QUERI

Nonpharmacologic Treatments for Menopause-associated Vasomotor Symptoms

July 2016

Prepared for:
Department of Veterans Affairs
Veterans Health Administration
Quality Enhancement Research Initiative
Health Services Research & Development Service
Washington, DC 20420

Prepared by:
Evidence-based Synthesis Program (ESP) Center
Durham VA Healthcare System
Durham, NC
John W. Williams, Jr., MD, MHSc, Director

Investigators:

Principal Investigators:
Karen M. Goldstein, MD, MSPH
Remy R. Coeytaux, MD, PhD
John W. Williams, Jr. MD, MHSc

Co-investigators:
Megan Shepherd-Banigan, PhD
Adam P. Goode, DPT, PhD
Jennifer R. McDuffie, PhD
Deanna Befus, BSN
Soheir Adam, MD
Varsha Masilamani, MBBS
Megan Van Noord, MSIS

Research Associates:
Avishek Nagi, MS
Liz Wing, MA

VA HEALTH CARE Defining EXCELLENCE in the 21st Century

PREFACE

The VA Evidence-based Synthesis Program (ESP) was established in 2007 to provide timely and accurate syntheses of targeted healthcare topics of particular importance to clinicians, managers, and policymakers as they work to improve the health and healthcare of Veterans. QUERI provides funding for four ESP Centers, and each Center has an active University affiliation. Center Directors are recognized leaders in the field of evidence synthesis with close ties to the AHRQ Evidence-based Practice Centers. The ESP is governed by a Steering Committee comprised of participants from VHA Policy, Program, and Operations Offices, VISN leadership, field-based investigators, and others as designated appropriate by QUERI/HSR&D.

The ESP Centers generate evidence syntheses on important clinical practice topics. These reports help:

- Develop clinical policies informed by evidence;
- Implement effective services to improve patient outcomes and to support VA clinical practice guidelines and performance measures; and
- Set the direction for future research to address gaps in clinical knowledge.

The ESP disseminates these reports throughout VA and in the published literature; some evidence syntheses have informed the clinical guidelines of large professional organizations.

The ESP Coordinating Center (ESP CC), located in Portland, Oregon, was created in 2009 to expand the capacity of QUERI/HSR&D and is charged with oversight of national ESP program operations, program development and evaluation, and dissemination efforts. The ESP CC establishes standard operating procedures for the production of evidence synthesis reports; facilitates a national topic nomination, prioritization, and selection process; manages the research portfolio of each Center; facilitates editorial review processes; ensures methodological consistency and quality of products; produces “rapid response evidence briefs” at the request of VHA senior leadership; collaborates with HSR&D Center for Information Dissemination and Education Resources (CIDER) to develop a national dissemination strategy for all ESP products; and interfaces with stakeholders to effectively engage the program.

Comments on this evidence report are welcome and can be sent to Nicole Floyd, ESP CC Program Manager, at Nicole.Floyd@va.gov.

Recommended citation: Goldstein KM, Coeytaux RR, Williams Jr JW, Shepherd-Banigan M, Goode AP, McDuffie JR, Befus D, Adam S, Masilamani V, Van Noord MG. Nonpharmacologic Treatments for Menopause-Associated Vasomotor Symptoms. VA ESP Project #09-009; 2016.

This report is based on research conducted by the Evidence-based Synthesis Program (ESP) Center located at the **Durham VA Medical Center, Durham, NC**, funded by the Department of Veterans Affairs, Veterans Health Administration, Office of Research and Development, Quality Enhancement Research Initiative. The findings and conclusions in this document are those of the author(s) who are responsible for its contents; the findings and conclusions do not necessarily represent the views of the Department of Veterans Affairs or the United States government. Therefore, no statement in this article should be construed as an official position of the Department of Veterans Affairs. No investigators have any affiliations or financial involvement (*eg*, employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties) that conflict with material presented in the report.

TABLE OF CONTENTS

TABLES

FIGURES

ABSTRACT

INTRODUCTION

Vasomotor symptoms (VMS), which include hot flashes and night sweats, are the most common symptoms reported during the menopausal transition and are experienced by as many as 80% of women. VMS can lead to increased healthcare encounters for symptom relief and reductions in quality of life. The degree to which VMS is bothersome is determined not only by how frequently it occurs but also by other factors such as duration of VMS, coexisting sleep problems, and the extent to which VMS interferes with daily activities or job-related activities. Hormone therapy (HT) is an effective treatment for reducing VMS, but use of this therapeutic approach must be individualized through weighing benefits with known risks, such as cardiovascular events or uterine and breast cancers. Based on age (> 45 years), currently half of the approximately 360,000 women Veterans who use Veterans Health Administration healthcare are perimenopausal or postmenopausal. Due in part to concerns about possible harms from long-term hormone therapy and in part to uncertain efficacy and safety of pharmacologic treatments, many women with VMS seek nonhormonal, nonpharmacologic treatment options.

METHODS

We developed the following research question in consultation with our stakeholders and a panel of technical experts:

> In women with vasomotor symptoms (VMS) that are associated with perimenopause or postmenopause, what are the effects on VMS, health-related quality of life, and adverse effects of the following nonpharmacologic, nonhormonal interventions: acupuncture; yoga, tai chi, and qigong; structured exercise; and meditation, mindfulness, hypnosis, and relaxation?

Given the multiple high-quality systematic reviews (SRs) in this topic area, we chose a method commonly known as an umbrella review, or review of reviews. We supplemented this approach by also evaluating randomized controlled trials (RCTs) published after the most recent good-quality SR for each of the eligible interventions. We grouped SRs and RCTs by intervention category. We graded the quality of SRs and new RCTs and prioritized the highest-quality SRs. Good-quality reviews had no important limitations and fair reviews had at least some important limitations. We reported intervention effects separately for inactive and active comparators and examined SRs for relevant subgroup analyses. When a new meta-analysis was indicated and feasible, we computed summary estimates of effect for each intervention. We explored the 2 primary outcomes: VMS and health-related quality of life. We assessed strength of evidence using the approach described in the Agency for Healthcare Research and Quality's *Methods Guide for Effectiveness and Comparative Effectiveness Reviews.*

RESULTS

We identified 239 unique SR citations; of these, 37 were retrieved for full-text review, and 10 were retained for data abstraction. We also identified 1135 unique RCTs; of these, 67 were retrieved for full-text review, and 14 articles plus 2 companion papers were retained for data

abstraction. The 10 SRs addressed acupuncture (n = 3), yoga (n = 1), structured exercise (n = 1), and meditation, mindfulness, hypnosis, relaxation, or mixed interventions (n = 5). The 14 new RCTs evaluated acupuncture (n = 4), yoga (n = 2), structured exercise (n = 3), and meditation, mindfulness, hypnosis, relaxation, or mixed interventions (n = 6). All studies reported the effects of interventions on VMS; quality of life and other outcomes were reported less frequently.

Acupuncture

We identified one good-quality SR published in 2013 that included women with VMS from any cause and 2 fair-quality SRs that included only women with any type of cancer. The good-quality SR included 15 relevant RCTs comprising 1127 participants. Comparison groups were no acupuncture, sham acupuncture, hormone therapy, or relaxation. Duration of acupuncture treatments ranged from 4 to 12 weeks across the RCTs. Traditional acupuncture was administered in 11 trials, electroacupuncture in 3 trials, and a combination of acupuncture and auricular acupuncture in a single trial. Twelve RCTs reported daily VMS by a self-report symptom diary. The remaining 3 assessed daily VMS frequency and severity as a hot flash score. Three RCTs assessed menopause-related symptoms considered by the authors of the SR to reflect quality of life. Applying the Cochrane guidance for overall risk of bias, 2 RCTs were rated as low risk and 11 were rated as high risk. This SR did not address durability of effect.

We identified 4 relevant RCTs with low risk of bias published since 2012 that were not included in the eligible SRs. Three examined the effect of acupuncture on VMS in perimenopausal and postmenopausal women, and one examined the effects of electroacupuncture in breast cancer survivors compared with the drug gabapentin, a placebo pill, or sham acupuncture. All 4 of these RCTs examined durability of effect. Three of the 4 new RCTS found that reductions in Hot Flash Composite Score and hot flash frequency were maintained at 12, 16, and 24 weeks posttreatment, respectively.

Our updated meta-analyses suggest that acupuncture is effective at reducing VMS frequency (SMD -0.66, 95% CI -1.06 to -0.26, I^2 = 61.7%, 5 trials) and VMS severity (SMD -0.49, 95% CI -0.85 to -0.13, I^2 = 18.1%, 4 trials), and improving quality of life (SMD 0.93, 95% CI -1.20 to -0.67, I^2 = 0.0%, 3 trials) when compared with no acupuncture, but not when compared with sham acupuncture. This translates to an average of 3 fewer hot flashes daily. Two of 4 trials reported that treatment benefit was maintained 16 to 24 weeks after end of treatment. The other 2 trials did not show sustained effects. Meta-analysis also suggests that acupuncture is associated with higher daily VMS frequency compared with hormone therapy (SMD 3.18, 95% CI 2.06 to 4.29, I^2 = 0%, 3 trials). Acupuncture and hormone therapy, however, are associated with similar changes in VMS severity (SMD 0.53, 95% CI -0.14 to 1.20, 2 trials).

Yoga, Tai Chi, and Qigong

We identified a single good-quality SR published in 2012 that included 5 RCTs (582 participants) evaluating the effect of yoga on VMS when compared with both inactive and/or active comparators. Yoga interventions were delivered in 34 to 40 sessions over 8 weeks to 4 months. All 5 interventions included yoga postures and meditation; 4 included breathing exercises, and 2 included lifestyle lectures. Risk of bias was determined to be low risk for 2 RCTs and high risk for 3 RCTs. Only one study included in the SR examined durability of effect

at 12 weeks post-treatment. We did not identify any eligible SRs that evaluated the effectiveness of tai chi or qigong.

We also identified 2 new relevant RCTs that assessed the impact of yoga on hot flash frequency and severity and quality of life among 102 perimenopausal or postmenopausal women, but both were judged to be of high risk of bias. In these, integral yoga or a Rusie Dutton course lasting 10 or 13 weeks was compared with inactive controls. Sessions were held weekly for 90 minutes. Outcomes were assessed using the Daily Hot Flash Diary and the Global Quality of Life Scale at 5 and 10 weeks in one RCT and subscales from the Menopause-specific Quality of Life instrument at 12 weeks in the other RCT. Neither of the new RCTs examined durability of treatment effects beyond the end of treatment. We did not identify any new RCTs that evaluated the effectiveness of tai chi or qigong.

The authors of the SR did not find an association between yoga and a reduction in VMS. We conducted an updated meta-analysis examining the effect of yoga compared with active and inactive controls on hot flash severity change scores using data from 2 new RCTs and 2 RCTs included in the SR. Results suggest that yoga is significantly associated with a reduction in hot flash severity (SMD -0.36, 95% CI -0.65 to -0.07, I^2 = 0.0%, 4 trials). Quality of life was not examined frequently. The single, small RCT that suggested some improvement in quality of life was statistically inconclusive.

Structured Exercise

We identified one good-quality and one fair-quality SR, both published in 2014, that included comparative evaluations of exercise, mindfulness, and relaxation. The good-quality SR included 5 RCTs that involved 762 sedentary, perimenopausal or postmenopausal women with VMS. In the majority of studies (n = 4), exercise was conducted with supervision or in a group session. The number and length of planned sessions ranged from 12 to 96 sessions over 12 to 36 weeks. The risk of bias was judged by the authors of the SR to be high risk for 4 RCTs and low risk for one RCT. The authors of the SR concluded there was no evidence that exercise reduced hot flash frequency or severity compared with inactive control (n = 4), yoga (n = 2), or hormone therapy (n = 1). Our updated meta-analyses did not find significant effects of structured exercise for either VMS frequency (SMD -0.08, 95% CI -0.33 to 0.16, I^2 = 0.0%, 4 trials) or severity (SMD -0.06, 95% CI -0.21 to 0.10, I^2 = 0.0%, 5 trials) compared with inactive control.

We identified 2 new RCTs that examined the effect of exercise compared with inactive control for VMS among perimenopausal or postmenopausal women. Both trials were judged to be low risk of bias. One studied women after breast cancer treatment exclusively. Exercise interventions were individually tailored programs focusing on aerobic exercise sessions lasting 40 to 90 minutes for 12 to 24 weeks. VMS was assessed using the Hot Flash Rating Scale in both trials. One RCT reported menopause-specific quality of life and found significantly lower sleep problem scores at 6 months' follow-up. The other trial found moderate benefit from structured exercise on health-related quality of life at end of treatment (effect size 0.46) and 12 weeks later (effect size 0.41).

Meditation, Mindfulness, Hypnosis, and Relaxation

We identified 5 SRs (2 good-quality, 3 fair-quality) that met eligibility criteria. One SR examined relaxation and mindfulness, one examined hypnosis exclusively, and 3 reported on mixed interventions. Of the RCTs included in the SRs, only one evaluated mindfulness-based stress reduction, and none assessed the effects of meditation. We focused on both the good-quality 2014 Cochrane Collaboration SR that addressed relaxation and paced respiration in 281 women with VMS and the fair-quality 2015 SR examining hypnosis in 274 women with breast cancer.

In the good-quality SR, 4 RCTs assessed VMS frequency, Kupperman Index scores, and mood. No studies reported on quality of life or adverse effects. Time of final follow-up assessment varied across RCTs from 4 weeks to 24 weeks with only one trial completing a follow-up assessment after completion of the intervention. No significant difference in reduction of hot flash frequency was found between any type of relaxation or mindfulness interventions compared with control groups for the primary outcomes of VMS frequency or quality of life. Authors of the SR did not report an overall risk of bias rating for individual studies but noted that the overall quality of evidence was “very low.”

In the fair-quality SR, 3 RCTs included the comparators of attention control, no treatment, and gabapentin. Outcomes measured included VMS frequency, health-related quality of life, mood, and adverse effects. The timing of outcome assessments was unclear for these trials. This SR presented qualitative findings only and found a significant decrease in hot flashes among women treated with hypnosis compared with no treatment; however, there was no difference when compared with gabapentin. Authors reported an overall low risk of bias in the included studies using the Cochrane Risk of Bias tool.

We identified 6 new RCTs that were not included in existing SRs. Three examined paced respiration and 3 examined effects of applied relaxation on VMS or quality of life among perimenopausal or postmenopausal women. Of the 3 new relaxation RCTs, one was low risk of bias, one was unclear risk, and one was high risk. We conducted a new meta-analysis comparing paced respiration with a control group. We found that paced respiration is not associated with a statistically significant decrease in hot flash frequency (SMD 0.04, 95% CI -0.73 to 0.82, I^2 = 56.6%) or severity (SMD 0.06, 95% CI -0.69 to 0.80, I^2 = 65.1%).

In general, our secondary outcomes of interest—sleep quality, depression, anxiety, and adverse effects—were rarely reported in the SRs addressing relaxation, and the results were mixed when they were reported.

CONCLUSION

Compared with waitlist controls, RCTs support acupuncture and yoga for reducing VMS and the impact of such symptoms on women’s activities and health-related quality life. The strength of evidence, however, is low to moderate. Moderately good evidence shows no benefit from structured exercise for VMS. The evidence in support of the effectiveness of mindfulness or relaxation is mixed, with some promising evidence that needs replication for hypnosis. There is insufficient evidence to draw conclusions about the effectiveness of these nonpharmacologic therapies for improving sleep, depression, or anxiety. The safety of the nonpharmacologic,

nonhormonal approaches evaluated in this report has not been rigorously examined, but there is no clear signal for a significant potential for harm. Overall, most of the data included in this report comes from smaller studies with homogenous participant populations. Larger trials of populations more reflective of the diversity of women experiencing VMS will be necessary to discern the effectiveness of nonpharmacologic interventions in symptomatic menopausal women.

ABBREVIATIONS TABLE

AHRQ	Agency for Healthcare Research and Quality
CI	Confidence interval
CINAHL	Cumulative Index to Nursing and Allied Health Literature
ESP	Evidence-based Synthesis Program
MD	Mean difference
MeSH	Medical Subject Heading
PICOTS	Population, intervention, comparator, outcome, timing, setting
RCT	Randomized controlled trial
SMD	Standardized mean difference
SOE	Strength of evidence
SR	Systematic review
VA	Veterans Affairs
VMS	Vasomotor symptoms
VHA	Veterans Health Administration

EVIDENCE REPORT

INTRODUCTION

Hot flashes and night sweats (or vasomotor symptoms) are 2 of the most common symptoms reported during the menopausal transition and are experienced by as many as 80% of women.[1] The mean age at onset of vasomotor symptoms (VMS) is 51 years,[2] and recent longitudinal studies of the menopausal transition revealed the median total duration can last more than 7 years.[3] Women's experience of VMS can vary based on personal characteristics. For example, more African-American women can expect to experience VMS than women from other racial/ethnic groups,[3] and those who gain weight in midlife are more likely to report hot flashes than those who do not.[4]

Prevalence and duration of VMS in perimenopausal and postmenopausal women are important women's health issues. VMS can lead to increased healthcare encounters for symptom relief[5] and reductions in quality of life.[6,7] The impact of VMS on quality of life can be worse for certain populations of women, such as those who undergo surgical rather than natural menopause.[8] Moreover, the degree to which VMS is bothersome is determined not only by how frequently it occurs but also by other factors such as duration of VMS, coexisting sleep problems,[9] and the extent to which VMS interferes with daily activities or job-related activities. For some defined populations of women, VMS exerts an especially strong, negative impact on quality of life; this appears to be the case for women Veterans.[6]

Hormone therapy (HT) is an effective treatment for reducing moderate to severe VMS, but use of this therapeutic approach must be individualized through weighing benefits with known risks, such as cardiovascular events or uterine and breast cancers.[10-12] FDA guidance defines moderate symptoms as "the sensation of heat with sweating, able to continue activity" and severe symptoms as those that cause "cessation of activity."[13] A common recommended strategy is to use HT with the "lowest effective dose for the shortest duration that is needed".[14,15] Additionally, HT is contraindicated for some women due to comorbid health conditions such as a history of breast cancer or being high risk for development of breast cancer, liver disease, or a history of blood clots. Based on age (> 45 years), currently half of the approximately 360,000 women Veterans who use Veterans Health Administration (VHA) healthcare are perimenopausal or postmenopausal. Many are at increased risk for complications of hormone therapy, such as cardiovascular disease, due to highly prevalent, known risk factors like obesity,[16] smoking,[17] and depression.[18] Despite this, recent evidence indicates that, compared with the general population of women in the United States, women Veterans using VHA are twice as likely to use hormone therapy for relief of menopausal symptoms.[19]

Due to the restriction on long-term hormone therapy use and the extended duration of expected VMS, many women with VMS are left in need of nonhormonal treatment options for many years. Thus, identifying effective nonhormonal interventions for improving quality of life among women Veterans with VMS is an important step to finding safe alternatives to hormone therapy and improved options for Veteran-centered care. Many perimenopausal and postmenopausal women are already using nonpharmacologic agents to manage VMS.[20-22] Nonpharmacologic treatments for VMS include herbal remedies (*eg*, black cohosh), mind and body practices (*eg*,

yoga, tai chi), structured exercise programs, and complementary and alternative medicine interventions (*eg*, acupuncture).[23]

In 2015, an Agency for Healthcare Research and Quality (AHRQ) systematic review[24] examined the comparative effectiveness of estrogens, isoflavones, selective serotonin reuptake inhibitors/serotonin-norepinephrine reuptake inhibitors, gabapentin, black cohosh, and ginseng for menopausal symptoms, including VMS. However, the AHRQ review did not address nonpharmacologic interventions such as mind and body practices, structured exercise programs, or complementary and alternative medicine interventions. Also in 2015, the North America Menopause Society released a position statement providing recommendations for many such intervention types and graded the level of evidence for their recommendations; however, this was not a formal systematic review of the literature.[25]

Our goal with this report is to summarize and update the evidence from systematic reviews (hereafter referred to as SRs) on selected nonpharmacologic approaches for the treatment of menopause-associated VMS and health-related quality of life.

