High
53	Estrogen (low/ultralow) vs. placebo	-0.55 (-0.61 to -0.48)	High
13	SSRI/SNRI vs. placebo	-0.35 (-0.46 to -0.24)	High
5	Gabapentin vs. placebo	-0.28 (-0.38 to -0.19)	Moderate
35	Isoflavones vs. placebo	-0.31 (-0.41 to -0.22)	Low
4	Black cohosh vs. placebo	-0.31 (-0.46 to -0.15)	Low
3	Ginseng vs. placebo	-0.17 (-0.43 to 0.09)	Low

CI = confidence interval; SMD = standardized mean difference; SNRI = serotonin-norepinephrine reuptake inhibitor; SSRI = selective serotonin reuptake inhibitor.

Quality of Life

Trials evaluating numerous agents—estrogens, isoflavones, SSRIs/SNRIs, ginseng, black cohosh, and dehydroepiandrosterone (DHEA)—reported some quality-of-life metric (Table B). Less than a third of trials (27.2%) were rated good or fair quality. Compared with placebo, improved quality-of-life scores accompanied estrogens, with SMDs exceeding 0.35 with high SOE; effect sizes for all other agents were lesser in magnitude or low SOE. Similarly, estrogens ranked highest in the network comparison. For estrogens, there were no apparent meaningful differences in effect according to dose or route of administration. Quality-of-life scores were reported from trials of many nonprescription agents, but results from single trials do not allow conclusions concerning effects.

We found improved global quality-of-life scores in women taking estrogens. Two of the larger trials, Women's International Study of long Duration Oestrogen after the Menopause (WISDOM)[20] and Women's Health Initiative (WHI)[21,22] reported no effect of estrogens on quality of life, a finding potentially attributable to older age and less symptom severity of enrolled women in these trials or the lack of employment of menopause-specific instruments.

For the larger body of comparisons in women receiving estrogens, despite between-trial variability, results were more consistent. The general pattern of comparative efficacy seen with quality-of-life scores paralleled results for vasomotor and other symptoms.

Table B. Magnitude and strength of evidence of treatments for quality of life: standardized mean differences from pairwise comparisons

Number of Comparisons	Comparators	Effect Size (SMD) (95% CI)	Strength of Evidence
5	Estrogen (high) vs. placebo	0.76 (0.48 to 1.03)	High
26	Estrogen (standard) vs. placebo	0.55 (0.41 to 0.69)	High
17	Estrogen (low/ultralow) vs. placebo	0.36 (0.27 to 0.45)	High
6	SSRI/SNRI vs. placebo	0.28 (0.17 to 0.39)	High
24	Isoflavones vs. placebo	0.27 (0.17 to 0.37)	Low
4	Black cohosh vs. placebo	0.26 (-0.15 to 0.66)	Insufficient
3	Ginseng vs. placebo	0.19 (0.01 to 0.36)	Low

CI = confidence interval; SMD = standardized mean difference; SNRI = serotonin-norepinephrine reuptake inhibitor; SSRI = selective serotonin reuptake inhibitor.

Psychological Symptoms

Just over one-third of trials examining symptom treatment reported at least one psychological outcome—depressive symptoms, anxiety, or global psychological well-being. Of these trials, 28.8 percent were rated good or fair quality. Approximately half specified some psychological symptom as a primary outcome. Generally, the samples were not selected to represent populations with clinical depression or anxiety. Compared with placebo, SMDs were in general not large (i.e., SMD between -0.5 and 0) for any of the agents studied for any psychological domain (Table C). The SOE was high that SSRIs/SNRIs and estrogens can effectively alleviate psychological symptoms in all domains.

An increased risk for depressive symptoms during the menopausal transition in the absence of prior depressive illness has been described[23] and may be associated with vasomotor symptoms.[24] The effects assessed here may provide guidance when menopausal women are experiencing psychological symptoms.