METHODS

Given the multiple high-quality SRs in this topic area, we chose a method commonly known as an umbrella review, or review of reviews. We supplemented this approach by also evaluating randomized controlled trials (RCTs) published after the most recent good-quality SR for each of the eligible interventions. We followed methodological guidance from the Cochrane Collaboration[26] and AHRQ's Evidence-based Practice Centers.[27]

TOPIC DEVELOPMENT

Numerous interventions could be considered nonpharmacologic treatments for VMS in perimenopausal and postmenopausal women, particularly variations of complementary and alternative medicine approaches to symptom control.[20] To focus the selection of interventions for this review, we invited individuals with expertise in the field of menopause management both from within the Veterans Health Administration (VHA) and outside the VHA to participate in a technical expert panel. This panel provided consultation during the process of reviewing and organizing a list of potential interventions originally generated by the primary review team based on published literature and clinical practice. Table 1 lists the eligible nonpharmacologic interventions and definitions for this umbrella review.

Table 1. Eligible Interventions and Definitions

Intervention category	Definitions and examples
Acupuncture, acupressure	Acupuncture from any tradition was considered, including auricular acupuncture, electroacupuncture, acupressure, and laser acupuncture. Excluded were studies where acupuncture was administered in conjunction with Chinese herbal therapies. Cupping therapy was excluded unless it was a component of an acupuncture intervention.

Intervention category	Definitions and examples
Yoga, tai chi, qigong (as defined by study investigators)	Yoga typically involves combinations of physical exercises and bodily positions or postures, breath control practices, and meditation.
	Tai chi typically involves a series of movements performed in a slow, focused manner accompanied by deep breathing.
	Qigong typically involves a combination of coordinated body postures and movement, breathing, and meditation.
Structured exercise, physical activity	Structured exercise is defined as regular physical activity done with the intention of improving or maintaining physical fitness or health, or performed as part of a class or with support from a health professional.
Meditation, mindfulness, relaxation	Practices include: • Alexander technique • Benson's relaxation response • Bernstein and Borkovec's progressive relaxation • Everly and Rosenfeld's passive relaxation • Guided imagery-based approaches • Hypnosis • Madder's release • Mindfulness-based stress reduction • Mitchell method • Osts's applied relaxation • Paced respiration • Poppen's behavioral relaxation training • Progressive relaxation • Roll breathing • 4-7-8 breath technique • Other approaches that focus on diaphragmatic breathing

The final key question (KQ), developed in consultation with stakeholders, was:

> In women with vasomotor symptoms (VMS) that are associated with perimenopause or postmenopause, what are the effects on VMS, health-related quality of life, and adverse events of the following nonpharmacologic, nonhormonal interventions?
>
> · Acupuncture
> · Yoga, tai chi, and qigong
> · Structured exercise
> · Meditation, mindfulness, hypnosis, and relaxation

We followed a standard protocol for this review,[28] and each step was pilot-tested to train and calibrate study investigators. The PROSPERO registration number is CRD42016029335.[29]

SEARCH STRATEGY

In consultation with an expert librarian, we conducted searches of MEDLINE (via PubMed) and the Cochrane Database of Systematic Reviews from January 2010 through November 2015 for SRs, as well as searches of PubMed, EMBASE, CINAHL, and the Allied and Complementary Medicine Database from January 2012 through February 2016 for RCTs. We used MeSH

keywords and selected free-text terms for the interventions and health conditions of interest as well as terms for SRs and RCTs. The exact search strategies used are in Appendix A. All eligible SRs along with relevant umbrella and narrative reviews identified during citation screening were reviewed for additional relevant RCTs. All citations were imported into 2 electronic databases (for referencing, EndNote® Version X7, Thomson Reuters, Philadelphia, PA; for data abstraction, DistillerSR; Evidence Partners Inc., Manotick, ON, Canada).

STUDY SELECTION

Using prespecified inclusion/exclusion criteria (Table 2), titles and abstracts of SRs and RCTs identified through our primary search were reviewed by 2 reviewers for potential relevance to the key question. Articles included by either reviewer underwent full-text screening. At the full-text screening stage, 2 independent reviewers were required to agree on a final inclusion/exclusion decision. Disagreements were resolved by discussion or by a third investigator. Articles meeting eligibility criteria were included for data abstraction.

Table 2. Inclusion and Exclusion Criteria

Study Characteristic	Inclusion/Exclusion Criteria
Population	Perimenopausal and postmenopausal women who are experiencing bothersome VMS. Perimenopause is defined as amenorrhea for > 60 days in the past 12 months; postmenopause is defined as being without a menstrual cycle due to spontaneous or surgical reasons for the preceding 12 months. Bothersome VMS is defined as any of the following: · Self-identified "bothersome" hot flashes · Moderate to severe VMS as defined by the FDA[13] · Hot flashes that occur at least 6 days in the previous 2 weeks[3] · Hot flashes associated with functional impairment (*eg,* impairment in role, social, emotional, or physical functioning)
Interventions[a]	· Acupuncture, acupressure · Yoga, tai chi, qigong (as defined by study investigators) · Structured exercise, physical activity · Meditation, mindfulness, hypnosis, relaxation
Comparators	Any inactive control (waitlist, attention, sham acupuncture, information control, or unenhanced usual care) or active comparator (including hormone treatments, antidepressants or other pharmacotherapies, dietary supplements, health education, and alternative forms of exercise)
Outcomes	Primary: · Frequency and severity of VMS · Overall and condition-specific health-related quality of life Secondary: · Psychological well-being (*ie*, depressive or anxiety symptoms) and sleep quality, prioritizing validated scales over unvalidated scales or single-symptom measures · Serious adverse effects and adverse effects

Study Characteristic	Inclusion/Exclusion Criteria
Timing	· For SRs, timing of outcome assessments as specified by the review · For RCTs, outcomes assessed at > 60 days after treatment assignment
Setting	Outpatient or community settings (and mixed settings inclusive of outpatient/community settings)
Study design	Included: SRs and RCTs that evaluate an eligible intervention for VMS that is associated with perimenopause or postmenopause Excluded: SRs of complementary and alternative therapies in general without a specific focus on the interventions of interest, and umbrella reviews
Publication type	Included: Full articles published in peer-reviewed journals from January 2010 to November 2015 for SRs, or from January 2012 to February 2016 for RCTs Excluded: Abstracts and dissertations

[a] Intervention definitions are in Table 1.
Abbreviations: RCT = randomized controlled trial; SR = systematic review; VMS = vasomotor symptoms

DATA ABSTRACTION

Data from published fair- or good-quality SRs and newly identified RCTs were abstracted into a customized DistillerSR database by one reviewer and overread by a second reviewer. Disagreements were resolved by discussion or by a third investigator. Data elements include descriptors to characterize the type of study, study population, intervention, comparator, outcomes reported, study quality, and author conclusions.

QUALITY ASSESSMENT

Two reviewers assessed the quality of SRs and RCTs (Appendix B). Disagreements were resolved by discussion or by a third investigator. For SRs, we used the following key quality criteria, adapted from ROBIS[30] and AMSTAR[7]: search methods adequate for replication and comprehensive; selection bias avoided; data abstracted reliably; characteristics of primary literature reported and quality assessed appropriately; results synthesized using appropriate methods; publication bias assessed; conflict of interest reported; and conclusions supported by results. Based on these criteria, SRs were categorized as good, fair, or poor quality. Good- and fair-quality SRs should provide sufficient information to assess the strength of the body of evidence using the GRADE criteria.[31,32] Poor-quality reviews were excluded from our review.

For newly identified RCTs, we used the Cochrane Collaboration's risk of bias (ROB) tool,[33] which categorizes biases by 6 domains (selection bias, performance bias, detection bias, attrition bias, reporting bias, and other bias) and makes a summary assessment. For each item, a summary rating (high, low, or unclear ROB) was made (Appendix C).

DATA SYNTHESIS

We grouped the SRs and RCTs by intervention. We prioritized SRs that are the highest quality and most relevant to the study question.[34] Relevance was assessed using the PICOTS (population, intervention, comparator, outcome, timing, setting) framework along with the search

date, review methods, and completeness of reporting. For each intervention, tables or graphical displays describe the studies included, study quality, and treatment effects. We report intervention effects separately for inactive and active comparators. We examined SRs for relevant subgroup analyses, including concurrent use of hormone therapy, effects in women with and without breast cancer, perimenopausal and postmenopausal women, and women with surgical versus natural menopause.

Although umbrella reviews typically do not search for new primary studies, our review incorporated this step in order to identify important new data. As a framework for determining whether a new quantitative synthesis is needed, we considered the number of new studies, the sample size, and the strength of evidence (SOE) domains.[31] When an updated or new meta-analysis was indicated and feasible, we computed summary estimates of effect for each intervention using end-of-treatment outcomes. When means and measures of dispersion were not reported in the text, they were approximated from figures with the use of Engauge Digitizer.[35] We imputed missing control group standard deviations (SDs) in one study (and for 2 outcome measures) by using the maximum of all available SDs from other arms in the same study, which were measured on the same scale.

One difficulty in performing new meta-analyses for the quality of life outcome was the large variation in the types of scales used to measure this variable. For example, a scale could have separate subscales labeled VMS, physical, and somatic, all of which we considered part of physical domain but not address the social, occupational or functional aspects, whereas other scales contained occupational, social, and emotional subscales, but did not address the physical domain.

We devised a working definition for health-related quality of life in order to bring some degree of homogeneity to our analyses. To be included, a scale must have had at least 2 of the following 5 domains that were determined to be major: emotional, functional, occupational, physical, and/or social. These 5 domains were found in 50% or more of all the scales represented in the studies that examined quality of life. Of the eight scales used in the studies included in our meta-analyses, only one scale, the MENQOL, contained only 2 of the 5 domains and only one scale, the SF-36, contained all 5 domains. All others contained at least 3 of the 5 major domains used in our working definition.

We used R version 3.1.2 (R Foundation for Statistical Computing, Vienna, Austria) with the metafor package[36] to calculate random-effects model summary estimates of effect with Knapp and Hartung adjustment of standard errors of the estimated coefficients.[37,38] Because outcomes were measured across the trials using different instruments, outcome measures were combined using standardized mean differences (SMDs) in a random-effects model. The SMD was calculated by dividing the difference in mean values by the pooled SDs of the 2 groups. SMDs of 0.2 can be considered small treatment effects; 0.5, moderate effects; and ≥0.8, large effects.[39] For hot flash frequency, we transformed the pooled SMD[26] into hot flashes/day by multiplying the SMD by the standard deviation of the original scale from a representative study. Consistency of findings across individual studies was assessed by standard chi-square tests and the I^2 statistic.

RATING THE BODY OF EVIDENCE

SOE was assessed using the approach described in the Agency for Healthcare Research and Quality (AHRQ)'s *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*.[27] We focused on the 2 primary outcomes of impact on VMS and health-related quality of life. The AHRQ approach requires assessment of 4 domains: risk of bias, consistency, directness, and precision. Additional domains are used when appropriate: coherence, dose-response association, impact of plausible residual confounders, strength of association (magnitude of effect), and publication bias. These domains were considered qualitatively, and a summary rating of high, moderate, or low SOE was assigned after discussion by 2 investigators. In some cases, a rating of high, moderate, or low was impossible or imprudent to make. In these situations, a rating of insufficient was assigned. This 4-level rating scale consists of the following definitions:

- High—We are confident that the true effect lies close to the estimate of effect.
- Moderate—We are moderately confident of the effect estimate. The true effect is likely to be close to the estimate of effect, but there is a possibility that it is substantially different.
- Low—Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of effect.
- Insufficient—We have very little confidence in the effect estimate. The true effect is likely to be substantially different from the estimate of effect.

PEER REVIEW

This report was reviewed by technical experts and clinical leadership. A transcript of their comments and our responses is provided in Appendix E.

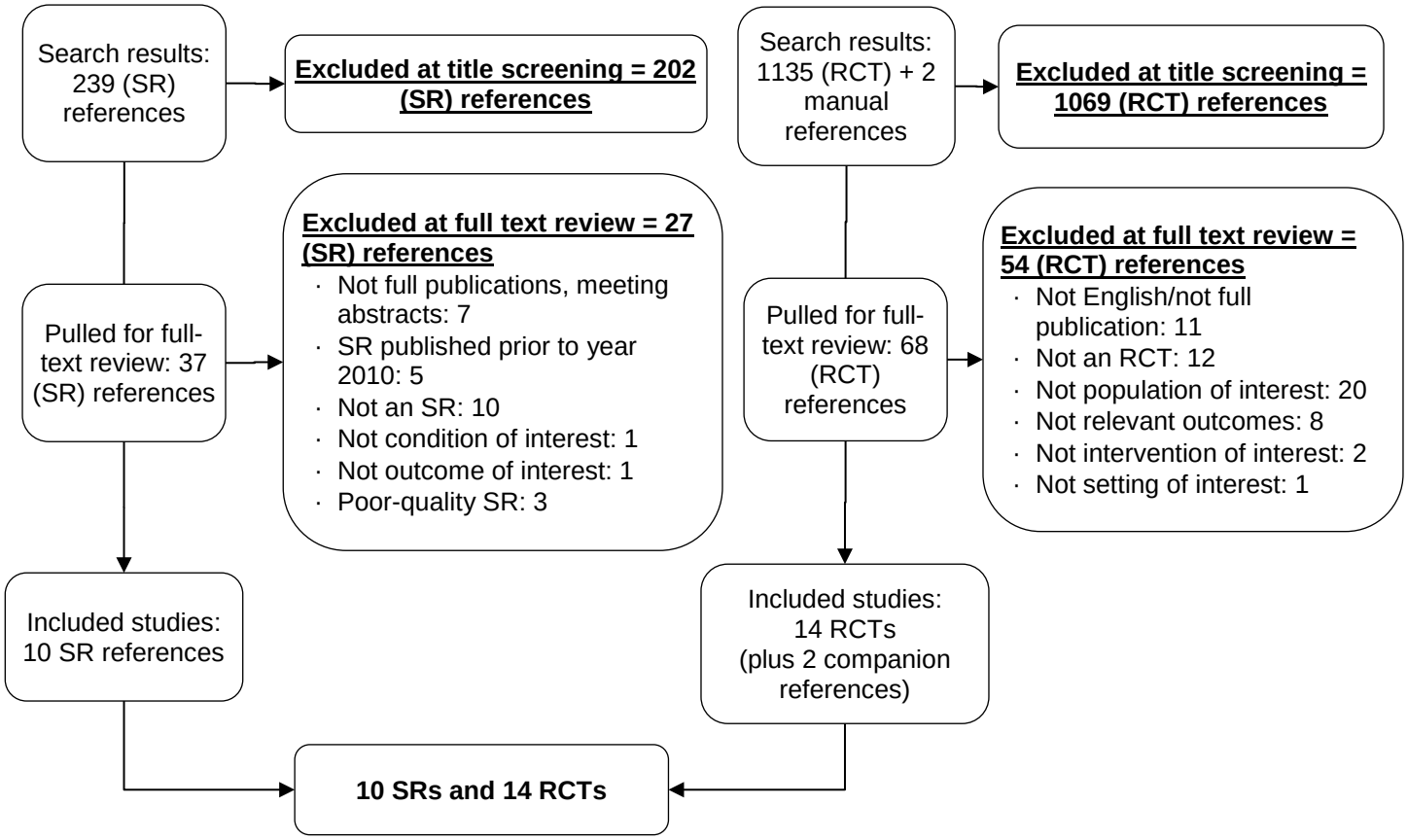
RESULTS

LITERATURE FLOW

Figure 1 shows the flow of articles through the literature search and screening process. The literature search for SRs identified 239 unique citations from a combined search of PubMed and the Cochrane Database of Systematic Reviews. After applying inclusion and exclusion criteria at the title-and-abstract screening level, 37 full texts were retrieved for further review. Of these, 10 were retained for data abstraction.

A separate literature search for RCTs published after January 2012 identified 1135 unique citations from a combined search of PubMed, EMBASE, CINAHL, and the Allied and Complementary Medicine Database. One additional citation was identified by review of study bibliographies or contact with authors. After applying inclusion and exclusion criteria at the title-and-abstract screening level, 68 full texts were retrieved for further review. After eliminating studies included in eligible reviews and applying eligibility criteria at the full-text review level, 14 were retained for data abstraction, and 2 were identified as companion papers (*ie*, secondary analyses to studies included in an eligible review).

Figure 1. Literature Flow Chart



CHARACTERISTICS OF INCLUDED STUDIES

Systematic Reviews

We identified 10 eligible SRs (5 good quality and 5 fair quality): 3 addressed acupuncture;[40-42] 1 addressed yoga;[43] 1 addressed structured exercise;[44] and 5 addressed meditation, mindfulness-based practices, hypnosis, relaxation, or mixed interventions.[45-49] All SRs evaluated interventions in women who were perimenopausal or postmenopausal and 5 required women to report bothersome frequency of VMS. Four restricted their focus to women with a history of breast cancer. Seven SRs included only RCTs. Four reported summary estimates of treatment effect from meta-analyses, while 6 synthesized results qualitatively.

New Randomized Controlled Trials

We identified 14 eligible RCTs published after the SRs (hereafter described as "new RCTs") that evaluated a range of interventions: 4 addressed acupuncture;[50-53] 2 addressed yoga;[54,55] 6 addressed relaxation, paced respiration, or hypnosis;[56-61] and 2 addressed structured exercise.[62,63] All new RCTs enrolled perimenopausal or postmenopausal women with VMS. Three RCTs focused on breast cancer survivors. Trials randomized 40 to 327 (median = 63 patients and reported outcomes at 5 to 24 (median = 11) weeks. All RCTs reported the effects of interventions on VMS; however, quality of life and other outcomes were reported less frequently.

For each of the 4 intervention categories, we focus our discussion on the SR(s) having the highest quality and most relevance to the primary outcomes of interest: VMS symptoms and health-related quality of life. Results from other SRs are described when they provided unique information. We then discuss results from the new RCTs and our updated syntheses (quantitative or qualitative) that included these additional studies. Results are further organized by inactive and active comparators. Last, we describe the findings for the secondary outcomes across all 4 intervention categories, which include adverse effects as well as effects on sleep, depression, and anxiety.

KQ: In women with vasomotor symptoms (VMS) that are associated with perimenopause or postmenopause, what are the effects on VMS, health-related quality of life, and adverse events of the following nonpharmacologic, nonhormonal interventions?