Table C. Magnitude and strength of evidence of treatments for psychological symptoms: standardized mean differences from pairwise comparisons

Domain	Number of Comparisons	Comparators	Effect Size (SMD) (95% CI)	Strength of Evidence
Global	6	SSRI/SNRI vs. placebo	-0.42 (-0.60 to -0.24)	High
Depressive symptoms	5	SSRI/SNRI vs. placebo	-0.43 (-0.60 to -0.26)	High
Anxiety symptoms	3	SNRI vs. placebo	-0.31 (-0.50 to -0.12)	High
Global	14	Estrogen vs. placebo	-0.26 (-0.40 to -0.13)	High
Depressive symptoms	18	Estrogen vs. placebo	-0.36 (-0.53 to -0.20)	High
Anxiety symptoms	13	Estrogen vs. placebo	-0.34 (-0.50 to -0.18)	High
Global	2	Gabapentin vs. placebo	-0.23 (-0.48 to 0.02)	Insufficient
Global	7	Isoflavones vs. placebo	-0.11 (-0.22 to 0.01)	Low
Depressive symptoms	9	Isoflavones vs. placebo	-0.29 (-0.49 to -0.09)	Low
Anxiety symptoms	7	Isoflavones vs. placebo	-0.30 (-0.46 to -0.14)	Moderate

CI = confidence interval; SMD = standardized mean difference; SNRI = serotonin-norepinephrine reuptake inhibitor; SSRI = selective serotonin reuptake inhibitor.

Sexual Function

Some measure of sexual function was reported in approximately one-third of trials; 41.4 percent of those trials specified the outcome as primary. Of these trials, 21 percent were rated good or fair quality. Outcomes were reported in four domains: pain (dyspareunia), a global metric, activity, and interest. Vaginal estrogens decreased pain most convincingly (high SOE), and lower pain scores were also reported with oral estrogens (moderate SOE) (Table D). There was improvement in global measures with all estrogens (high SOE). Estrogens appeared to enhance measures of interest, while SSRIs/SNRIs showed only modest improvement. Sexually satisfying episodes were more frequent with testosterone (7 out of the 8 trials administered testosterone through a patch) compared with placebo—slightly more than one extra episode reported every 4 weeks (moderate SOE). Overall, these results are generally consistent with evidence-informed expert clinical opinion.[1]

The Prevalence of Female Sexual Problems Associated with Distress and Determinants of Treatment Seeking (PRESIDE) study[25] estimated that approximately 15 percent of women age 45 to 64 years experienced some form of sexual distress. A cohort study, Study of Women's Health Across the Nation (SWAN),[26] reported that during the menopausal transition, there are significant decreases in sexual interest, frequency, and arousal along with increased pain during sex. One quantitative review on sexual outcomes during menopause included literature published between 1972 and 1992.[27] In this review by Myers, the effect of estrogen therapy on all four sexual function domains combined (108 studies) yielded an SMD of -0.67—somewhat larger in magnitude than that obtained in this review.

Table D. Magnitude and strength of evidence of treatments for sexual function: standardized mean differences from pairwise comparisons

Domain and Number of Comparisons	Comparators	Effect Size (SMD) (95% CI)	Strength of Evidence
Pain (lower is better)			
10	Vaginal estrogens vs. placebo	-0.54 (-0.73 to -0.34)	High
4	Oral estrogens vs. placebo	-0.22 (-0.35 to -0.09)	Moderate
Global (higher is better)			
15	All estrogens vs. placebo	0.27 (0.19 to 0.35)	High
2	SSRI/SNRI vs. placebo	0.27 (0.01 to 0.52)	Insufficient
4	Isoflavones vs. placebo	0.24 (-0.12 to 0.61)	Low
Interest (higher is better)			
7	All estrogens vs. placebo	0.18 (0.10 to 0.26)	Moderate
2	SNRI vs. placebo	0.16 (-0.07 to 0.39)	Insufficient
5	Isoflavones vs. placebo	0.26 (-0.001 to 0.52)	Insufficient
Pain, interest, global			
10	Estrogen route a vs. route b	Not estimated	Moderate
Activity (higher is better)		**SSE/4 weeks (95% CI)**	
8	Testosterone (7 patch, 1 oral), all trials	1.17 (0.88 to 1.46)[a]	Moderate

CI = confidence interval; SMD = standardized mean difference; SNRI = serotonin-norepinephrine reuptake inhibitor; SSE = satisfying sexual episode; SSRI = selective serotonin reuptake inhibitor.

[a]Number of satisfying sexual episodes per four weeks

Urogenital Atrophy

One-quarter of trials reported urogenital atrophy outcomes—a primary outcome in 56.3 percent. A minority of the trials (20%) were assessed as good or fair quality. Ospemifene, an estrogen agonist/antagonist, was approved by FDA in February 2013 to treat moderate to severe dyspareunia in postmenopausal women. Evidence from three clinical trials showed that ospemifene improved symptoms of vulvar and vaginal atrophy. Although multiple scales were employed and heterogeneity noted in the pooled estimate for vaginal route of administration, the SOE was high that either oral or vaginal estrogens improve symptoms (Table E). The SOE was low for isoflavones.

The conclusions here are similar to those provided to clinicians[1] when considering treatment of symptoms that may be experienced by as many as 40 percent of postmenopausal women.[28] A 2006 Cochrane review including 19 trials concluded that vaginal or oral estrogens were similarly effective for treating vaginal atrophy symptoms.[29] These results, albeit indirectly based on placebo comparisons, indicate a greater magnitude of effect for vaginal compared with oral administration.

Table E. Magnitude and strength of evidence of treatments for urogenital atrophy: standardized mean differences from pairwise comparisons

Number of Comparisons	Comparators	Effect Size (SMD) (95% CI)	Strength of Evidence
3	Ospemifene vs. placebo	-0.75 (-1.05 to -0.45)	High
12	Vaginal estrogen vs. placebo	-0.44 (-0.65 to -0.23)	High
14	Nonvaginal estrogen vs. placebo	-0.35 (-0.44 to -0.26)	High
5	Isoflavones vs. placebo	-0.48 (-0.77 to -0.18)	Low

CI = confidence interval; SMD = standardized mean difference.

Sleep

Many trials ascertained self-reported sleep outcomes, but a single trial examined a drug approved by FDA for use in insomnia (eszopiclone). Compared with placebo, the SMD for improved sleep measures was approximately threefold greater with eszopiclone than with estrogens or any other agent. This suggests that modestly improved sleep accompanies other agents, including estrogens, used to treat menopausal symptoms (Table F). Of the trials reporting sleep outcomes, 11 percent were rated good or fair quality.

Although sleep disturbances during menopause are common,[30] how often they are secondary to menopausal symptoms is not well defined. Sedative hypnotic agents are not generally used to treat menopausal symptoms and so were not represented in the trials identified. Reported improvement in sleep evident with other agents such as estrogens is possibly due to treatment of vasomotor symptoms but requires evidence not considered here.

Table F. Magnitude and strength of evidence of treatments for sleep: standardized mean differences from pairwise comparisons

Number of Comparisons	Comparators	Effect Size (SMD) (95% CI)	Strength of Evidence
1	Eszopiclone vs. placebo	1.08 (0.53 to 1.62)	Not rated[a]
24	Estrogen vs. placebo	0.32 (0.24 to 0.46)	High
2	SSRI vs. placebo	0.46 (0.24 to 0.69)	Low
2	Gabapentin vs. placebo	0.33 (0.18 to 0.49)	Low
6	Isoflavones vs. placebo	0.37 (0.10 to 0.64)	Low
2	Ginseng vs. placebo	0.13 (-0.05 to 0.32)	Insufficient

[a]Eszopiclone, an oral sedative hypnotic used to treat insomnia, was included as a referent. With a single trial comparing eszopiclone with placebo, a rating could not be made.

CI = confidence interval; SMD = standardized mean difference; SSRI = selective serotonin reuptake inhibitor.

Compounded Hormone Therapies

Compounded hormone therapies are commonly prescribed, often in combination with some testing for hormone levels, with effectively no direct evidence base. We identified a single RCT examining pharmacokinetics in 40 women studied for 16 days.[31] No studies were identified examining the safety of the compounding practices for hormone therapies.

Limitations of the Evidence on Symptom Relief

The body of evidence synthesized for Key Question 1 was large, with many trials rated poor quality. However, the challenges of synthesizing this evidence extend far beyond trial quality to limitations incompletely incorporated in SOE assessments. These include—

- Use of different outcome scales or metrics
- Necessity of calculating SMDs and inherent difficulties estimating from publications
- Potential differences in populations represented by trial samples
- Potential for selective outcome reporting

Interpreting results when presented with continuous measures and multiple scales requiring the use of SMDs is challenging. It is difficult to infer proportions of women achieving minimal clinically important improvements.[32,33] Calculating SMDs is also not without challenges. There were a number of ways to obtain effect sizes from the continuous measures reported. Unbiased ANCOVA (analysis of covariance) effect estimates[13,34] were not typically reported, requiring the use of change score or sometimes end-of-followup comparisons.