- **Acupuncture**
- **Yoga, tai chi, and qigong**
- **Structured exercise**
- **Meditation, mindfulness, hypnosis, and relaxation**

ACUPUNCTURE

Key Points

- We identified one good-quality SR published in 2013 and 2 fair-quality SRs published in 2013 and 2015 that met eligibility criteria. The good-quality review included women with menopausal VMS from any cause, and the 2 fair-quality reviews included only women with any type of cancer. All 3 SRs included RCTs that compared acupuncture with no acupuncture, sham acupuncture, hormone therapy, or relaxation.

- The Dodin et al meta-analysis suggests that acupuncture is effective at reducing VMS frequency (SMD -0.66, 95% CI -1.06 to -0.26, I^2 = 61.7%, 5 trials) and VMS severity (SMD -0.49, 95% CI -0.85 to -0.13, I^2 = 18.1%, 4 trials) and improving quality of life (SMD -0.93, 95% CI -1.20 to -0.67, I^2 = 0.0%, 3 trials) when compared with no acupuncture. This translates to an average of 3 fewer hot flashes daily.

- Updated meta-analysis suggests that acupuncture is not effective at reducing VMS frequency (SMD -0.21, 95% CI -0.49 to 0.07, I^2 = 53.5%, 10 trials) compared with sham acupuncture, but there is a modest, positive effect associated with acupuncture for reducing VMS severity compared with sham acupuncture (SMD -0.35, 95% CI -0.70 to 0.01, I^2 = 66.5%, 8 trials). Acupuncture does not appear to be more effective than sham acupuncture in improving quality of life as assessed by symptom inventories (SMD -0.23, 95% CI -1.40 to 0.95, I^2 = 77.0%, 5 trials).

- Meta-analysis suggests that acupuncture is associated with much higher daily VMS frequency compared with hormone therapy (SMD 3.18, 95% CI 2.06 to 4.29, I^2 = 0%, 3 trials). This translates to 13 more hot flashes daily. Acupuncture and hormone therapy, however, did not differ in effects on VMS severity (SMD 0.53, 95% CI -0.14 to 1.20, 2 trials), but trials were small and 95% CI wide.

- We identified 4 relevant, low risk of bias RCTs published since 2012 that were not included in the eligible SRs. Two RCTs compared 8- or 12-week courses of acupuncture with sham acupuncture among 327 and 20 women with VMS; one compared 6 months of acupuncture with a waitlist control among 209 women with VMS; and one 4-arm trial randomized 120 breast cancer survivors to 8 weeks of electroacupuncture versus gabapentin versus sham acupuncture versus placebo gabapentin.

Systematic Reviews

Study characteristics

We identified 3 eligible SRs evaluating the effectiveness of acupuncture for VMS and quality of life. One was a good-quality Cochrane Collaboration SR by Dodin et al[40] published in 2013, and 2 were fair-quality SRs that were qualitative summaries of RCTs published in 2013[64] and 2015[41] evaluating acupuncture among women with cancer (primarily breast cancer) who reported VMS. The Dodin et al SR included all but 3 of the individual RCTs that were also included in the 2 fair-quality SRs. Dodin et al excluded one RCT because of insufficient information provided about VMS experienced by the participants,[65] and 2 RCTs were indexed in the literature databases after the search date of January 15, 2013.[66,67]

For the purposes of this report, we focus on the SR by Dodin et al[40] because it did not limit inclusion to trials of women with a history of cancer, it included all but 3 RCTs also included in the other 2 SRs, it conducted meta-analyses pertinent to this report, and it was judged to be of higher quality than the other 2 SRs. Additionally, the other SRs did not provide a quantitative synthesis of findings.

Dodin et al included 16 RCTs[68-83] comparing any type of acupuncture with no acupuncture, sham acupuncture, or other treatments for reducing menopausal hot flashes and improving the quality of life of symptomatic perimenopausal and postmenopausal women. Of the 16 RCTs, one is not pertinent to our report because the study intervention was moxibustion alone,[70] which is not technically acupuncture. Collectively, the 15 relevant RCTs included 1127 women. The minimum number of daily VMS for inclusion in the trials was 2 to 7 in the 8 RCTs that reported this inclusion criterion. All RCTs included perimenopausal or postmenopausal women. Five exclusively enrolled women who had recently completed treatment for breast cancer, and one included only women who had undergone bilateral oophorectomy.

Duration of acupuncture treatments ranged from 4 to 12 weeks across the RCTs. Traditional acupuncture was administered in 11 trials, electroacupuncture in 3, and a combination of acupuncture and acupressure in one RCT. Standardized or semistandardized acupuncture treatment protocols were employed in 12 RCTs, with the number of acupuncture points needled at each treatment session ranging from 4 to 13. Individualized treatments were administered in

the other 3 RCTs. Although 6 of the 15 applicable trials reported durability of effect at 3 to 12 months, the SR by Dodin et al[40] did not address durability of treatment effect.

Twelve RCTs reported daily VMS as assessed by a self-report symptom diary. The remaining 3 also assessed daily VMS frequency but reported the findings as a Hot Flash Score, calculated by multiplying the number of VMS per day by self-reported VMS severity. Three RCTs assessed health-related quality of life using the Menopause-specific Quality of Life (MENQOL) instrument, SF-36, or a visual analog scale. Sleep quality or quantity was assessed in 4 RCTs using the Women's Health Initiative Insomnia Rating Scale, Pittsburgh Sleep Quality Index, or number of hours of sleep per night. The Kupperman Index was used in 3 RCTs to assess menopause-related symptoms. Two RCTs assessed symptoms of depression using the Beck Depression Inventory, 2 assessed symptoms of anxiety, one used the Greene Climacteric Scale, and one used the Women's Health Questionnaire. Authors of the SR considered the MENQOL, the Kupperman Index, the Greene Climacteric Scale, and the Women's Health Questionnaire as measures of quality of life.

Characteristics of the SR and included RCTs are summarized in Table 3.

Table 3. Study Characteristics: Dodin et al 2013[40]

Characteristic	Value
Systematic review	
Number of included trials	15[a]
Number of patients	1127
Date of SR literature search	January 2013
Age range, years	40-76
RCTs included in the SR	
Study years	2006-2013
Countries	
USA	6
Sweden	3
South Korea	2
Norway	2
China	1
Denmark	1
Population	
Perimenopausal or postmenopausal	13
Women who had completed treatment for breast cancer	1
Women who had ovaries surgically removed	1
Acupuncture interventions	
Traditional acupuncture	11
Electroacupuncture	3
Acupuncture plus acupressure	1
Planned number of acupuncture treatments	5-36
Length of intervention period	4-12 weeks
Training of persons administering acupuncture	
Licensed acupuncturist	7
Physiotherapist	4
Traditional Korean medical doctor	2
Not reported	2

Characteristic	Value
Comparisons[b]	
No acupuncture or usual care only	5
Sham acupuncture	10
Hormone therapy	3
Relaxation	1
Outcomes	
Daily VMS diary	15
VMS frequency	12
VMS severity	5
Hot flash score (frequency times severity)	3
Kupperman Index	3
MENQOL	2
Sleep quality (various measures)	4
Timing of last outcome assessment after randomization	
3 days after each treatment	1
8-24 weeks	10
6-24 months	4

[a] We excluded one of the 16 RCTs because the study intervention was moxibustion, which is not technically acupuncture.
[b] Some RCTs included 3 arms and thus had 2 comparators.

Risk of bias—Systematic review

In the SR, authors assessed each included RCT using the Cochrane Risk of Bias Tool across 6 domains (2 selection bias domains, detection bias, attrition bias, reporting bias, and other). Of the 12 trials that contributed data to meta-analyses reported in the SR, 10 were judged by the authors to have 4 or more low risk of bias domains (out of 6) and 2 were judged to have 3 low risk of bias domains. Applying the Cochrane guidance for assessing overall risk of bias, 2 RCTs would be rated as low risk and 10 would be rated as high risk.

New Randomized Controlled Trials

Study characteristics

We identified 4 relevant RCTs published after 2012 that assessed the impact of acupuncture on VMS or quality of life among perimenopausal or postmenopausal women.[50-53] Avis et al[52] compared acupuncture with no acupuncture for the treatment of VMS. In this RCT, 209 perimenopausal or postmenopausal women who reported an average of 4 or more VMS per day were randomized to receive up to 20 acupuncture treatments within 6 weeks or to a waitlist control. In a 4-arm trial, Mao et al[51] compared acupuncture with sham acupuncture and randomized 120 breast cancer survivors experiencing bothersome hot flashes to gabapentin versus placebo versus 10 treatments (8 weeks) of electroacupuncture, compared with sham acupuncture. Ee et al[50] randomized 327 perimenopausal and postmenopausal women experiencing bothersome hot flashes to 10 treatments (8 weeks) of acupuncture or sham acupuncture. Nedeljkovic et al randomized 40 postmenopausal women with at least 20 hot flashes per week to 12-weeks of acupuncture, sham acupuncture, Chinese herbal medicine, or placebo.

Outcome measures

In Avis et al[52] primary outcomes were daily hot flash frequency and severity, as measured by a daily hot flash diary. Secondary outcomes were hot flash interference with daily life, sleep quality, depressive symptoms, somatic and other symptoms, anxiety, and quality of life assessed at end of treatment and 6 months post-treatment. In Mao et al[51] the primary outcome was the once per week average hot flash composite score as measured by a daily hot flash diary. Participants documented their hot flash frequency and severity each day starting from baseline until end of intervention and then for 1 week at 12 and 24 weeks (16 weeks post-treatment). In Ee et al,[50] the primary outcome was a hot flash score reflecting frequency and severity at the end of treatment. Secondary outcomes included quality of life, anxiety, depression, and adverse events. Participants recorded the number of daily mild, moderate, severe, and very severe hot flashes for 7 days using a validated hot flash diary. Participants were assessed for secondary outcomes at 4 weeks, the end of treatment, and then 3 and 6 months after the end of treatment. Hot flash frequency and severity, measured with a daily diary at 12 and 24 weeks, were the primary outcomes in the study by Nedeljkovic et al. Menopause-related quality of life was assessed using the Menopause Rating Scale-II.

Characteristics of the 4 new RCTs are summarized in Table 4.

Table 4. Study Characteristics of New RCTs

Study Country	**Population** # Women randomized Type of menopause # Hot flashes Mean age in years (range)	**Intervention** Category/type Session frequency/duration	**Comparator** Category/type Session frequency/duration	**General Outcomes** Instruments[a]
Ee 2016[50] Australia	327 Perimenopausal and postmenopausal ≥7 moderate per day Intervention: 55.2 (4.3) Comparator: 54.8 (4.2)	Acupuncture 10 treatments over 8 weeks	Sham acupuncture (Park Sham device) 10 sessions over 8 weeks	8 and 24 weeks: · Hot flash frequency · Hot flash severity · Hot flash score · Menopause QOL · Hospital Anxiety and Depression Scale · Adverse events
Avis 2016[52] USA	209 Perimenopausal and postmenopausal ≥4 per day 53.8 (3.5)	Acupuncture Up to 20 treatments over 24 weeks	Waitlist control: usual care for 6 months followed by 6-month course of acupuncture treatments	Several time points up to 26 weeks: · Hot flash frequency · Hot flash severity · Hot flash Daily Interference · Pittsburgh Sleep Quality Index and PROMIS short form Sleep Disturbance · Women's Health Questionnaire · Center for Epidemiologic Studies Depression Scale · General Anxiety Disorder Scale and PROMIS short form Anxiety · Perceived Stress Scale · Medical Outcomes Study short form, Physical and Mental Health Component scores

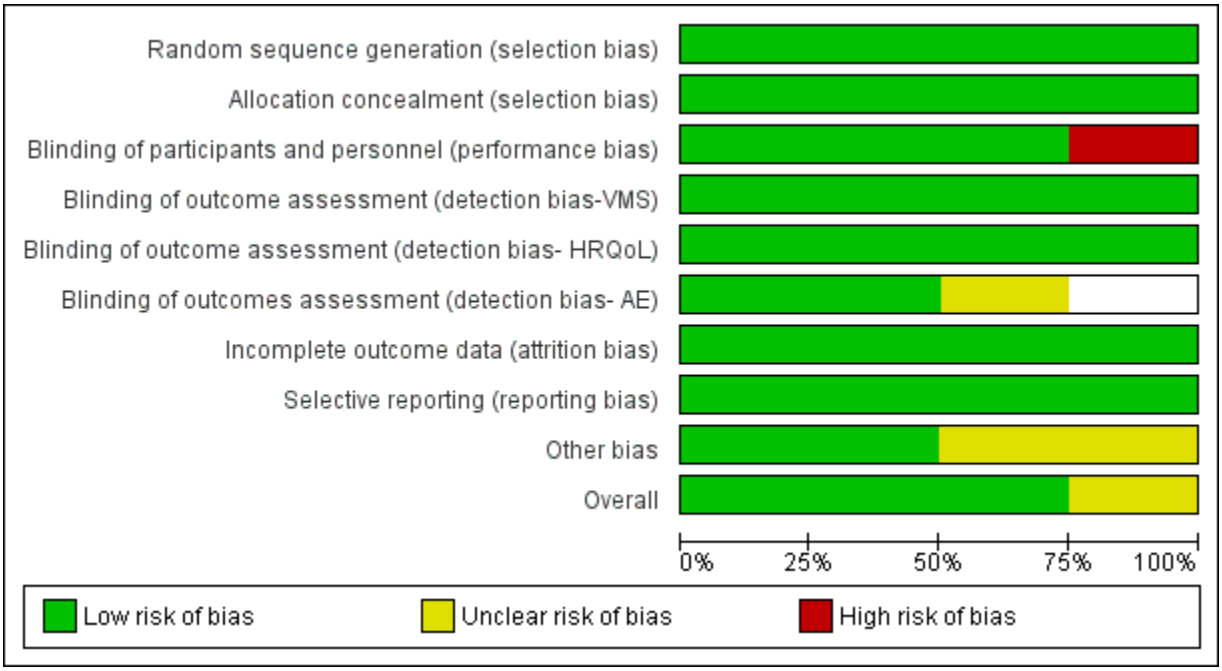
Study Country	**Population** # Women randomized Type of menopause # Hot flashes Mean age in years (range)	**Intervention** Category/type Session frequency/duration	**Comparator** Category/type Session frequency/duration	**General Outcomes** Instruments[a]
Mao 2015[51] USA	120 Breast cancer survivors ≥2 per day of bothersome severity 53.4 (3.4)	Electroacupuncture 10 treatments over 8 weeks	(1) Sham acupuncture 10 sessions over 8 weeks (2) Gabapentin 900 mg/day for 8 weeks (3) Placebo pill Gelatin capsules filled with lactose monohydrate	8 and 24 weeks: · Hot flash composite score · Adverse events
Nedeljkovic 2014[53] Switzerland	40 Postmenopausal ≥20 per week 53.1 (3.5)	Acupuncture 12 weekly treatments	(1) Sham acupuncture, 12 weekly sessions (2) Chinese herbal medicine, 3 capsules twice daily and 4 clinic visits (3) Placebo, 3 capsules twice daily and 4 clinic visits	12 and 24 weeks: · Hot flash frequency · Hot flash severity · Menopause Rating Scale-II

[a] Menopause instruments and scales are described in Appendix D.

Risk of bias—New RCTs

All 4 new RCTs were judged to be low risk of bias. Risk of bias domains for the 4 RCTs are reported in Figure 2.

Figure 2. Risk of Bias Ratings for New Acupuncture RCTs[a]



[a] A white bar indicates that the outcome was not reported.

Synthesis of Findings: Primary Outcomes

Vasomotor symptoms

Acupuncture versus no acupuncture or usual care

In the SR, Dodin et al[40] conducted a meta-analysis of 3 trials that compared acupuncture with no acupuncture, the results of which demonstrated a statistically significant effect associated with acupuncture for both VMS frequency (SMD -0.50, 95% CI -0.69 to -0.31, I^2 = 0.0%) and severity (SMD -0.54, 95% CI -0.73 to -0.35, I^2 = 0.0%) compared with no acupuncture. Subgroup and sensitivity analysis were not performed. A meta-analysis reported in the review also suggests that quality of life, as assessed by the Women's Health Questionnaire, Menopausal Rating Scale, or the Menopause Specific Quality of Life Questionnaire, is improved with acupuncture (SMD -0.93, 95% CI -1.20 to -0.67, I^2 = 32%, 3 trials) compared with no acupuncture. Each of the 3 trials included in these meta-analyses was judged by the authors of the SR to have low risk of bias for 5 of the 6 Cochrane Collaboration's Risk of Bias domains assessed. However, confidence intervals reported for pooled estimates of effect in this review are likely overly precise because Dodin et al did not use methods to account for small sample effects.[84]

We updated the meta-analysis reported in the SR by adding both the results at the end of treatment from Avis et al[52] and the results of the waitlist control group at the end of treatment from the 3-arm RCT by Avis et al[80] in the SR by Dodin (which was not included in the meta-analyses of acupuncture versus no acupuncture in that SR). Our updated meta-analysis generated an estimate of the SMD of -0.66 (95% CI -1.06 to -0.26; I^2 = 61.7%, 5 trials) for reduction in VMS frequency (Figure 3) and SMD -0.49 (95% CI -0.85 to -0.13, I^2 = 18.1%, 4 trials) for reduction in VMS severity (Figure 4). This translates to an average of 3 fewer hot flashes daily. Acupuncture resulted in a statistically significant decrease in hot flash frequency at 6 months that was maintained until 12 months after baseline.[52]

Figure 3. Forest Plot of Acupuncture versus No Acupuncture on Change in Hot Flash Frequency at End of Treatment

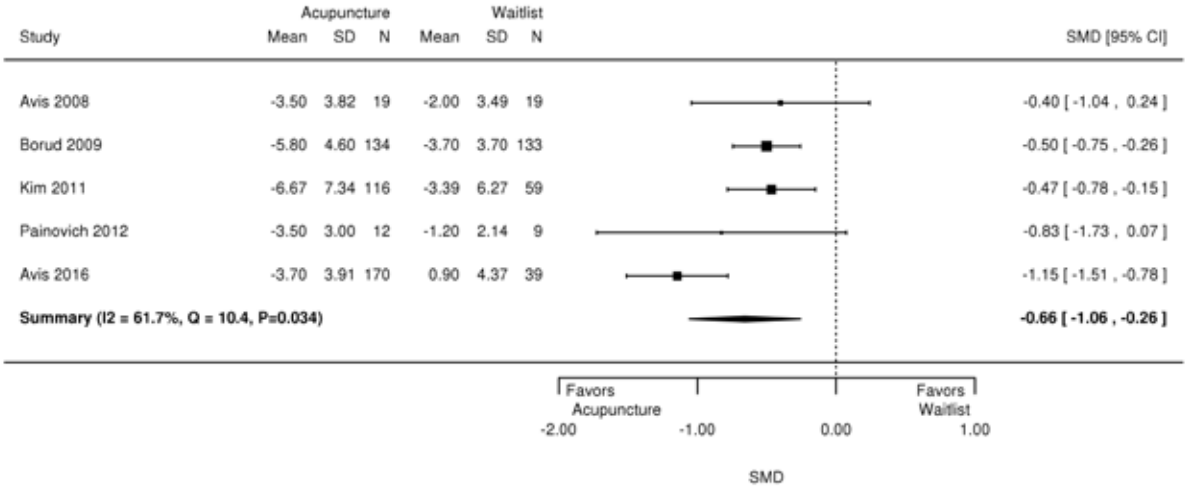
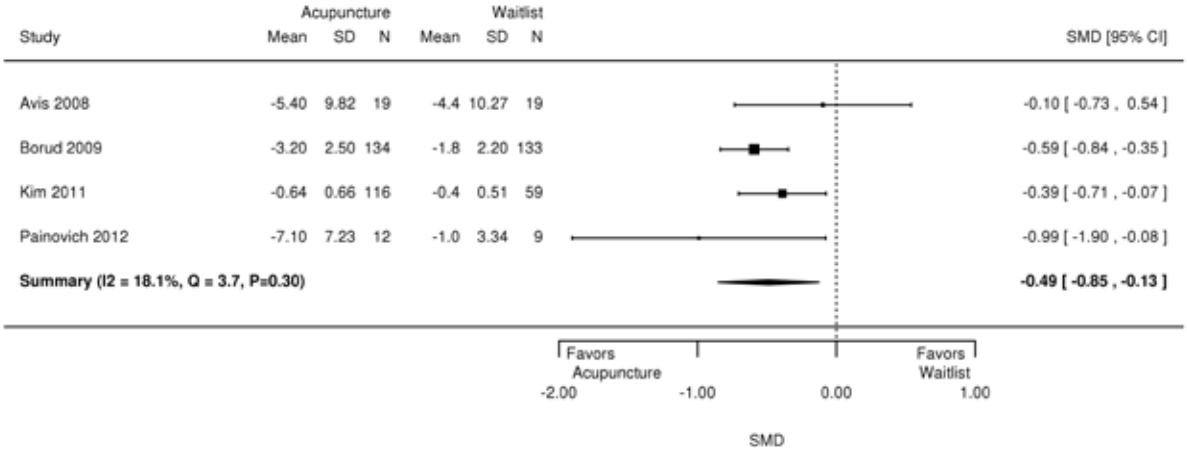


Figure 4. Forest Plot of Acupuncture versus No Acupuncture on Change in Hot Flash Severity at End of Treatment



Acupuncture versus sham acupuncture

Several different procedures designed to simulate the subjective experience of receiving an acupuncture treatment have been used as semi-inert comparisons to true acupuncture in the context of clinical trials. Dodin et al[40] conducted meta-analyses using data from 8 trials[68,69,72,74,76,77,80,83] that compared acupuncture to a sham control and reported daily VMS frequency or severity. These meta-analyses revealed no significant difference for daily VMS frequency (SMD -1.13, 95% CI -2.55 to 0.29; I^2 = 70%) but significantly improved severity of hot flashes (SMD -0.45, 95% CI -0.84 to -0.05, I^2 = 62%) compared with sham acupuncture. Subgroup analyses demonstrated that heterogeneity was partially explained by the trials involving women with breast cancer and trials with duration of treatment less than or greater than 12 weeks. Collectively, the 8 trials included in the meta-analyses had 37 low-risk, 3 high-risk, and 8 unclear risk of bias domains, as judged by the authors of the SR.

In the new RCT by Mao et al,[51] electroacupuncture (which involves administering low-dose electrical current through some of the acupuncture needles) reduced the hot flash composite score at 24 weeks by 7.4 units, compared with 5.9 with sham electroacupuncture, 5.2 with gabapentin, and 3.4 with the placebo medication ($p<0.001$). In the new RCT by Ee et al,[50] hot flash scores decreased in both groups by approximately 40% from baseline to end of treatment and were sustained for 6 months. There was no evidence of an advantage of electroacupuncture over sham acupuncture on quality of life. In the new RCT by Nedeljkovic et al,[53] acupuncture was more effective than sham acupuncture in reducing hot flash frequency and severity, and in improving menopause-related quality of life.

We updated the meta-analysis reported in the SR by adding findings from these 3 new RCTs.[50,51,53] Our updated meta-analysis generated an estimate of the SMD of -0.21 (95% CI -0.49 to 0.07, I^2 = 53.5%, 10 trials) for reduction in VMS frequency (Figure 5) and -0.35 (95% CI -0.70 to 0.01, I^2 = 66.5%, 8 trials) for reduction in VMS severity (Figure 6).

Figure 5. Forest Plot of Acupuncture versus Sham Acupuncture on Change in Hot Flash Frequency at End of Treatment

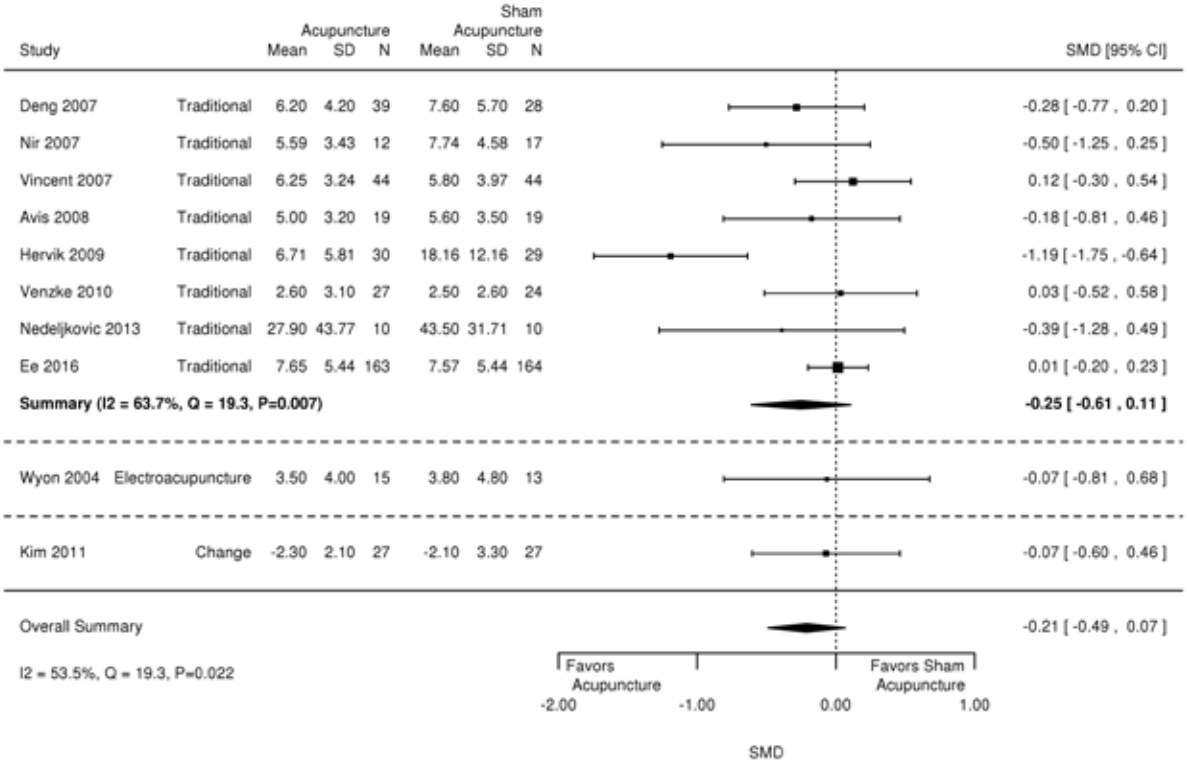
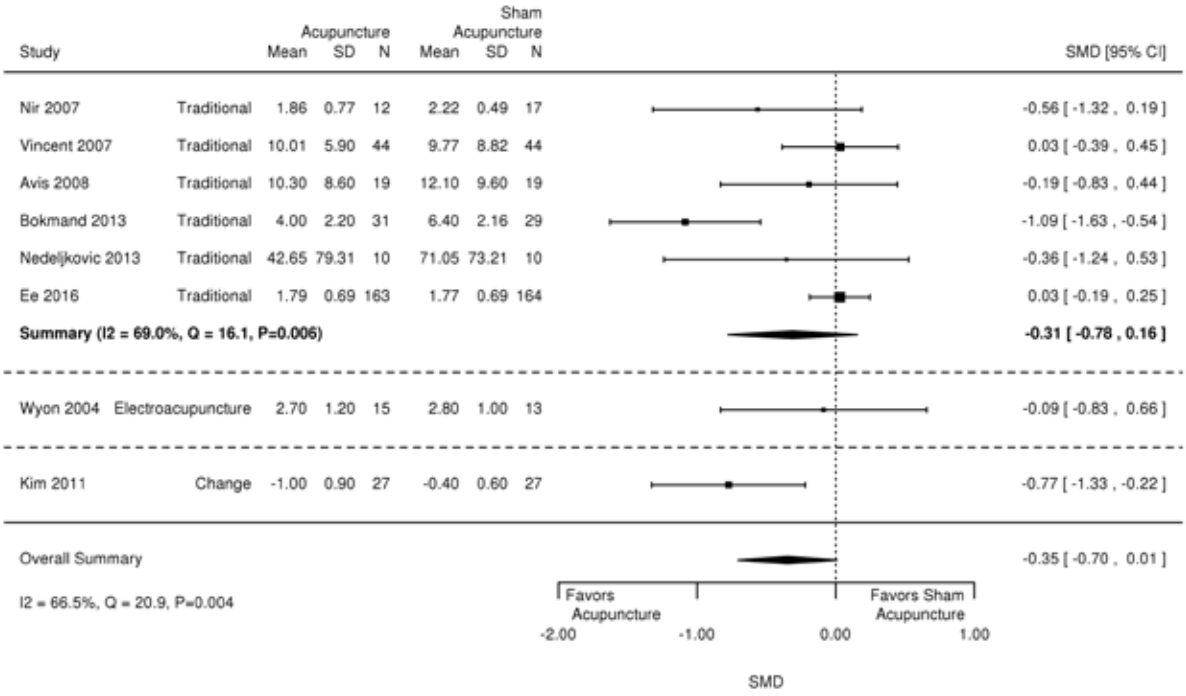
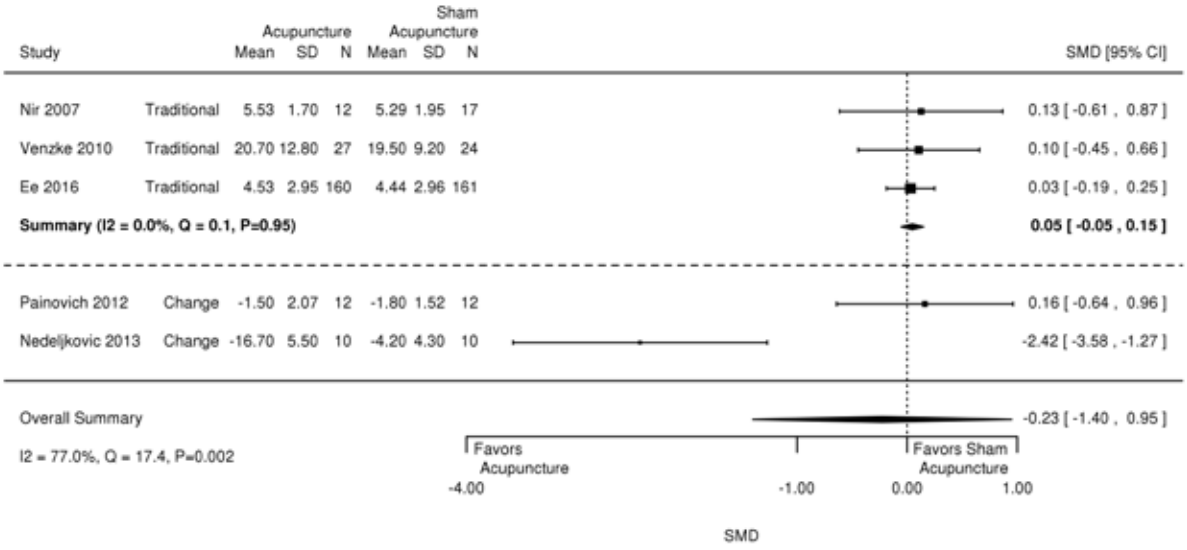


Figure 6. Forest Plot of Acupuncture versus Sham Acupuncture on Change in Hot Flash Severity at End of Treatment



In the 2 RCTs that assessed quality of life in the SR,[69,72] no significant difference in quality of life (as assessed by the Menopausal Specific Quality of Life questionnaire or the Greene Climacteric Scale) was found between acupuncture and sham acupuncture (SMD 0.11, 95% CI -0.33 to 0.55). Our updated meta-analysis that included results of changes in the Menopause Specific Quality of Life questionnaire in 2 of the new RCTs[50,53] generated an estimate of the SMD of -0.23 (95% CI -1.40 to 0.95, I^2 = 77.0%, 5 trials) for change in quality of life (Figure 7).

Figure 7. Forest Plot of Acupuncture versus Sham Acupuncture on Change in Quality of Life at End of Treatment



Acupuncture versus hormone therapy

We did not identify new RCTs that compared acupuncture with hormone therapy. The SR, however, included 3 RCTs that compared acupuncture with hormone therapy.[81-83] Among these, acupuncture was associated with significantly higher daily VMS frequency compared with hormone therapy (SMD 3.18, 95% CI 2.06 to 4.29, $I^2 = 0\%$, 3 trials). There was no statistically significant difference in VMS severity between acupuncture and hormone therapy (SMD 0.53, 95% CI -0.14 to 1.20, 2 trials). Two of these RCTs were judged to be of low risk for 4 of 6 domains, and one was judged to have 3 low-risk, 2 high-risk, and 1 unclear risk of bias domains. Quality of life as assessed by the Kupperman Index was significantly greater in the hormone therapy group than in the acupuncture group (MD 0.11, 95% CI 0.01 to 0.21, 1 trial) in the single study included in the review that reported this outcome.[82]

Acupuncture versus relaxation

We did not identify any new RCTs that compared acupuncture with relaxation. However, a single RCT[73] included in the SR compared electroacupuncture with relaxation, finding no significant between-group differences for either daily VMS frequency (MD -0.40, 95% CI -2.18 to 1.38) or severity (MD 0.20, 95% CI -0.85 to 1.25). This RCT is discussed in greater detail in the Meditation, Relaxation, and Mindfulness section of this report.

Summary of Findings for Acupuncture

We found evidence that a course of acupuncture improves VMS and quality of life compared with no acupuncture, but not more than sham acupuncture. This suggests that acupuncture may be useful as an adjunct therapy and that the observed clinical benefits associated with acupuncture in the context of clinical trials may be attributable in part or in whole to nonspecific effects. We also found evidence that acupuncture may improve sleep quality, with one of the

new RCTs demonstrating significant improvement in sleep compared with a waitlist control on 2 of 3 sleep outcome measures. There is inconclusive evidence of acupuncture's effectiveness in alleviating symptoms of depression or anxiety among women who experience menopause-related VMS. The only RCT included in our review that systematically assessed the safety of acupuncture relative to no acupuncture or medical management found that acupuncture was associated with significantly fewer adverse effects than gabapentin. To make these determinations, we drew from an existing, good-quality SR that included 15 pertinent trials and 4 new low risk of bias RCTs. Three of the 4 new RCTS found that reductions in Hot Flash Composite Score and hot flash frequency were maintained at 12, 16, and 24 weeks post-treatment, respectively.[51-53]

YOGA, TAI CHI, AND QIGONG

Key Points

- We identified one eligible good-quality SR, published in 2012, that included 5 RCTs comparing yoga with inactive (waitlist) and/or active (exercise, stretching) comparators, and 2 new RCTs published since 2012 that assessed the impact of yoga on hot flash frequency, severity, and quality of life. We did not identify any eligible SRs or new RCTs that evaluated the effectiveness of tai chi or qigong.

- We conducted a meta-analysis examining the effect of yoga compared with active and inactive comparators on hot flash severity change scores using 2 new RCTs and 2 RCTs included in the SR. Meta-analysis results suggest that yoga is significantly associated with a reduction in hot flash severity (SMD -0.36, 95% CI -0.65 to -0.07, $I^2 = 0.0\%$).

- One RCT examined the effect of yoga versus control on quality of life; results were inconclusive due to the small study sample, but suggest that yoga may improve some measures of quality of life. Adverse effects were reported as an outcome in only one RCT and were infrequent.

Systematic Review

Study characteristics

We identified one good-quality SR by Cramer et al[43] published in 2012 that included 5 RCTs[85-89] (582 total participants) evaluating the effectiveness of yoga for bothersome menopause symptoms. Yoga interventions were delivered in 34 to 40 sessions over 8 weeks to 4 months. Outcomes included overall menopause symptoms and quality of life;[85-87] VMS;[85,88,89] psychological symptoms;[85-89] and the adverse effects of yoga.[87] Only one study[88] examined the durability of treatment effects beyond end of treatment (at 12 weeks post intervention); this study only reported findings related to VMS using a daily diary (total score of hot flash frequency and severity) and not menopause symptoms or quality of life.

Authors of the SR[43] conducted a meta-analysis for outcomes reported by more than 2 RCTs. Meta-analyses evaluated the effect of yoga compared with an active comparator and yoga compared with no treatment on (1) total menopause symptoms, (2) psychological symptoms, and (3) VMS. Subgroup meta-analyses also considered the effect of yoga versus an attention control

and yoga versus no treatment separately on those 5 outcomes. One RCT was not included because it did not report sufficient data for quantitative analysis;[88] all RCTs were included in the qualitative analysis.

Characteristics of the SR and included RCTs are summarized in Table 5.

Table 5. Study Characteristics: Cramer et al 2012[43]

Characteristic	Value
Systematic review	
Number of included trials	5
Number of patients	582
Date of SR literature search	April 2012
Mean age range in years (median)	45.6 to 54.9 (49.0)
RCTs included in the SR	
Study years	2007-2011
Countries	
USA	2
India	2
Brazil	1
Population	
Perimenopausal or postmenopausal	4
Women who had completed treatment for breast cancer	1
Yoga interventions	
Yoga postures	5
Breathing	4
Meditation	5
Lifestyle lectures	2
Planned number of yoga treatments	34-40
Length of intervention period	8 weeks to 4 months
Training of yoga instructors	
Certified and experienced	2
Not reported	3
Comparisons	
Inactive—waitlist	4
Active—walking, stretching, lifestyle lectures	3
Inactive and active (3-arm trial)	2
Outcomes	
Total menopause symptoms	3
Vasomotor symptoms	3
Psychological symptoms	5
Sleep/insomnia	0
Adverse effects of yoga	1
Timing of last outcome assessment after randomization	
Short-term (taken closest to 12 weeks after randomization)	5
Long-term (taken closest to 12 months after randomization)	1

Risk of bias—Systematic review

In the SR,[43] risk of bias was assessed for each included RCT using the 12 criteria recommended by the Cochrane Back and Neck Review Group. RCTs that met 6 of the 12 criteria were rated as having low risk of bias. Of the 5 included RCTs, 2 were rated as low risk of bias[86,88] and 3 as high risk of bias.[85,87,89] Key issues included lack of blinding of participants or providers, inadequate intention-to-treat analysis, and inadequate disclosure of the full study protocol.