A separate issue is that, although trial populations included women experiencing menopause, there were some differences in mean age, length of followup, and symptom severity. While the initial intent was to examine subgroups according to characteristics such as the presence of a uterus, lack of reporting did not allow us to do so. Results, then, apply to average women across all trials.

It is also difficult to evaluate potential selective outcome reporting from the included trials. Vasomotor symptoms were reported in about three-quarters of trials, but all other outcomes were reported in less than half. While some trials, such as those of sexual function or vaginal atrophy, were clearly not designed to primarily assess all outcomes, insignificant results may have gone unreported. For some of the outcomes reported, the outcome was stated as primary in only half of the studies. Results do not allow assessment of whether effects on different outcomes are independent.

We did not include studies examining effects among breast cancer survivors—women frequently affected by troublesome symptoms, including hot flushes. Although effects of nonhormonal agents on hot flushes may be similar regardless of breast cancer history, cancer survivors constitute a different patient population. Accordingly, these results are not intended to apply to those women. Further, the results are not intended to apply to women experiencing menopause at an early age due to ovarian insufficiency.

Other Benefits and Harms

Summary results are presented first for hormone therapies, then for nonhormone therapy preparations, followed by a discussion of limitations of the evidence base for other benefits and harms.

Menopausal Hormone Therapy Preparations

Evidence for this Key Question included the recent report for the USPSTF by Nelson and colleagues[35] and results from the Danish Osteoporosis Prevention Study (DOPS), which were published after the report by Nelson and colleagues. A majority of evidence in that report was derived from WHI trials, representing an older population without severe menopausal symptoms, but one that overlaps with the population for this review. Therefore, findings from large observational studies with younger populations were incorporated to inform the discussion on applicability. The picture of long-term effects emerges with some clarity, as summarized in Table G.

The USPSTF review reported differences in event rates with estrogen/progestin or estrogen compared with placebo. However, extrapolating absolute rates from the WHI samples to the target population of this review is problematic. In broad absolute terms, gallbladder disease is the most frequent occurrence, with thromboembolic events, stroke, and breast cancer less frequent.

Table G. Summary of long-term effects of menopausal hormone therapy preparations

Outcome	Risk	Treatment vs. Placebo	Strength of Evidence	Comment
Breast cancer	↑	Estrogen/progestin	High	
	↓	Estrogen	Low	Inconsistent
Gallbladder disease[a]	↑	Estrogen/progestin	Moderate	Consistency unknown with 1 trial
	↑	Estrogen	Moderate	Consistency unknown with 1 trial
Venous thromboembolic events[b]	↑	Estrogen/progestin	Moderate	Consistency unknown with 1 trial
	↑	Estrogen	High	
Stroke	↑	Estrogen/progestin	Moderate	Consistency unknown with 1 trial
	↑	Estrogen	High	
Ovarian cancer	↑	Estrogen/progestin	Low	Consistency unknown with 1 trial; imprecise with few cases
Colorectal cancer	↓	Estrogen/progestin	Low	Consistency unknown with 1 trial; imprecise with wide confidence interval
	—	Estrogen	Moderate	Consistency unknown with 1 trial
Coronary heart disease	↑	Estrogen/progestin	Moderate	Consistency unknown with 1 trial
	—	Estrogen	Moderate	Consistency unknown with 1 trial
Endometrial cancer	—	Estrogen/progestin	Moderate	Imprecise
Osteoporotic fractures	↓	Estrogen/progestin	Moderate	Inconsistency between 2 trials
	↓	Estrogen	Moderate	Consistency unknown with 1 trial

Risk: ↑increased, ↓ decreased, — no change
[a]Risk may be lower with transdermal estrogen administration.
[b]Risk may not be increased with transdermal estrogen administration.