New Randomized Controlled Trials

Study characteristics

We identified 2 new RCTs published after 2012 that assessed the impact of yoga on VMS or quality of life among perimenopausal or postmenopausal women. Avis et al[54] conducted a 3-arm trial comparing the effects of 10 weeks of integral yoga versus health-and-wellness education and waitlist control on frequency and severity of hot flashes. Ngowsiri et al[55] conducted a trial comparing a 13-week Rusie Dutton course (which we determined to have similar components to yoga) with a waitlist control on menopause symptoms, including VMS. Outcomes examined among the 2 RCTs included VMS,[54,55] health-related quality of life,[54] psychological symptoms,[54,55] and sleep disturbance.[54,55] Yoga was compared with an attention health-and-wellness education control in one RCT[54] and with an inactive waitlist control in both RCTs. Sample sizes were 54[54] and 48.[55] All participants were perimenopausal or postmenopausal. One RCT did not report cause of menopause or breast cancer history.[55] Mean participant age ranged from 50.7 to 53.5 years. Yoga interventions included a combination of postures and breathing. One RCT included relaxation/meditation techniques in addition to yoga.[54] The duration of yoga treatment ranged from 10 weeks to 13 weeks; sessions were held weekly for 90 minutes in both trials and all participants were encouraged to practice at home.

Outcome measures

In the 2 new RCTs, VMS was assessed using the Daily Hot Flash Diary (DHFD) at 5 and 10 weeks post-randomization[54] and subscales from the MENQOL (Thai version) at 12 weeks post-randomization.[55] Quality of life was measured using the Global Quality of Life scale at 5 and 10 weeks post-randomization and the Hot Flash Related Daily Interference Scale at 5 and 10 weeks post-randomization.[54] Psychological symptoms were examined using the CESD-10 at 5 and 10 weeks post-randomization[54] (Avis 2014) and subscales from the MENQOL at 12 weeks post-randomization (Ngowsiri 2014).[55] Persistence of treatment effects beyond end of treatment was not examined in either study. Characteristics of the 2 new RCTs are summarized in Table 6.

Table 6. Study Characteristics of New RCTs

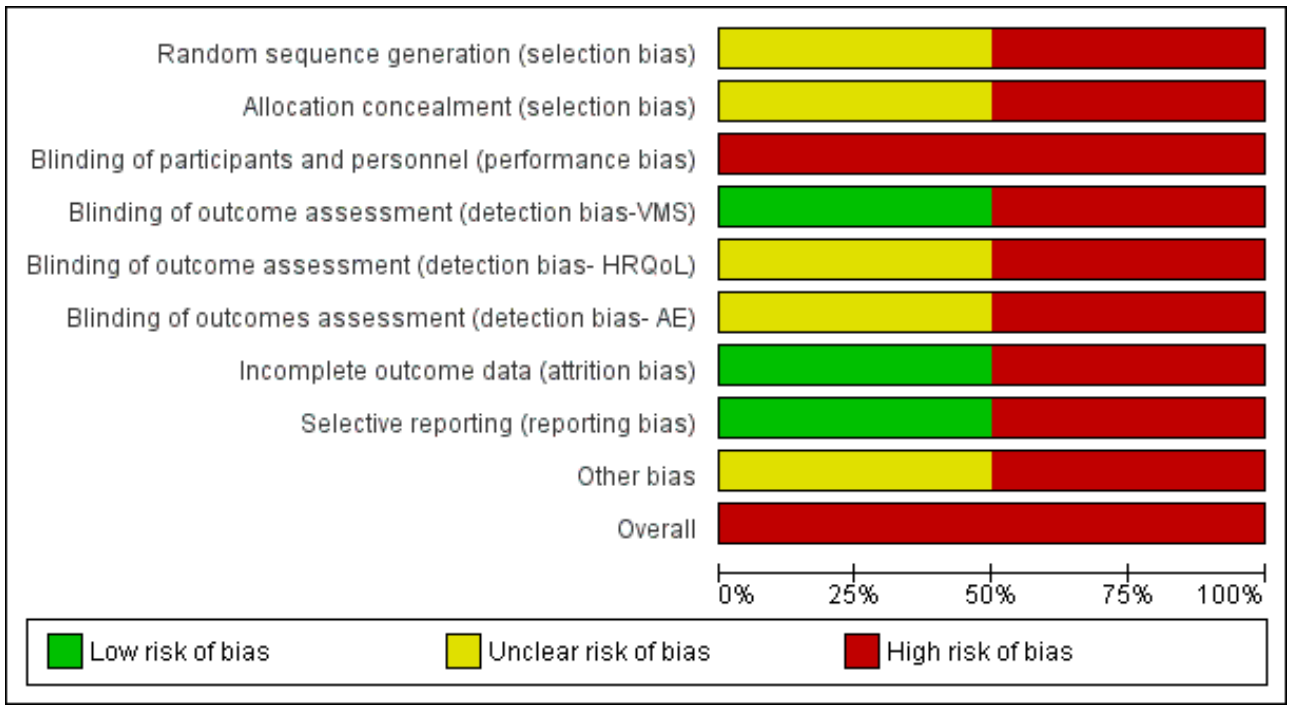
Study Country	**Population** # Women randomized Type of menopause # Hot flashes Mean age (range)	**Intervention** Category/type Session frequency/duration	**Comparator** Category/type Session frequency/duration	**General Outcomes** Instruments[a]
Avis 2014[54] USA	Yoga: 18 Health and wellness: 19 Waitlist: 17 Natural; surgical; and women treated for breast cancer (not currently receiving treatment) ≥4 hot flashes per day on average for ≥4 weeks and ≥2 months since last menses Yoga: 53.5 (0.7) Health and wellness: 52.8 (0.7) Waitlist: 53.5(0.7)	Yoga: integral yoga, adapted from Satchidananda, with breathing, relaxation, and postures 1 group session per week for 10 weeks; 2-3 home practice per week for 15 minutes 90 minutes duration	Attention: health and wellness education 10 weekly educational classes; materials to read at home (to match time spent in home practice among individuals in intervention group) 90 minutes duration Inactive comparator: Waitlist; no initiation of new therapies for hot flashes or participation in classes for 10 weeks	· Frequency/severity of hot flashes (DDHF) · Sleep quality (Global Sleep Quality index) · Anxiety (HSCL) · Depressive symptoms (CESD) · Global quality of life (SF-36)
Ngowsiri 2014[55] Thailand	Rusie Dutton: 24 Waitlist: 26 Type of menopause NR Mild to moderate menopause symptoms per Self-Assessment for Menopause Symptoms (developed from MRS) Yoga: 52.9 (4.3) Waitlist: 50.7 (3.6)	Yoga: Rusie Dutton 1 group class per week for 13 weeks; 2 home practice per week 90 minutes duration	Inactive comparator: Waitlist; offered class 13 weeks after post-test assessment was completed	Menopause specific quality of life subscales (Thai version of MENQOL subscales: vasomotor, psychological)

[a] Menopause instruments and scales are described in Appendix D.

Risk of bias—New RCTs

Both new RCTs were judged to be high risk of bias. Of primary concern was the use of a weak control (waitlist). Also, the study methods were not clearly described, including the method for random sequence generation, allocation concealment, and participant, provider, and assessor blinding (Figure 8).

Figure 8. Risk of Bias Ratings for New Yoga RCTs



Synthesis of Findings: Primary Outcomes

Vasomotor symptoms

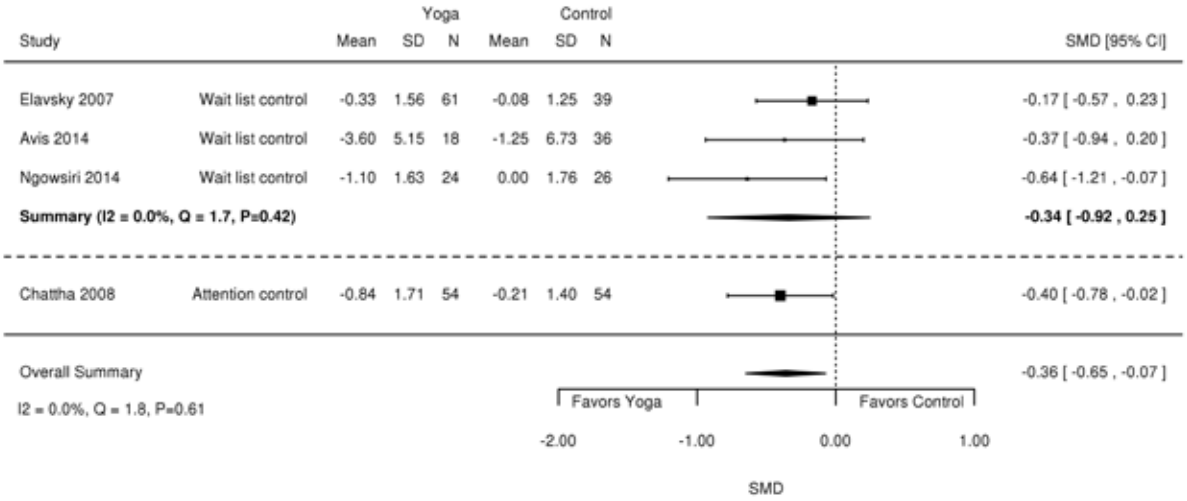
Yoga versus inactive control

In the SR by Cramer et al, 2 studies examined the effect of yoga versus a waitlist control on VMS.[85,88] The study by Carson et al was not included in the quantitative analysis because of incomplete reporting; therefore the SR did not conduct a meta-analysis that examined the effect of yoga compared with a waitlist control only on VMS. The Carson et al study was the only trial in the SR that examined the durability of treatment effect comparing yoga versus a waitlist control on VMS[88]; qualitative results suggest that the positive effects of yoga on VMS persisted at 12 weeks after the end of treatment (20 weeks post-randomization).

Our updated meta-analysis included data from the SR and both new RCTs. We summarized the findings of 3 RCTs that examined the effect of yoga compared with a waitlist control on changes in hot flash severity.[54,55,85] These trials included the 3-armed trial[85] identified in the SR, a newly identified 3-armed RCT that also included inactive and attention control arms,[54] and another newly identified RCT that compared yoga with a waitlist control.[55] We pooled the 3-arm trial samples and, unlike the SR, analyzed them as part of the inactive control comparator subgroup

analysis. In our meta-analysis, we found no effect of yoga (3 trials; SMD -0.34, 95% CI -0.92 to 0.25, $I^2 = 0.0\%$) (Figure 9).

Figure 9. Forest Plot of Yoga versus Control on Change in Hot Flash Severity at End of Treatment



Yoga versus attention control

Authors of the SR conducted a meta-analysis of 2 RCTs; one compared yoga with an attention control[89] and the other, a 3-armed trial, included an attention and inactive control.[85] Only data from the attention control arm was used; the analysis examined yoga versus an attention control and found that, compared with exercise, stretching, and lifestyle discussions, yoga was not associated with a reduction in VMS (2 trials, SMD -0.13, 95% CI -0.58 to 0.33, $I^2 = 67\%$).[85,89]

For one of the new RCTs[54] – a 3-arm trial (yoga vs waitlist control vs health-and-wellness education attention control) – the data from the 2 control arms were pooled and analyzed as part of the inactive comparator subgroup analysis. In this trial,[89] yoga was more effective than the waitlist and attention control arms combined (SMD -0.40, 95% CI -0.78 to -0.02) (Figure 9).

Overall effect of yoga versus all controls

Of the 2 new RCTs, one found significant differences in vasomotor pre/post scores on the MENQOL Thai version subscale ($p = 0.04$).[55] The other RCT did not find any statistically significant effects of yoga on VMS (Table 6).[54] We updated the meta-analysis in the SR by adding data reported in the 2 new RCTs[54,55] to assess the effect of yoga on changes in hot flash severity. We included inactive (waitlist)[54,55,85] and attention (health-and-wellness education and exercise) controls;[54,89] attention and inactive controls were pooled. We found that yoga was associated with an overall statistically significant decrease in hot flash severity (4 trials, SMD -0.36, 95% CI -0.65 to -0.07, $I^2 = 0.0\%$) (Figure 9).

In the SR,[43] authors conducted a meta-analysis of 2 RCTs[85,89] and found that yoga compared with exercise[89] and waitlist control[85] was not associated with short-term benefits for VMS (2

trials, SMD -0.04, 95% CI -0.68 to 0.60, I^2 = 81%). The effects were inconsistent across the 2 RCTs: one showed a small benefit and the other showed a small harm, but in both, the effect size was not statistically significant. However, in one RCT that included a longer-term follow-up (20 weeks) for yoga compared with no treatment, significant effects were found for VMS.[88]

Health-related quality of life

Of the 5 RCTs identified by the SR, only one RCT used a scale (the Greene Climacteric Scale) that by our judgment assessed the concept of health-related quality of life,[85] which was conceptualized as total menopause symptoms in the SR. The other 2 RCTs that assessed total menopause symptoms used the Kupperman Index[87] and the Menopause Rating Scale.[86] In an RCT comparing yoga with a walking attention control and a waitlist control,[85] yoga was not found to have an impact on quality of life. Of the 2 new RCTs, one did not assess quality of life[55] and the other used the SF-36 and found no effect of yoga on quality of life.[54]

Summary of Findings for Yoga

We identified 1 good-quality SR and 2 new RCTs that examined the effect of yoga on bothersome menopause symptoms and quality of life. In general, RCTs demonstrated inconsistent effects of yoga for reducing VMS and increasing quality of life. Meta-analysis results from the SR suggest that short-term effects of yoga are not significantly associated with changes in VMS. However, results from our meta-analysis suggest that yoga is associated with a decrease in hot flash severity. Our results likely differ from those in the SR because of the addition of the 2 new RCTs and because we used a treatment effect that adjusted for baseline differences in symptom severity.[85] Yoga demonstrated no impact on quality of life.

STRUCTURED EXERCISE

Key Points

- We identified 2 SRs (1 good quality and 1 fair quality) that met our eligibility criteria. One SR focused exclusively on exercise for the frequency and severity of hot flashes among perimenopausal and postmenopausal women, and the other SR included evaluations of exercise and mindfulness and relaxation. All of the studies in the good-quality SR were included in the fair-quality SR.

- In the good-quality SR, no statistical evidence was found to support exercise versus control (n = 4 studies) or exercise versus yoga (n = 2 studies). One RCT (n = 14) included in this SR found hormone therapy to be statistically superior to exercise.

- We identified 2 new RCTs that examined exercise for VMS frequency and severity of hot flashes among perimenopausal or postmenopausal women. One of these included only women after breast cancer treatment. Both RCTs were judged to be low risk of bias.

- Our new meta-analysis compared exercise as an intervention for hot flash frequency and severity. We found that exercise did not produce a significant effect on either hot flash frequency (SMD -0.08, 95% CI -0.33 to 0.16, I^2 = 0.0%, 4 trials) or severity (SMD -0.06, 95% CI -0.21 to 0.10, I^2 = 0.0%, 5 trials).

- One new RCT reported on the outcome of menopause-specific quality of life and found significantly lower sleep problem scores at the 6-month follow-up. The other new RCT measured overall health-related quality of life and reported moderate benefit on short-term (effect size 0.46) and long-term (effect size 0.41) physical function outcomes. A companion study reported that exercise significantly improved sleep quality and decreased hot flashes during sleep.

- One RCT reported on adverse events, finding no serious adverse events and no differential incident adverse events between the exercise and control groups.

Systematic Reviews

Study characteristics

We identified 2 eligible SRs evaluating the effectiveness of structured exercise for VMS and quality of life.[44,46] The good-quality SR by Daley et al[44] published in 2014 included 5 RCTs evaluating the effectiveness of structured exercise for menopausal hot flashes.[85,90-93] The fair-quality SR by Woods et al[46] published in 2014 included 4 RCTs that were not included in the SR by Daley et al. Of those, 3 were excluded by Daley et al because women were not symptomatic at baseline, and one was excluded because participants were taking hormone therapy at baseline. Two RCTs[85,92] were included in both SRs.

For the purposes of this report, we focus only on the good-quality Cochrane Collaboration SR by Daley et al[44] because it included more studies, employed more inclusive eligibility criteria, contained a meta-analysis of studies, and was of higher quality. In this SR, sample sizes of the 5 included RCTs ranged from 37 to 355. The population was middle-aged women with VMS. Most studies required women to be sedentary or at low activity levels at baseline. Hormone therapy was restricted in most studies (n = 3) to the previous 2 months to 6 months. The comparator in most studies was an inactive control (n = 4) or yoga (n = 2).

A total of 762 women were enrolled across the 5 RCTs with an age range of 40 to 63 years. Three of the 5 RCTs were 3-arm trials resulting in a total of 8 comparators. Comparisons in most studies were made between inactive controls (n = 4) or yoga (n = 2). The outcome was VMS frequency and severity for all 5 RCTs with only one study reporting adverse effects. The majority of the RCTs included symptomatic perimenopausal or postmenopausal women (n = 4), with one RCT including women with surgically removed ovaries. In the majority of studies (n = 4) exercise was conducted with supervision or in a group session. The number of planned sessions and length of the intervention period varied greatly from 12 to 96 weeks and 12 to 36 weeks, respectively. Only 2 RCTs used a certified or trained exercise instructor during the exercise sessions. The majority of studies were short-term (≤12 weeks).

Authors of the SR[44] conducted a meta-analysis using random-effects models when 2 or more RCTs reported the same outcomes. The primary outcome was frequency and severity of VMS and the secondary outcome was adverse effects of the intervention. For the primary outcome, summary estimates were generated for exercise compared with no treatment or control (n = 3 studies)[85,90,91] and exercise compared with yoga (n = 2 studies).[85,91] Qualitative analysis was conducted for one RCT[93] that compared structured exercise with hormone therapy and for one RCT in which the authors determined that insufficient data were provided to be included in the

meta-analysis. For the secondary outcome, qualitative analyses were provided for one RCT that included data on adverse events of the intervention.

Characteristics of the SR and included RCTs are summarized in Table 7.

Table 7. Study Characteristics: Daley et al 2014[44]

Characteristic	Value
Systematic review	
Number of included trials	5
Number of patients	762
Date of SR literature search	March 2014
Mean age range in years (median)	40-63
RCTs included in the SR	
Study years	2002-2011
Countries Iran Finland Sweden USA	 1 1 1 2
Population Perimenopausal or postmenopausal Women who had completed treatment for breast cancer Women who had ovaries surgically removed	 4 0 1
Exercise interventions Exercise unsupervised Exercise with supervision Exercise in a group	 1 2 2
Planned number of exercise treatments	12 to 96
Length of intervention period in weeks	12 to 36
Training of exercise instructors No instructor or supervision Certified and trained Not reported	 2 2 1
Comparisons Inactive control Active (yoga or hormone therapy)	 5 3
Outcomes Hot flash frequency and severity Adverse effects	 5 1
Timing of last outcome assessment after randomization 3 months 6 to 9 months	 3 2 (6 and 9 months)

Risk of bias—Systematic review

In the SR,[44] risk of bias was assessed for each included RCT using the Cochrane Risk of Bias Tool. The risk of bias was judged to be high among 4 RCTs[85,90,92,93] and low in one.[91] Key issues included unclear specification of sequence generation (n = 2 studies), unclear outcome

assessment blinding (n = 3 studies), high attrition rate (n = 1 study), and incomplete data reporting (n = 2 studies).

New Randomized Controlled Trials

Study characteristics

We identified 2 relevant RCTs that were not included in the SR by Daley et al.[62,63] Both compared exercise with an inactive control. One RCT[62] was published in 2015 after the SR, and the other,[63] published in 2012, was not included in the SR because it evaluated only women who had primary breast cancer. The new RCT by Daley et al[62] was a 3-arm trial randomizing participants to (1) exercise and DVD support, (2) exercise and community-based social support, or (3) inactive control. The outcomes were measured at both 6 months (immediately following completion of the intervention) as well as at 12 months' post-randomization. The 2 exercise intervention groups consisted of face-to-face consultations with a physical activity facilitator, with one of the groups additionally receiving a DVD of menopause education and written materials to encourage regular exercise.

The RCT by Duijts et al[63] was a 4-arm trial randomizing participants to (1) exercise alone, (2) exercise plus cognitive behavioral therapy, (3) cognitive behavioral therapy alone, or (4) a waitlist control group. The exercise program was a 12-week, individually tailored, home-based, self-directed exercise program of 2.5 to 3 hours per week. A primary outcome for both of these studies was hot flash frequency and severity. Outcomes were measured at both 12 weeks (immediately following completion of the intervention) as well as at 6 months post-randomization.

Among both RCTs, there were 576 participants; all were perimenopausal or postmenopausal. One RCT[63] included only women with primary breast cancer not currently on cancer treatment but experiencing VMS. The other[62] included women with VMS who were not receiving hormone therapy in the past 3 months. Mean participant age ranged from 47.7 to 57.7 years. Exercise interventions were group sessions[63] and individually tailored programs,[62] with both studies focusing on aerobic exercise. The intervention in one RCT[63] included one treatment arm with cognitive behavioral therapy and a treatment arm with combined cognitive behavioral therapy and exercises; for the purposes of this review, it was excluded from the meta-analysis. The exercise intervention in the other RCT[62] included supplemental DVD education in one arm and social support in the other arm. We combined these 2 arms and compared them against the control group in our updated meta-analysis because we were primarily concerned with the efficacy of exercise. From both RCTs, the duration of exercise interventions ranged from 12 weeks to 24 weeks; sessions were held weekly for 40 to 90 minutes, and all participants were encouraged to exercise at home on their own.

Outcome measures

In both new RCTs, VMS was assessed using the Hot Flush Rating Scale. Outcomes were assessed at 6 and 12 months post-randomization[62] and at 12 weeks and 6 months follow-up.[63] One RCT included generic health-related quality of life measured with the SF-36 Short Health Assessment Form as a secondary outcome[63] and the other RCT included menopause-specific

quality of life using subscales from the Women's Health Questionnaire as a secondary outcome.[62]

Characteristics of the 2 new RCTs are summarized in Table 8.

Table 8. Study Characteristics of New RCTs

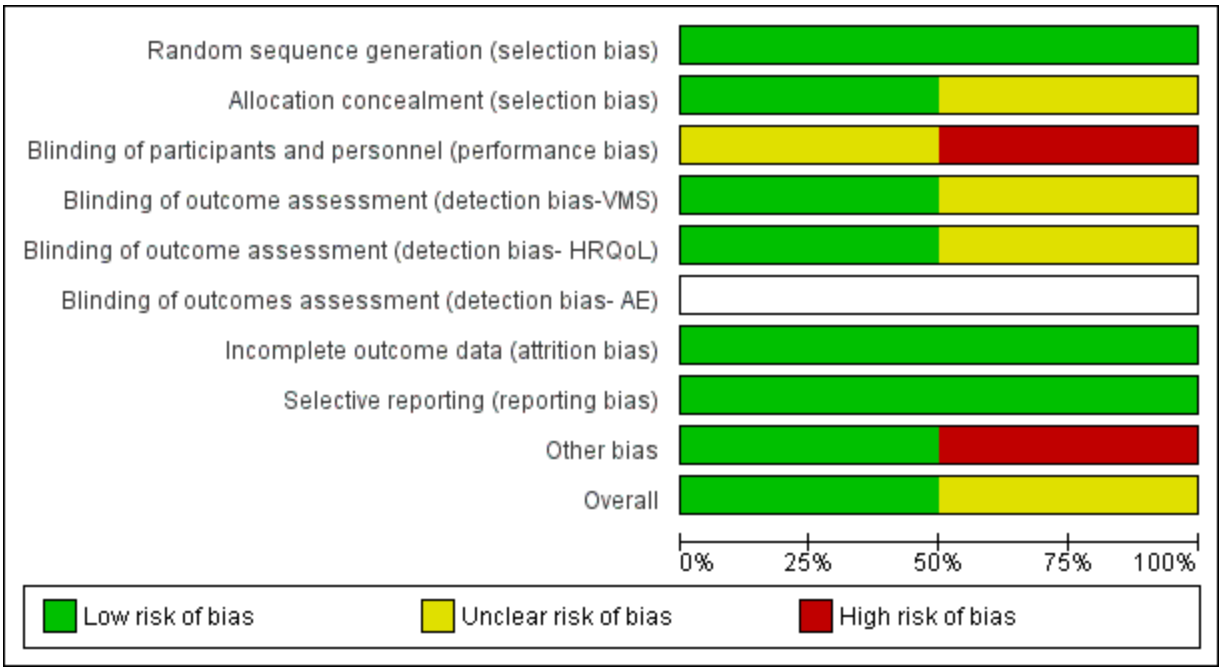
Study Country	**Population** # Women randomized Type of menopause # Hot flashes Mean age (range)	**Intervention** Category/type Session frequency/duration	**Comparator** Category/type Session frequency/duration	**General Outcomes** Instruments[a]
Duijts 2012[63] Netherlands	Exercise + CBT: 106 Exercise: 104 Waitlist: 103 Primary breast cancer and not currently on treatment 2 or 3 VMS symptoms "sometimes" or 1 "often" in previous 2 weeks Exercise + CBT: 49.0 Exercise: 47.7 Control: 47.8	Cognitive behavioral therapy; 6 weekly group sessions of 90 minutes, including relaxation exercises Physical exercise: 12 weeks, individually tailored, home-based, self-directed program of 2.5 to 3 hours per week	Waitlist control with no instructions	· Hot flashes and night sweats measured using Hot Flush Rating Scale · Generic quality of life measured using SF-36 Short Form Health Survey
Daley 2015[62] Thailand	Exercise + DVD: 87 Exercise + social support: 87 Control: 87 Perimenopausal and postmenopausal women experiencing ≥ 5 hot flashes/night sweats per day and have not taken hormone therapy in previous 3 months Exercise + DVD: 52.3 (2.4) Exercise + social support: 52.7 (2.5) Control: 52.0 (2.4)	Both intervention groups: 2 face-to-face consultations with a physical activity facilitator to support engagement in regular exercise Exercise + DVD group received a menopause-specific information DVD and written materials to encourage regular exercise Exercise + support group offered opportunity to attend exercise social support groups in their communities	Inactive comparator (control) offered the opportunity to have an exercise consultation and were given a pedometer at the end of their involvement in the study	· Hot flashes and night sweat frequency measured using Hot Flush Rating Scale · Menopause-specific quality of life subscales from the Women's Health Questionnaire

[a] Menopause instruments and scales are described in Appendix D.

Risk of bias—New RCTs

Both new RCTs were judged to be low risk of bias. Of primary concern in one RCT was the use of a weak control and poor compliance with the intervention.[62] Both studies had high or unclear risk of bias for items related to blinding of the intervener or the participant. However, studies using exercise as an intervention are commonly unable to implement blinding and therefore comments and ratings should be interpreted in light of these constraints (Figure 10).

Figure 10. Risk of Bias Ratings for New Structured Exercise RCTs[a]



[a] A white bar indicates that the outcome was not reported.

Synthesis of Findings: Primary Outcomes

Vasomotor symptoms

Structured exercise versus inactive control

In the SR, 4 RCTs were evaluated that compared structured exercise with an inactive control.[85,90,91,93] Three of the these were included in a meta-analysis; one[93] was excluded due to inadequate reporting of the control group. The random-effects meta-analysis of the other 3 RCTs demonstrated no significant effect of exercise on the outcome of frequency or severity of VMS when compared with no treatment or control. The RCT excluded in this meta-analysis[93] included 37 women and reported a significant decrease ($p<0.05$) in hot flash scores in the exercise plus soy milk group compared with control.

Our updated meta-analysis provided separate pooled estimates for hot flash frequency and severity. We summarized 3 RCTs[85,90,91] from the SR examining a change in hot flash frequency plus one new RCT[62] assessing a follow-up for hot flash frequency. Effects were fairly consistent and studies were homogeneous; however, we found no statistical evidence that exercise

decreased hot flash frequency when compared with no treatment or control (SMD -0.08, 95% CI -0.33 to 0.16, I^2 = 0.0%, 4 trials) (Figure 11). Only one study, the new RCT by Daley et al,[62] measured hot flash frequency at 12 months, 6 months following end of treatment, finding no significant difference between the intervention arm and control arm. For hot flash severity, we summarized 3 RCTs[62,63,85] on follow-up in hot flash severity and 2 RCTs[90,91] on a change in hot flash severity. Again, effects were fairly consistent and studies were homogeneous; however, we found no statistical evidence that exercise decreased hot flash severity when compared to no treatment or control (SMD -0.07, 95% CI -0.26 to 0.13, I^2 = 1.8%, 3 trials) (Figure 12). Both Duijts et al and Daley et al followed participants after the end of treatment, showing no significant durability of effect for exercise as an intervention for hot flash severity.

Figure 11. Forest Plot of Exercise versus Control on Change in Hot Flash Frequency at End of Treatment.

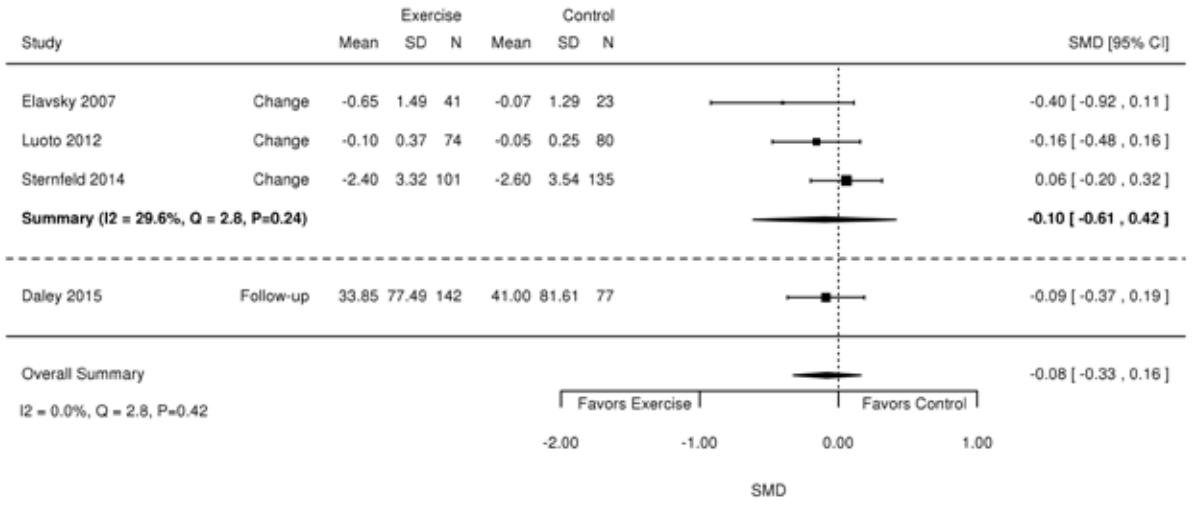
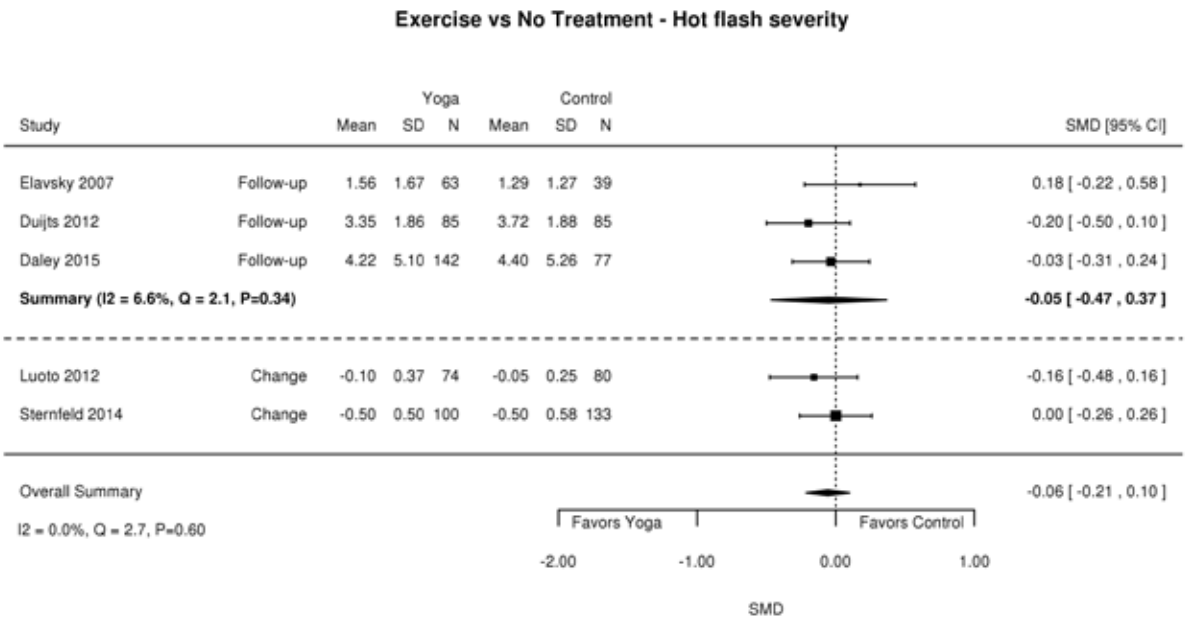


Figure 12. Forest Plot of Exercise versus Control on Change in Hot Flash Severity



Structured exercise versus yoga

Authors of the SR[44] identified 2 RCTs that compared structured exercise with yoga.[85,91] In a random-effects meta-analysis no evidence was found that structured exercise was significantly better than yoga for reducing a combined frequency or severity score for hot flashes (SMD 0.03, 95% CI -0.45 to 0.38, I^2 = 61%, 2 trials).

Structured exercise versus hormone therapy

Only one RCT was identified in the SR evaluating structured exercise compared with hormone therapy.[92] This RCT reported a larger reduction in the frequency of hot flashes per 24 hours in the hormone treatment group than in the exercise group (mean difference 5.80, 95% CI 3.17 to 8.43) among 14 participants at a 12-week follow-up.

Sensitivity analyses

Authors of the SR[44] found that exclusion of the single low risk of bias RCT did not affect the overall findings for any outcome.

Summary of Findings for Exercise

Overall, there were 7 RCTs evaluating exercise as an intervention among perimenopausal or postmenopausal women experiencing hot flashes. Five of these were identified in one good-quality SR with a meta-analysis examining the effect of exercise on the frequency and severity of hot flashes among perimenopausal or postmenopausal women.[44] We also identified 2 new RCTs and conducted a new meta-analysis separating the outcomes of hot flash frequency and severity.[62,63] The meta-analysis conducted in the SR along with our meta-analysis indicated no

significant effect of exercise for decreasing hot flash frequency or severity. Statistically significant findings were shown in 2 new RCTs for the secondary outcomes of generic health-related quality of life[63] and menopause-specific quality of life.[62] A companion follow-up paper indicated that exercise significantly improved sleep quality and decreased hot flashes during sleep. Only one RCT reported adverse events, finding no serious or differential adverse events between exercise and control groups.[91] The sample sizes of included studies in the SR were small, many of the primary studies included in the SR were judged to be of low quality, and few studies addressed long-term outcomes. The new RCTs were of greater sample size and judged by our group to be low risk of bias[63] and unclear risk of bias.[62] We conclude that the there is no evidence that structured exercise is beneficial over an inactive control for the outcomes of VMS frequency or severity, and limited evidence suggests that exercise is not beneficial over an active comparator or hormone therapy.

MEDITATION, MINDFULNESS, HYPNOSIS, AND RELAXATION

Key Points

- We identified 5 SRs (2 good quality, 3 fair quality) that met eligibility criteria. Four SRs examined relaxation and mindfulness. Three of these reported on trials of multiple types of interventions. One examined hypnosis exclusively.

- No significant difference in reduction of hot flash frequency was found between any types of relaxation or mindfulness interventions compared with control groups for the primary outcomes of VMS frequency or quality of life. One small RCT found that severity of hot flashes decreased more in a modified applied relaxation training group compared with traditional applied relaxation. One trial of mindfulness-based stress reduction compared with a waitlist control found significant improvements in anxiety and sleep quality. However, the quality of evidence overall was poor.

- One fair-quality SR included 3 RCTs examining the impacts of hypnosis on breast cancer-related symptoms, including hot flashes, and found a significant decrease in hot flashes compared with no treatment (effect size ranged from 0.479 to 1.25), but no difference with hypnosis when compared to an active control (gabapentin).

- We identified 6 new RCTs that were not included in existing SRs. Three RCTs examined paced respiration and 3 examined applied relaxation. Overall, the quality of these new studies was mixed.

- We conducted 2 new meta-analyses examining 4 RCTs comparing paced respiration with a control group. We found that paced respiration is not associated with a statistically significant decrease in hot flash frequency (SMD 0.04, 95% CI -0.73 to 0.82, I^2 = 56.6%, 3 trials) or severity (SMD 0.06, 95% CI -0.69 to 0.80; I^2 = 65.1% 3 trials).

Systematic Reviews

Study characteristics

We identified 5 eligible SRs evaluating the effectiveness of meditation, mindfulness, or relaxation interventions for bothersome menopause symptoms.[45-49] Of these, one good-quality SR focused solely on relaxation,[45] and one fair-quality SR focused solely on hypnosis,[49] while the other 3 evaluated multiple interventions.[46-48] Of the latter SRs, one was good quality[48] and 2 were fair quality.[46,47]

For the purposes of this report, we focus on the 2014 Cochrane Collaboration SR by Saensak et al[45] addressing relaxation, and on the 2015 SR by Cramer et al,[49] which presented qualitative findings only of studies that examined hypnosis. We include findings from the latter primarily because it was the only SR available on this intervention type. The other fair-quality SRs evaluated additional studies that were not included in our 2 prioritized SRs, but the individual studies were generally of poor quality. One unique RCT evaluating relaxation[94] was included in the good-quality SR[48] of multiple interventions and is noted in our synthesis. Only one RCT evaluating mindfulness-based stress reduction[95] was identified and it was described in 2 of the fair-quality SRs.

The good-quality SR from 2014[45] by Saensak et al included 4 RCTs[96-99] comparing relaxation-based interventions and either nontreatment or other nonhormonal treatments. The authors focused on identifying relaxation techniques that were based on physiological principles of either somatic or cognitive relaxation, or both. These included more commonly known types of relaxation techniques such as paced respiration and muscle relaxation, though the authors also looked for a wide variety of relaxation approaches. Six databases and the grey literature were searched through February 19, 2014. The 4 RCTs included 281 menopausal women ranging in age from 30 to 77 years. Women were included if their last menstrual period was at least 6 months ago,[98] and one RCT was limited women to those whose last menstrual period was at least 12 months ago.[97] Menopausal status was confirmed with labs including follicle-stimulating hormone and estradiol levels in 2 RCTs.[98,99] Two RCTs targeted women with primary breast cancer who were bothered by hot flashes, one of which specified that women taking aromatase inhibitors or hormone therapies other than tamoxifen were included.[96,98] The reason for menopause onset (spontaneous versus surgical) was not specified in one study.[99]

VMS status in these RCTs was described as follows: one trial specified hot flashes that were troublesome,[96] one trial required at least 5 hot flashes per day,[97] one trial required at least 2 hot flashes per 24 hour period,[98] and one trial required hot flashes severe enough to request treatment.[99] Outcomes studied included VMS frequency per 24 hours (n = 3),[97-99] Kupperman Index (n = 2),[96,98] and mood (n = 1)[96]; no studies reported on quality of life or adverse effects. Two RCTs included in the SR reported duration of treatment sessions and number of sessions planned.[98,99] In both of those, the applied relaxation arms received 12 weekly training sessions of 60 minutes per session. Duration and frequency of relaxation treatments in the other 2 RCTs were not reported. Only one of the 4 trials included in the SR assessed the durability of treatment effects, reporting outcomes 12 weeks after completion of the intervention.[98] Final follow-up visit data from this trial and that from a second trial of relaxation[99] were pooled despite differences in time since baseline.