Nonhormone Therapy Preparations

The evidence base informing other potential benefits and harms of nonhormone therapies in women treated for menopausal symptoms is limited but does not suggest that harmful long-term effects are likely for those agents studied (Table H). We identified large trials examining vitamin E, small trials of isoflavones, and observational studies evaluating antidepressants. Some studies of the long-term use of antidepressants did not distinguish risks for the different classes of agents used to treat symptoms and therefore did not meet our inclusion criteria. Although no salient long term benefits were identified, neither were safety signals apparent. However, given the large numbers of women potentially taking these agents, some caution is advised, particularly for nonprescription agents. For example, the possibility of increased mortality with high-dose vitamin E has been raised.[36] Additionally, case reports of hepatotoxicity with black cohosh have been published.[37] This association has been debated,[38] but surveillance for adverse effects of nonprescription agents is generally inadequate. Safety data are also needed for the broad array of herbs and botanicals used to treat menopausal symptoms.

Table H. Summary of long-term effects of nonhormone therapy preparations

Outcome	Risk	Treatment vs. Placebo	Strength of Evidence
Breast cancer	—	Vitamin E	High
Breast cancer		SSRI	Insufficient
Colorectal cancer	—	Vitamin E	High
Cardiovascular events	—	Vitamin E	High
Cardiovascular death	↓	Vitamin E	Low
Osteoporotic fractures	↑	SSRI	Low
Osteoporotic fractures		Isoflavones	Insufficient
Ovarian cancer		Vitamin E	Insufficient

Risk: ↑ increased, ↓ decreased, — no change. SSRI = selective serotonin reuptake inhibitor.

Limitations of the Evidence Base on Other Benefits and Harms

One limitation of the evidence base concerning long-term outcomes of hormone therapies derives from the necessity to rely on results of RCTs. There are well-described discrepant conclusions between observational studies and RCTs concerning long-term outcomes accompanying hormone therapies.[39] The discrepancies have been largely attributed to selection bias and time-varying confounding.[40-42] Although the association with cardiovascular outcomes has been most scrutinized, difficulties assessing causal effects of menopausal hormone therapy from observational data appear to extend to other outcomes, including hip fractures[40] and colorectal cancer.[42] As noted throughout, trials have been conducted in a target population overlapping with the one for this review, creating some challenges for assessing applicability.

There are several limitations to the evidence base of nonhormone therapies to consider. Many studies included women of all ages and therefore were excluded unless subgroup analyses on older women or menopausal women were specified. Much of the research available on the long-term effects of isoflavones and vitamin E consisted of population-based dietary studies and therefore did not meet inclusion criteria. Intermediate outcomes were reported in many of the studies: for example, bone density rather than osteoporotic fractures, and cholesterol levels rather than cardiovascular events. Finally, in studies that included all women rather than focusing on menopausal women, it was difficult to discern if exposure (e.g., to SSRIs/SNRIs or isoflavones) occurred during menopausal years.

Symptom Relief in Subgroups

A small subset of trials identified for Key Question 1 reported subgroup analyses on symptom relief: 10 for hormone therapies, 2 for nonhormone prescription therapies, and 4 for nonprescription therapies. No subgroup analyses could be pooled, as no two trials had the same comparators, definitions of subgroups, and outcomes. The sparse evidence did not allow rating of SOE.

Discussion

This section addresses research gaps, implications for clinical policy and decisionmaking, limitations of the Comparative Effectiveness Review process, and conclusions.

Research Gaps

The principal gaps in the evidence on symptom relief include lack of common validated instruments and assessment of meaningful clinical improvement; safety data on nonprescription agents; lack of evidence on compounded hormone therapies; and potential for predicting treatment response for nonhormonal agents:

- The trials comprising the body of evidence included in this review had in common the evaluation of outcomes on continuous scales using multiple instruments. A standard set of common data elements using validated instruments would facilitate evidence synthesis and interpreting results across trials. In place of, or in addition to, summary continuous effect measures, reporting differences in proportions of women achieving defined clinically meaningful improvements

would be more informative for decisionmaking. Reporting only summaries of continuous effect measures challenges interpretation for patients and providers.

- A large number of nonprescription agents were studied in individual trials. The Dietary Supplement Health and Education Act requires manufacturers of these agents to determine their products' safety and efficacy, but the manufacturers are not required to submit the safety or efficacy data to FDA. As women may elect to use these agents, the data need to become available.
- Millions of women use compounded hormone treatments. Yet there is a stark absence of evidence concerning compounded hormone therapies and the methods used to determine the personalized dosages. Although the gap is most concerning regarding safety, efficacy issues are important as well.
- For nonhormonal interventions for which there is moderate evidence of efficacy, identifying predictors of response would likely be helpful.

Many important previous gaps in the evidence concerning long-term effects of hormone therapies have been filled. For some nonhormone therapies, with reasonable certainty (i.e., moderate or greater SOE), significant safety issues have not been apparent; the same cannot be said for the entirety of the nonprescription agents.

Finally, estrogen therapy has efficacy relieving many symptoms but is accompanied by other potentially important harms (varying according to whether combined with progestogen). Given the number of outcomes to consider with different exposure effects (e.g., duration of use), the overall risk-benefit calculus is not simple. Juxtaposing evidence concerning symptom relief (as obtained here) with models for the long-term harms and potential benefits[43] according to patient characteristics (e.g., lower risk of hip fracture in blacks) could facilitate informed decisions by women and health care providers.

Implications for Clinical and Policy Decisionmaking

The implications of this review for clinical decisionmaking follow from better defining evidence supporting the multiple treatment options for different yet overlapping menopausal symptoms, each treatment option having different potential harms. The results provide a guide to comparative efficacy alongside potential long-term harms and benefits; all are weighed in clinical decisions.

For vasomotor symptoms and quality of life, the review provides clinicians with efficacy comparison for the most commonly used treatments. Although evidence concerning potential long-term benefits is included as they are part of the decisionmaking process, this review did not specifically address use of therapies for those purposes.

Limitations of the Comparative Effectiveness Review Process

This review was a large undertaking with many complexities. These included the variable manner in which trials reported results, multiple trial arms, and multiple treatments, along with the goal of not excluding results for any a priori potentially arbitrary reason. Obtaining standardized effects can be challenging.[44] Furthermore, given multiple trial arms and multiple outcomes, the number of calculations required was substantial. Confidence intervals and SOE ratings do not incorporate this analytical uncertainty. Pooled estimates should be interpreted with this understanding.

Analyses of the multiple treatments required imposing some classification scheme that has limitations. For example, the estrogen dose categorization scheme did not consider progestogen or distinguish between combined and sequential progestogen administration. Progestogen use was problematic to distinguish because trials may have not given the agent to women without a uterus yet reported an effect for the entire sample.

Finally, interpreting network and pairwise meta-analyses deserves comment. In the pairwise meta-analyses, only direct randomized comparisons are included; the network analyses incorporate both direct and indirect evidence. Underlying the network of comparisons is assumed similarity of study characteristics and patients (transitivity) as well consistency of effects throughout the network. All enrolled women were menopausal or perimenopausal, but there were some differences in studies and samples as noted in the review. However, across all studies the assumption was likely satisfied. The closeness of most network and pairwise estimates shows that inconsistencies are likely small.

Conclusions

Women experiencing symptoms of menopause can consider a number of potential treatments of varying efficacy. From a large body of evidence, there is considerable certainty that estrogens are the most effective treatment for relieving vasomotor symptoms and are

accompanied by the greatest improvement in quality-of-life measures. For other common symptoms—psychological, urogenital, and sleep disturbance—although estrogens are effective, some nonhormonal agents compare favorably. Estrogens are accompanied by potential long-term harms that require consideration. There is limited evidence on the potential consequences of long-term use of nonhormonal agents when those agents are used to treat menopausal symptoms.