Characteristics of this SR and included RCTs are summarized in Table 9.

Table 9. Study Characteristics: Saensak et al 2014[45]

Characteristic	Value
Systematic review	
Number of included trials	4
Number of patients	281
Date of SR literature search	February 2014
Age range in years	30 to 77
RCTs included in the SR	
Study years	1992-2008
Country, number of studies Sweden United Kingdom United States	 2 1 1
Population Excluded women who had stopped hormone therapy within 6 months Surgical menopause allowed Breast cancer patients	 2 3 2
Interventions Paced respiration Applied relaxation	 1 3
Planned number of sessions 12 sessions of 60 minutes Not reported	 2 2
Duration of intervention 4 weeks 12 weeks	 1 3
Training of instructors Not reported	 4[a]
Comparisons Control Electro acupuncture Superficial needle insertion/electroacupuncture/oral estrogen Alpha-wave biofeedback	 1 1 1 1
Outcomes VMS frequency per 24 hours Quality of life Menopause-related quality of life Mood Adverse effects	 3 0 2 1 0
Timing of last outcome[b] 4 weeks 6 months	 2 1

[a] No restriction was placed on training type of person delivering intervention.
[b] Timing of outcomes for one RCT was not reported.

The fair-quality SR from 2015[49] included 13 RCTs examining hypnosis among women with breast cancer, women who were survivors of breast cancer, women undergoing diagnostic breast

biopsy, and postmenopausal women without a history of breast cancer. Five databases were searched from inception through February 25, 2014, as well as a review of bibliographies. Of the 13 RCTs, 3 were relevant to our review and included a total of 274 women.[100-102] One RCT compared hypnosis with attention control among 187 postmenopausal women.[100] One compared hypnosis with no treatment among breast cancer survivors (n = 60)[101]; and one compared hypnosis with moderate-dose gabapentin in breast cancer survivors and women at increased risk of developing breast cancer (n = 27).[102] Outcomes measured by these RCTs included VMS frequency (n = 2),[100,102] health-related quality of life (n = 3),[100-102] mood (n = 1),[101] and adverse effects (n = 2).[100,102] Timing of follow-up assessments during these RCTs was not clearly stated.

Characteristics of this SR and included RCTs are summarized in Table 10.

Table 10. Study Characteristics: Cramer et al 2015[49]

Characteristic	Value
Systematic review	
Number of included trials	3
Number of patients	274
Date of SR literature search	February 2014
Age range in years	54.5 to 58.2[a]
RCTs included in the SR	
Study years	2008-2013
Country, number of studies United States	 3
Population Breast cancer survivors Postmenopausal women without a breast cancer history	 2 1
Interventions Hypnosis	 3
Planned number of sessions 5 weekly sessions 3 weekly sessions	 2 1
Duration of intervention 3 to 5 weeks	 3
Training of instructors Not reported	 3[b]
Comparisons No treatment Attention control Gabapentin	 1 1 1
Outcomes VMS frequency Health-related quality of life Mood Adverse effects	 2 3 1 2
Timing of last outcome Not reported	 3

[a] One study did not report mean age of participants.
[b] No restriction was placed on training type of person delivering intervention.

Saensak at al[45] pooled the data from 2 RCTs for a meta-analysis comparing the effectiveness of relaxation versus acupuncture.[98,99] Too few studies were identified for additional sensitivity analysis. Cramer et al[49] presented only qualitative descriptions of included studies. None of the other SRs conducted any meta-analyses; instead, results of comparisons were described qualitatively.

Risk of bias—Systematic reviews

In the SR by Saensak et al,[45] authors did not report an overall risk of bias rating for individual studies. They noted that the overall quality of the evidence was "very low," implying high risk of bias. Concerns included lack of data, imprecision, failure to report methods adequately, and 2 RCTs whose data were not suitable for meta-analysis. They also noted that the included RCTs did not report important outcomes of interest including adverse events, night sweats, and sleep disturbances.

The SR by Cramer et al[49] reported that, overall, the studies they included had a low risk of bias using the Cochrane Risk of Bias tool.[33] However, in the risk of bias assessment for the 3 relevant studies for our review, one was noted to be at high risk for performance, detection, attrition, and other biases.[102] The second RCT was at unclear risk of performance and detection bias and high risk of reporting bias.[101] The third RCT was at unclear risk of selection, performance, and detection bias.[100]

New Randomized Controlled Trials

Study characteristics

We identified 6 relevant RCTs not included in the previous SRs that assessed the impact of meditation, mindfulness, hypnosis, or relaxation interventions on VMS or quality of life among perimenopausal or postmenopausal women.[56-61] Of the 6 RCTs, 3 examined paced respiration[56,57,60] and 3 examined applied relaxation.[58,59,61] We did not find any new RCTs on hypnosis. There was a total of 650 participants across the 6 RCTs, 446 for comparisons of paced respiration and 177 for relaxation.

Characteristics of the 6 new RCTs are summarized in Table 11.

Table 11. Study Characteristics of New RCTs

Study Country	**Population** # Women randomized Type of menopause # Hot flashes Mean age in years (range)	**Intervention** Category/type Session frequency/duration	**Comparator** Category/type Session frequency/duration	**General Outcomes** Instruments[a]
Paced respiration				
Carpenter 2013[56] USA	218 Mixed population[a] ≥2 hot flashes per day of moderate or greater severity 53.44 (6.84)	Paced respiration Home-based practice 15 minutes twice daily	Attention control: fast shallow breathing Twice daily Usual care: waitlist	· Hot flash frequency · Hot flash severity · Profile of Mood States-short form · Pittsburgh Sleep Quality Index · Adverse events
Huang 2015[57] USA	123 Perimenopausal or postmenopausal ≥4 per day 53.4 (3.4)	Paced respiration Brief, in-person instruction (<15 minutes); ≥15 minutes per day home practice	Attention control: music listening Brief, in-person instruction (<15 minutes); ≥15 minutes per day home practice	· Hot flash frequency · Hot flash severity · Menopause quality of life · Insomnia Severity Index · Beck Depression Inventory · Hospital Anxiety and Depression Scale
Sood 2013[60] USA	105 Mixed population[b] ≥14 per week, at least 1 month prior to enrollment 51	Paced breathing (2 arms) Initial in-person instruction: (1) 15 minutes once daily and (2) 15 minutes twice daily	Attention control: usual breathing Initial in-person instruction; 10 minutes per day	· Hot flash frequency · Hot flash severity · Profile of Mood States · Pittsburgh Sleep Quality Index · Adverse effects

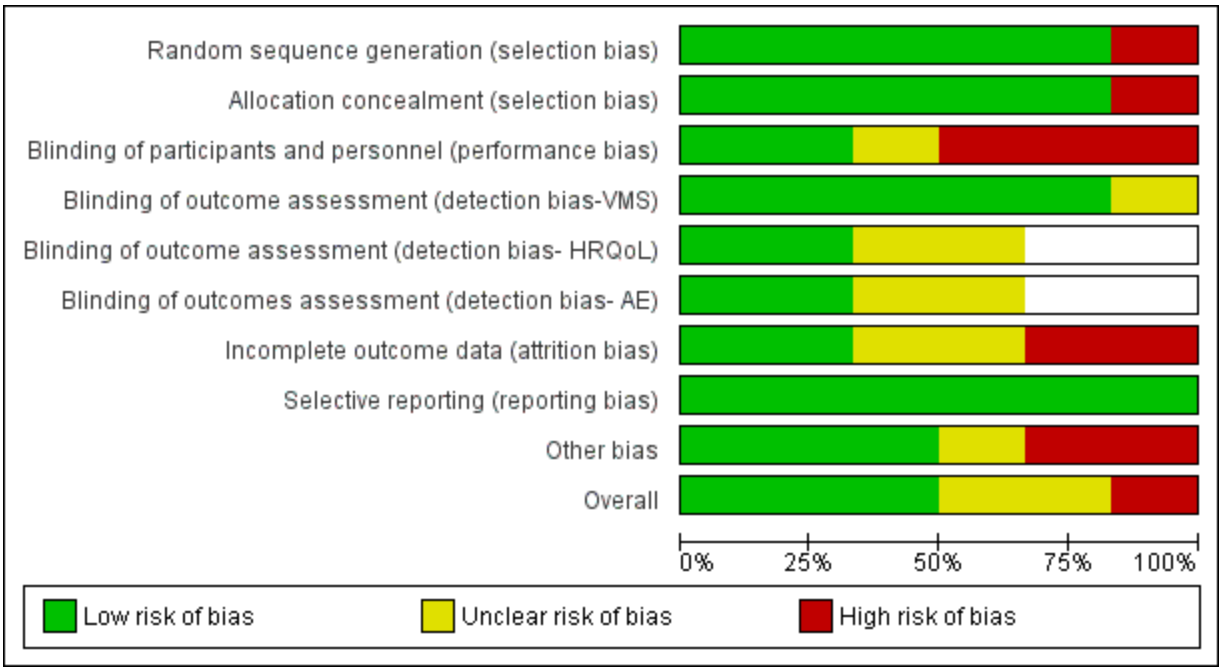
Study Country	**Population** # Women randomized Type of menopause # Hot flashes Mean age in years (range)	**Intervention** Category/type Session frequency/duration	**Comparator** Category/type Session frequency/duration	**General Outcomes** Instruments[a]
Applied relaxation				
Lindh-Astrand 2015[61] Sweden	46 Postmenopausal ≥7 moderate/severe per 24 hours Mean age: NR	Applied relaxation, internet-delivered 10 weeks; 9 weekly modules via group sessions Instructional CD	Untreated control	10 weeks: · Hot flash frequency
Lindh-Astrand 2013[58] Sweden	60 Postmenopausal ≥7 moderate to severe per 24 hours Applied relaxation: 54.0 (5.7) Control: 56.0 (5.1)	Applied relaxation 10 therapist-led group sessions over 12 weeks; 60 minutes per session Home practice	Untreated control: waitlist control	12 weeks: · Hot flash frequency · Women's Health Questionnaire: · Sleep subscale · Adverse events
Saensak 2013[59] Thailand	71 Perimenopausal and postmenopausal women[c] ≥ Hot flash severity score Modified relaxation: 49.8 (3.8) Applied relaxation: 52.5 (5.1)	Applied relaxation, modified relaxation version One 60-minute session of therapist-led training; daily home practice of 15-20 minutes per day at least 5 days per week Weekly phone contact with therapist	Applied relaxation: Conventional 12 weekly sessions of therapist-led training; 60 minutes per session; daily home practice; weekly phone contact with therapists	12 weeks: · Hot flash frequency · Hot flash severity · Sleep disturbance

[a] Menopause instruments and scales are described in Appendix D.
[b] Mixed population includes those with and without a history of breast cancer.
[c] Surgical or natural.

Risk of bias—New RCTs

Of the 6 new RCTs that we identified, one was judged to be at high risk of bias,[61] 2 at unclear risk of bias,[58,60] and 3 at low risk of bias[56,57,59] (Figure 13). For the new RCTs of paced respiration, there were 2 low risk[56,57] and one at unclear risk.[60] Of the 3 new relaxation RCTs, one was low risk,[59] one was unclear risk,[58] and one was high risk.[61] The one study at high risk was stopped prematurely due to an unacceptably high dropout rate (>60%). There were concerns regarding blinding of participants in all new RCTs except for one,[59] which was comparing 2 versions of applied relaxation. Similar to other types of interventions considered in this review, it is not uncommon to be unable to fully blind participants in relaxation, mindfulness, or hypnosis interventions. Another common concern among these trials was the measurement of self-reported outcomes among participants who were unblinded or incompletely blinded. Again, this concern is shared across all the intervention types included in this review.

Figure 13. Risk of Bias Ratings for Meditation, Mindfulness, Hypnosis, and Relaxation New RCTs[a]



[a] A white bar indicates that the outcome was not reported.

Synthesis of Findings: Primary Outcomes

For the primary outcomes of VMS and quality of life, we have organized the findings by intervention type (*ie*, meditation, mindfulness, hypnosis, and relaxation). We focus on findings from the good-quality SRs – Saensak et al[45] and Rada et al[48] – and supplement by using results from the new RCTs as applicable.

Meditation

We did not find an SR that addressed meditation, nor did we identify any recently published RCTs on this intervention type for treatment of VMS in perimenopausal or postmenopausal women.

Mindfulness

We did not find an SR that addressed mindfulness specifically, nor did we identify any recently published RCTs on this intervention type. One RCT[95] included in the fair-quality SR by Woods et al[46] compared mindfulness-based stress reduction with a waitlist control. In this trial of 110 women (mean age 53.1 years), there was no difference between groups in hot flash frequency or intensity. This same study noted that women in the mindfulness-based stress reduction group did experience significant improvements ($p<0.01$) in sleep quality and anxiety compared with a waitlist control at 9 weeks.

Hypnosis

Based on one fair-quality SR by Cramer et al[49] that included 3 trials of hypnosis on women with a history of breast cancer and those with hot flashes in the setting of menopause, there is evidence of a significant decrease in hot flashes compared with no treatment (effect size ranged from 0.479 to 1.25), though no difference for hypnosis was found compared with treatment with gabapentin.[100-102]

Active comparators

One RCT (n = 27) in Cramer et al[102] compared hypnosis with 900mg daily of gabapentin among breast cancer survivors and those at increased risk for breast cancer. Authors found that the median number of hot flashes decreased in both arms (80% versus 33.3% respectively), though the between-group difference was not statistically different. A similar reduction in hot flash severity as reported by the Hot Flash Severity Score in both arms was similarly not significant between groups.

Inactive comparators

The 2 RCTs[100,101] comparing hypnosis with no treatment/attention control in Cramer et al[49] found that hypnosis led to a larger decrease in hot flashes than no treatment/attention control, with effect sizes of 0.479 ($p < 0.001$) in one RCT of breast cancer survivors,[101] and 0.52 to 1.25 ($p < 0.001$) in an RCT of postmenopausal women without a history of breast cancer.[100]

Relaxation

The evidence, based on findings from the 4 RCTs in Saensak et al[45] plus one RCT from the multiple intervention SR by Rada et al[48] along with 6 of the new RCTs we identified,[56-61] does not support a significant reduction in hot flash frequency with relaxation techniques versus comparator interventions or placebo. One new RCT (n = 60)[58] noted a significant decrease in hot flash frequency among women treated with applied relaxation compared with control at 12 weeks (5.0 fewer hot flashes/24 hours compared with 1.9 fewer hot flashes/24 hours respectively [$P<0.001$]).

We conducted a new meta-analysis of 3 new RCTs[56,57,60] plus one RCT[97] included in the Saensak et al SR, all of which examined paced respiration compared with control. This meta-analysis supported the prior conclusion that paced respiration as a relaxation technique is not associated with a statistically significant decrease in hot flash frequency or severity. While quality of life was among the outcomes reported to be measured by some of these RCTs, it was measured less often than hot flash frequency or severity, and the results were rarely noted in the SRs.

Active comparators

Of the 4 RCTs included in Saensak et al,[45] 2 compared applied relaxation with acupuncture and were included in a meta-analysis that examined changes in frequency and severity of hot flashes.[98,99] Of note, follow-up data from these 2 RCTs were combined even though the timing of outcome assessments differed between studies (12 weeks and 24 weeks). In one of these, data from both the superficial needle insertion and electroacupuncture groups were combined into one "acupuncture" group.[99] There was no significant difference in the change of number of hot flashes per 24 hours between relaxation versus acupuncture (MD 0.05, 95% CI -1.33 to 1.43, I^2 = 0%, 2 trials). In a sensitivity analysis that excluded individuals who were in the superficial needle insertion group, effects were similar (MD 0.16, 95% CI -1.35 to 1.68, I^2 = 0%). Similarly, no significant difference was found in severity of hot flashes between groups using the Kupperman index (MD -1.32, 95% CI -5.06 to 2.43, I^2 = 0%, 2 trials). Saensak et al also included a 3-arm study (n = 11 per arm) evaluating paced respiration, muscle relaxation, and alpha-wave biofeedback (control).[97] The paced respiration group experienced a significant decrease in frequency of VMS per 24-hour period from the pretest period to the posttest period ($p<0.02$) while the muscle relaxation group did not. The difference between the groups was not significant.

One new RCT (n = 71)[59] randomized perimenopausal and postmenopausal women to a traditional 12-week applied relaxation training plus daily home practice compared with a modified version of applied relaxation that had a single in-person training session followed by daily home practice. Authors found similar reductions in hot flash frequency in both groups, though hot flash severity decreased more in the modified applied relaxation training group (p = 0.02).

Inactive comparators

Two RCTs[96,97] included in Saensak et al compared relaxation to inactive controls; however, they were not suitable for meta-analysis due to differences in outcomes reported and large standard deviations in one RCT. Neither found a significant difference in VMS frequency between groups. The SR by Rada et al[48] provided one additional small RCT comparing deep breathing and guided imagery with no treatment among women with breast cancer (n = 16),[94] finding no significant differences in the frequency of daily hot flashes (MD 0.50; 95% CI -1.23 to 2.23) or severity (MD 0.64; 95% CI -2.10 to 3.38) between groups. This RCT also found no difference in quality of life between groups.

Two new RCTs also made an inactive comparison.[58,61] One randomized healthy postmenopausal women (n = 60) to 10 sessions of therapist-led applied relaxation versus a waitlist control condition.[58] Authors found a significant decrease in hot flash frequency in the applied relaxation group compared with the control group at 12 weeks that was sustained at 3 months (5.0 fewer hot flashes/24 hours compared with 1.9 fewer hot flashes/24 hours respectively [$p < 0.001$]). A second RCT (n = 46) comparing internet-delivered applied relaxation with control found no differences in hot flash frequency between groups at 10 weeks.[61] This RCT was limited due to early termination in the setting of more than 60-percent dropout rate. Only 3 (13%) of the women assigned to internet-delivered applied relaxation completed at least 6 of the 9 planned modules.

We combined the findings from one RCT[97] included in Saensak et al with data from 3 new RCTs comparing paced respiration with control for 2 new meta-analyses.[56,57,60] For these meta-analyses, we combined control groups for the comparison as follows: attention control groups from 3 trials[56,57,60] were combined with an alpha-wave feedback control group from a fourth trial,[97] and 2 attention control comparator groups from one RCT[60] were combined together. We combined data from end-of-treatment across trials (range 4 to 16 weeks). None of these 4 RCTs included follow-up after completion of the intervention to assess for duration of treatment effects. In each of our 2 meta-analyses, only 3 of the 4 RCTs are included. Freedman et al[97] measured only hot flash frequency and although Sood et al[60] measured both frequency and severity, the data on frequency was not sufficient to be included in the meta-analysis. We found that paced respiration was not associated with a statistically significant decrease in hot flash frequency (SMD 0.04, 95% CI -0.73 to 0.82, I^2 = 56.6%, 3 trials) (Figure 14) or in hot flash severity (SMD 0.06, 95% CI -0.69 to 0.80, I^2 = 65.1%, 3 trials) (Figure 15), but 95% CIs were wide and do not exclude a small effect.