References

1. North American Menopause Society. Menopause Practice: A Clinician's Guide. 4th ed., Mayfield Heights, OH: North American Menopause Society; 2010.
2. 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):1159-68. PMID: 22344196.
3. McKinlay SM, Brambilla DJ, Posner JG. The normal menopause transition. Maturitas. 1992 Jan;14(2):103-15. PMID: 1565019.
4. Nelson HD, Haney E, Humphrey L, et al. Management of menopause-related symptoms. Evid Rep Technol Assess (Summ). 2005 Mar;(120):1-6. PMID: 15910013.
5. Woods NF, Mitchell ES. Symptoms during the perimenopause: prevalence, severity, trajectory, and significance in women's lives. Am J Med. 2005 Dec 19;118(Suppl 12B):14-24. PMID: 16414323
6. Proceedings from the NIH State-of-the-Science Conference on Management of Menopause-Related Symptoms, March 21-23, 2005, Bethesda, Maryland, USA. Am J Med. 2005 Dec 19;118(Suppl 12B):1-171. PMID: 16414320.
7. Politi MC, Schleinitz MD, Col NF. Revisiting the duration of vasomotor symptoms of menopause: a meta-analysis. J Gen Intern Med. 2008 Sep;23(9):1507-13. PMID: 18521690.
8. Freeman EW, Sammel MD, Lin H, et al. Duration of menopausal hot flushes and associated risk factors. Obstet Gynecol. 2011 May;117(5):1095-104. PMID: 21508748.
9. Gold EB, Block G, Crawford S, et al. Lifestyle and demographic factors in relation to vasomotor symptoms: baseline results from the Study of Women's Health Across the Nation. Am J Epidemiol. 2004 Jun 15;159(12):1189-99. PMID: 15191936.
10. Organization MMaSS. Introductory Guide to MeDRA Version 14.0. Chantilly, VA: International Federation of Pharmaceutical Manufacturers and Associations; 2011.
11. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality; January 2014. Chapters available at www.effectivehealthcare.ahrq.gov.
12. Harris RP, Helfand M, Woolf SH, et al. Current methods of the US Preventive Services Task Force: a review of the process. Am J Prev Med. 2001 Apr;20(3 Suppl):21-35. PMID: 11306229.
13. Fu R, Vandermeer BW, Shamliyan TA, et al. Handling continuous outcomes in quantitative synthesis. In: Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality; January 2014. Chapters available at www.effectivehealthcare.ahrq.gov.
14. Hedges LV, Vevea JL. Fixed- and random-effects models in meta-analysis. Psychol Meth. 1998;3(4):486-504.
15. Rucker G, Schwarzer G, Carpenter JR, et al. Undue reliance on I(2) in assessing heterogeneity may mislead. BMC Med Res Methodol. 2008;8:79. PMID: 19036172.
16. Song F, Altman DG, Glenny AM, et al. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. BMJ. 2003 Mar 1;326(7387):472. PMID: 12609941.
17. Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. J C Clin Epidemiol. 2011 Apr;64(4):401-6. PMID: 21208779.
18. Treadwell JR, Uhl S, Tipton K, et al. Assessing equivalence and noninferiority. J Clin Epidemiol. 2012 Nov;65(11):1144-9. PMID: 22732455.
19. Furness S, Roberts H, Marjoribanks J, et al. Hormone therapy in postmenopausal women and risk of endometrial hyperplasia. Cochrane Database Syst Rev. 2009;(2):CD000402. PMID: 19370558.
20. Welton AJ, Vickers MR, Kim J, et al. Health related quality of life after combined hormone replacement therapy: randomised controlled trial. BMJ. 2008;337:a1190. PMID: 18719013.
21. Brunner RL, Gass M, Aragaki A, et al. Effects of conjugated equine estrogen on health-related quality of life in postmenopausal women with hysterectomy: results from the Women's Health Initiative Randomized Clinical Trial. Arch Intern Med. 2005 Sep 26;165(17):1976-86. PMID: 16186467.
22. Hays J, Ockene JK, Brunner RL, et al. Effects of estrogen plus progestin on health-related quality of life. N Engl J Med. 2003 May 8;348(19):1839-54. PMID: 12642637.
23. Freeman EW, Sammel MD, Lin H, et al. Associations of hormones and menopausal status with depressed mood in women with no history of depression. Arch Gen Psychiatry. 2006 Apr;63(4):375-82. PMID: 16585466.
24. Cohen LS, Soares CN, Vitonis AF, et al. Risk for new onset of depression during the menopausal transition: the Harvard study of moods and cycles. Arch Gen Psychiatry. 2006 Apr;63(4):385-90. PMID: 16585467.
25. Shifren JL, Monz BU, Russo PA, et al. Sexual problems and distress in United States women: prevalence and correlates. Obstet Gynecol. 2008 Nov;112(5):970-8. PMID: 18978095.
26. Avis NE, Brockwell S, Randolph JF Jr, et al. Longitudinal changes in sexual functioning as women transition through menopause: results from the Study of Women's Health Across the Nation. Menopause. 2009;16(3):442-52. PMID: 19212271.

27. Myers LS. Methodological review and meta-analysis of sexuality and menopause research. Neurosci Biobehav Rev. 1995 Summer;19(2):331-41. PMID: 7630587.

28. Krychman ML. Vaginal estrogens for the treatment of dyspareunia. J Sex Med. 2011 Mar;8(3):666-74. PMID: 21091878.

29. Suckling J, Lethaby A, Kennedy R. Local oestrogen for vaginal atrophy in postmenopausal women. Cochrane Database Syst Rev. 2006;(4):CD001500. PMID: 17054136.

30. National Sleep Foundation. 2007 Women and Sleep. http://sleepfoundation.org.

31. Sood R, Warndahl RA, Schroeder DR, et al. Bioidentical compounded hormones: a pharmacokinetic evaluation in a randomized clinical trial. Maturitas. 2013 Apr;74(4):375-82. PMID: 23384975.

32. Johnston BC, Patrick DL, Thorlund K, et al. Patient-reported outcomes in meta-analyses-part 2: methods for improving interpretability for decision-makers. Health Qual Life Outcomes. 2013;11:211. PMID: 24359184.

33. Guyatt GH, Thorlund K, Oxman AD, et al. GRADE guidelines: 13. Preparing summary of findings tables and evidence profiles-continuous outcomes. J Clin Epidemiol. 2013 Feb;66(2):173-83. PMID: 23116689.

34. Senn S. Statistical Issues in Drug Development. 2nd ed.; Chichester, England; Hoboken, NJ: John Wiley & Sons; 2007.

35. Nelson HD, Walker M, Zakher B, et al. Menopausal Hormone Therapy for the Primary Prevention of Chronic Conditions: Systematic Review To Update the 2002 and 2005 U.S. Preventive Services Task Force Recommendations. Evidence Synthesis No. 93. AHRQ Pub. No. 12-05168-EF-1. Rockville, MD; Agency for Healthcare Research and Quality; May 2012.

36. Miller ER 3rd, Pastor-Barriuso R, Dalal D, et al. Meta-analysis: high-dosage vitamin E supplementation may increase all-cause mortality. Ann Intern Med. 2005 Jan 4;142(1):37-46. PMID: 15537682.

37. Mahady GB, Low Dog T, Barrett ML, et al. United States Pharmacopeia review of the black cohosh case reports of hepatotoxicity. Menopause. 2008 Jul-Aug;15(4 Pt 1):628-38. PMID: 18340277.

38. Teschke R, Schmidt-Taenzer W, Wolff A. Spontaneous reports of assumed herbal hepatotoxicity by black cohosh: is the liver-unspecific Naranjo scale precise enough to ascertain causality? Pharmacoepidemiol Drug Saf. 2011 Jun;20(6):567-82. PMID: 21702069.

39. Prentice RL. Observational studies, clinical trials, and the Women's Health Initiative. Lifetime Data Anal. 2007 Dec;13(4):449-62. PMID: 17943443.

40. Tanaka S, Matsuyama Y, Shiraki M, et al. Estimating the effects of time-varying treatments: incidence of fractures among postmenopausal Japanese women. Epidemiology. 2007 Sep;18(5):529-36. PMID: 17700241.

41. Hernán MA, Alonso A, Logan R, et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. Epidemiology. 2008 Nov;19(6):766-79. PMID: 18854702.

42. Morois S, Fournier A, Clavel-Chapelon F, et al. Menopausal hormone therapy and risks of colorectal adenomas and cancers in the French E3N prospective cohort: true associations or bias? Eur J Epidemiol. 2012 Jun;27(6):439-52. PMID: 22644109.

43. Salpeter SR, Buckley NS, Liu H, et al. The cost-effectiveness of hormone therapy in younger and older postmenopausal women. Am J Med. 2009 Jan;122(1):42-52 e2. PMID: 19114171.

44. Gøtzsche PC, Hrobjartsson A, Maric K, et al. Data extraction errors in meta-analyses that use standardized mean differences. JAMA. 2007 Jul 25;298(4):430-7. PMID: 17652297.

Full Report

This executive summary is part of the following document: Grant MD, Marbella A, Wang AT, Pines E, Hoag J, Bonnell C, Ziegler KM, Aronson N. Menopausal Symptoms: Comparative Effectiveness of Therapies. Comparative Effectiveness Review No. 147. (Prepared by Blue Cross and Blue Shield Association Technology Evaluation Center Evidence-based Practice Center under Contract No. 290-2007-10058-I.) AHRQ Publication No. 15-EHC005-EF. Rockville, MD: Agency for Healthcare Research and Quality; March 2015. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

AHRQ Pub. No. 15-EHC005-1-EF
March 2015