Figure 14. Forest Plot of Paced Respiration versus Control on Change in Hot Flash Frequency

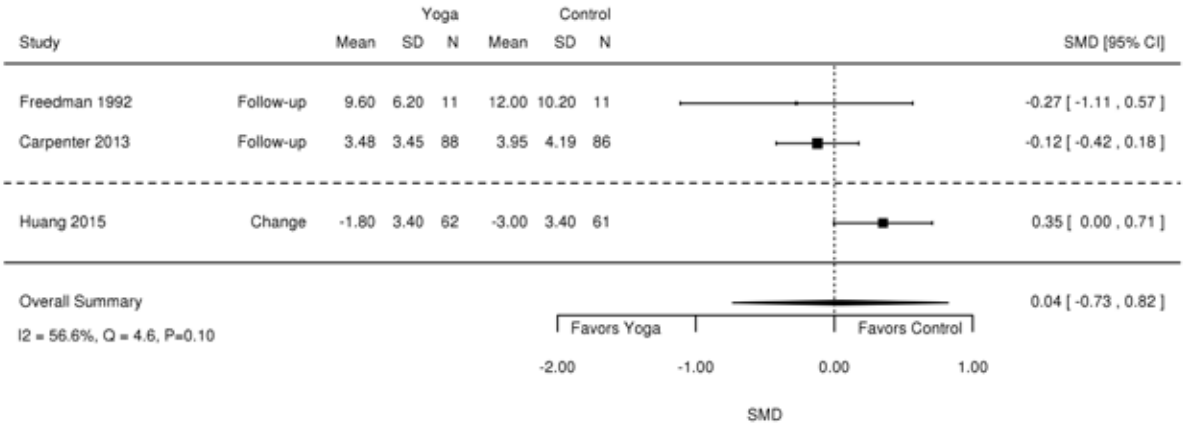
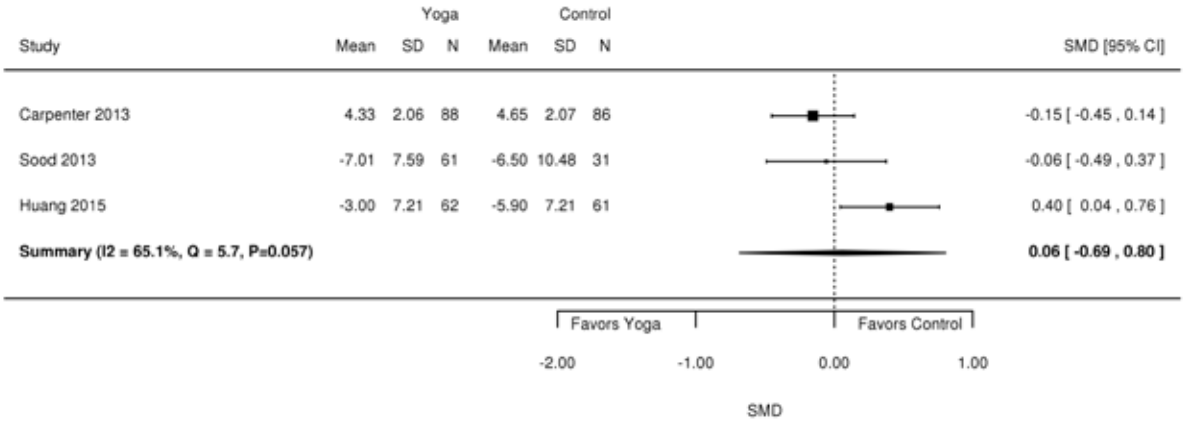


Figure 15. Forest Plot of Paced Respiration versus Control on Change in Hot Flash Severity at End-of-Treatment



Summary of Findings for Meditation, Mindfulness, Hypnosis, and Relaxation

We found that there is some support for the use of hypnosis to treat menopause-associated VMS, but the evidence does not support the use of relaxation techniques to alleviate VMS, and there is insufficient data on the role of mindfulness. To make this determination, we drew from 5 existing SRs (2 good and 3 fair quality)[45-49] and 6 new RCTs.[56-61] Overall, the quality of data is poor to mixed. For relaxation, there were 11 trials including 4 from one good-quality SR by Saensak et al,[45] 1 from a good-quality SR by Rada et al,[48] and 6 new RCTs.[56-61] Relaxation techniques studied were mostly applied relaxation and paced respiration. We added to the existing literature by conducting a new meta-analysis of paced respiration trials by combining data from 3 new RCTs and one RCT from the SR.[45] Similar to conclusions drawn in the SR, our meta-analysis found no significant treatment effect on VMS frequency (SMD 0.04, 95% CI -0.73 to 0.82; n = 162, I^2 = 56.6%) or severity (SMD 0.06, 95% CI -0.69 to 0.80; n = 211; I^2 = 65.1%). One small RCT[59] found that severity of hot flashes decreased more in a modified applied relaxation training group compared with traditional applied relaxation. One fair-quality SR of hypnosis[49] included 3 RCTs and found a significant decrease in hot flashes compared with no treatment (effect size ranged from 0.479 to 1.25), though no difference when compared to treatment with gabapentin. No new hypnosis trials were found. A fair-quality SR[46] included the only RCT[95] of mindfulness-based stress reduction compared with a waitlist control and found no difference in VMS frequency or severity.

SECONDARY OUTCOMES ACROSS ALL INTERVENTIONS

Acupuncture

Adverse effects

Four studies included in the SR by Dodin et al[40] reported mild bruising at the site of acupuncture needle insertion. Four other studies reported no adverse effects associated with acupuncture. Adverse effects, if any, reported in the other 7 RCTs were not discussed by Dodin et al. Adverse effects were not reported in the RCT by Avis et al.[52] No significant adverse events were reported

in the RCT by Ee et al,[50] but among the 327 patients allocated to either true or sham acupuncture, 8 (2%) reported bleeding or bruising, 7 (2%) reported pain, 3 (1%) reported syncope or presyncope, 3 (1%) reported worsening of symptoms, 1 reported tingling near an acupuncture point, 1 reported swelling around an acupuncture point and itching of the whole arm, 1 reported skin sensitivity and feeling hot, 1 reported nervousness, and 1 reported essential tremor. The RCT by Mao et al[51] reported 5 (17%) adverse effects in the real acupuncture arm, 1 (3%) in the sham arm, 8 (27%) in the placebo pill group, and 13 (43%) in the gabapentin arm (p value across the 4 groups was 0.005). All of the adverse effects were graded as mild.

Other secondary outcomes

Secondary outcomes such as sleep quality/quantity and symptoms of depression or anxiety were evaluated in some of the RCTs included in the SR by Dodin et al,[40] but the findings were not reported. Of the 3 new RCTs, the one by Avis et al[52] had the most extensive reporting of secondary outcomes. At the 6-month end-of-treatment assessment period in this trial of 209 perimenopausal or postmenopausal women, acupuncture was associated with significant improvement of sleep, when compared with a waitlist control, using the Pittsburgh Sleep Quality Index score, the sleep domain of the Women's Health Questionnaire, and the Patient-Reported Outcomes Measurement Information System (PROMIS) Sleep Disturbance short form. No between-group differences were found at 6 months by the different study instruments designed to assess depressive or anxiety symptoms.

Yoga

Adverse effects

Of the 5 RCTs in the SR by Cramer et al[43] plus the 2 new RCTs, only one reported assessing adverse events that are possibly associated with yoga, and found none.[87]

Other secondary outcomes

The SR[43] found positive long- and-short term effects of yoga on psychological symptoms, but no effect of yoga on sleep symptoms. Among the 2 new RCTs, one[55] found positive effects of yoga on psychological symptoms, but did not examine sleep outcomes. The other[54] was not sufficiently powered to detect differences and thus findings were inconsistent. However, participants in the yoga intervention did not experience meaningful changes in sleep quality, perceived stress, or depression.

Structured Exercise

Adverse effects

No serious adverse events were reported in the 5 RCTs included in the SR by Daley et al[44] for any of the outcomes. In one RCT, the proportion of incident adverse events were similar across the exercise group (17%) and the usual activity group (18%).[91]

Other secondary outcomes

One new RCT[62] measured menopause-specific quality of life using subscales from the Women's Health Questionnaire. That study found significantly lower sleep problem scores (mean difference -0.11, 95% CI -0.21 to -0.01) among the exercise group when compared with control

at the 6-month follow up. No statistically significant effects were found for anxiety (mean difference -1, 95% CI -2.0 to 0.02) at the 12-month follow-up.

One new RCT[63] measured generic quality of life using the SF-36 Short Form Health Survey. Moderate effect sizes were found for short (effect size 0.46, $p<0.001$) and long-term follow-up (effect size 0.41, $p = 0.002$) among the exercise group compared with control for physical function.

Four years after the intervention, Luoto et al[90] published a companion paper using the original cohort[103] examining long-term quality of life measured by the SF-36 Health Survey Questionnaire. In the companion study, the exercise group had a significantly higher physical functioning (OR 1.41, 95% CI 1.00 to 1.99) compared with the control group. No significant associations were found for higher good role functioning (OR 1.21, 95% CI 0.88 to 1.67), physical health (OR 1.33, 95% CI 0.96 to 1.84), or general health (OR 1.14, 95% CI 0.81 to 1.62).[103]

An additional companion publication to Luoto et al[90] focused on sleep quality.[104] Outcomes for this study were sleep quality and the amount of hot flashes during sleep measured by mobile phone with a 1 (poor) to 5 (good) scale response to each question. Sleep quality was significantly improved in the exercise intervention group compared with the control (OR 1.02, 95% CI 1.0 to 1.05, $p = 0.043$).[104] The amount of hot flashes related to sleep diminished ($p = 0.004$) by the end of the intervention.

Meditation, Mindfulness, Hypnosis, and Relaxation

Adverse effects

Overall, adverse effects were rarely reported and when noted were mild. For interventions of hypnosis, the SR by Cramer et al[49] found 3 adverse events in the gabapentin control group in one included RCT,[102] and no dropouts due to adverse events or no adverse events related to the intervention in 2 other RCTs.[100,101] One new RCT of paced breathing compared with control[60] noted that more women in the paced breathing arm reported mild dizziness than in the control arm.

Other secondary outcomes

In general, secondary outcomes of interest were rarely reported in the SRs addressing relaxation, hypnosis, or mindfulness, and the results were mixed when they were reported. In our examination of reported secondary outcomes in the SRs and new RCTs, the findings were generally not significant, with a few exceptions. In one RCT included by Cramer et al,[101] hypnosis decreased depression and anxiety compared with no treatment, though no specific results were given. In the one RCT of mindfulness included in the SR by Woods et al,[46] mindfulness-based stress reduction significantly improved both anxiety and sleep quality at 9 weeks compared with a waitlist control group ($p<0.01$).

SUMMARY AND DISCUSSION

We evaluated a broad range of nonpharmacologic interventions for perimenopausal and postmenopausal women with VMS, evaluating effects on VMS, health-related quality of life, and adverse events. Our review identified good- and fair-quality SRs for all of the eligible intervention categories. In addition, we identified 14 new RCTs, representing a 42% increase in RCTs in the past 4 years. Based on the number of trials, the evidence base is most developed for acupuncture, followed by relaxation, then exercise and yoga. Except for acupuncture, most trials used waitlist controls as the comparator and reported outcomes at 6 months or earlier. There were few comparative effectiveness trials. However, the growing evidence base allowed for updated meta-analyses and estimates of effect for most interventions compared with inactive or attention controls. Almost all studies reported effects on VMS frequency or severity; fewer reported effects on quality of life, insomnia, or psychological symptoms. Adverse effects were rare but often not reported systematically.

LIMITATIONS

Limitations of the Umbrella Review

The novel approach of supplementing a review of reviews with findings from recently published RCTs allowed for assessment of a broad range of interventions. This approach made it possible for us to synthesize both quantitatively and qualitatively the most current information for each of the 4 categories of nonpharmacologic interventions for VMS. A significant limitation to this approach is that we relied on the authors' judgments about risk of bias for individual trials and the appropriateness of their search strategies, eligibility criteria, and synthesis of the evidence. We confirmed (and at times corrected) data included in meta-analysis that we updated, but we did not confirm all of the study characteristics and outcomes data reported in the SRs. We limited our review to English-language publications, which may have excluded potentially informative evidence.

Limitations of the Studies

Most of the RCTs discussed in our report were relatively small, short-term trials. Adverse events were often not reported. All of the studies used self-report assessments, and most did not mask participants to intervention allocation, thereby introducing the risk that patients allocated to the active interventions might exaggerate clinical improvement. Few of the trials compared 2 or more active interventions, which could have informed clinical decision-making for patients or healthcare providers faced with deciding among various therapeutic options. Several of the trials used usual care or waitlist controls, which do not control for nonspecific effects of a given intervention, and as such do not provide insights about an intervention's potential mechanisms of action. However, usual care controls serve as an appropriate control for trials that aim to inform patients or healthcare providers what to expect from a given course of treatment relative to not undergoing that course of treatment.[105] Trials that include a usual care arm introduce a risk of performance bias, which may either overestimate or underestimate the true effect size of an intervention.[106]

Approximately half of the acupuncture trials used sham acupuncture controls. The purpose of a sham procedure or attention control is to evaluate whether a given intervention's mechanisms of

action involve physiological processes that are independent of nonspecific effects that can be attributed to health care providers' care and attention or individual patients' beliefs and expectations. In the case of acupuncture, there is considerable debate about the appropriate control.[107,108] Some argue that sham procedures may not be physiologically inert, and that a usual care arm may be a more appropriate comparison for trials that aim to inform clinical practice, as opposed to determining the specific effects of an intervention.

Clinical interpretation of the findings from this report was also hampered by the use of many different outcome measures across trials, thereby requiring us to report standardized mean differences, which are difficult to interpret clinically. A more readily interpretable outcome would be the proportion of women achieving the minimum clinically important response, but this value has not been established for any of the primary outcomes used in these studies. This was especially true for quality-of-life measures, with the additional caveat that there is no universally accepted quality-of-life assessment instrument specific to menopausal symptoms. Authors of the SRs frequently considered symptom inventories or functional status questionnaires as health-related or menopausal-specific quality-of-life instruments. We used the same terms reported by the authors of the SRs, with the understanding that the underlying constructs measured are not likely to be truly health-related or menopausal-specific quality of life.

Unexplained heterogeneity evident in some of our meta-analyses represents another limitation of the existing evidence. The source of this heterogeneity across studies is probably multifactorial, with poor study quality, variable patient eligibility criteria, the use of a wide variety of different outcome instruments, differing doses and duration of treatments, and other factors all likely contributing to statistical heterogeneity.

SUMMARY OF EVIDENCE

Strength of Evidence (SOE)

We found evidence that acupuncture (moderate SOE) and yoga (low SOE) improve VMS more than controls. A trial that compared acupuncture to enhanced self-care in women with hot flashes and breast cancer was published after our search date, and also found benefit for hot flashes.[109] However, traditional acupuncture was no better than sham acupuncture. Hypnosis was effective in women with breast cancer (low SOE). Neither structured exercise (moderate SOE) nor paced respiration (low SOE) improved VMS. Evidence was insufficient for applied relaxation and mindfulness-based stress reduction. No trials evaluated qigong, tai chi, or meditation.

Effects of interventions on health-related quality of life were reported less frequently, more inconsistently, and using measures that did not always conform to standard definitions for this construct. Consistent with effects on VMS, acupuncture improved health-related quality of life when compared with a waitlist control (moderate SOE) but not when compared with sham acupuncture (moderate SOE). Yoga did not improve health-related quality of life (insufficient SOE). The SOE for adverse effects was rated insufficient because of inconsistent reporting.

In Table 12, we summarize the SOE for effects of interventions compared with control on VMS—the priority symptom for many perimenopausal and postmenopausal women.[110] Few trials compared these interventions with active interventions such as hormone therapy or

gabapentin, and none of the interventions were compared with antidepressants. Because of the sparseness of evidence for these comparisons, we did not rate the SOE.

Table 12. Strength of Evidence for Effects of Interventions on VMS

Comparison	# RCTs (Patients)	Findings	Strength of Evidence (Rationale by Domain)
Acupuncture vs waitlist	4 (501)	SMD 0.66 lower (1.06 lower to 0.26 lower)	Moderate Low ROB, consistent, direct, imprecise
Acupuncture vs sham acupuncture	8 (644)	SMD 0.35 lower (0.70 lower to 0.01 higher)	Moderate Low ROB, consistent, direct, imprecise
Yoga vs control	4 (157)	SMD 0.36 lower (0.65 lower to 0.07 lower)	Low Moderate ROB, consistent, direct, imprecise
Structured exercise vs control	4 (431)	SMD 0.08 lower (0.33 lower to 0.16 higher)	Moderate Moderate ROB, consistent, direct, precise
Paced respiration vs control	3 (161)	SMD 0.04 higher (0.73 lower to 0.82 higher)	Low Low ROB, inconsistent, direct, imprecise
Applied relaxation vs control	2 (82)	1 RCT showed small benefit and 1 RCT showed no effect	Insufficient Moderate ROB, inconsistent, direct, imprecise
Hypnosis vs control	3 (274)	No pooled estimate. Effect size ranged from 0.479 to 1.25	Low Moderate ROB, consistent, direct, imprecise
Mindfulness-based stress reduction vs control	1 (110)	No reduction in hot flashes	Not rated

Abbreviations: RCT = randomized controlled trial; ROB = risk of bias; SMD = standardized mean difference

CLINICAL IMPLICATIONS

The VA does not have a current guideline that addresses the management of menopause in general, and few clinical practice guidelines specifically address nonpharmacologic treatment strategies for VMS. One exception is a 2015 position statement from the North American Menopause Society.[25] This statement recommends cognitive behavioral therapy, hypnosis, and, with caution, mindfulness-based stress reduction. Exercise, yoga, paced respiration, relaxation, and acupuncture were considered but not recommended. Our updated analysis shows benefit for yoga, acupuncture, and hypnosis compared (primarily) with no treatment and only very limited evidence for mindfulness-based stress reduction. These updated results should be considered when making clinical or policy recommendations.

RESEARCH GAPS/FUTURE RESEARCH

For some interventions such as acupuncture and yoga, there is evidence of benefit. However, SOE is low to moderate, so larger, high-quality trials are needed. Comparative effectiveness trials would be more likely to inform policy and clinical decision-making than sham- or placebo-controlled effectiveness trials. This may be especially true in the search for alternatives to pharmacologic approaches to managing menopausal symptoms, where clinical effectiveness

outcomes may need to be counterbalanced by other outcomes of importance to women, healthcare providers, and policymakers, such as potential harm, cost, overall utility, and women's preferences. Research is needed, as well, to better understand considerations related to treatment options that women with menopausal symptoms consider to be most important.

The design of some of the RCTs included in our review can inform future research. For example, the internet-based, applied relaxation trial by Lindh-Astrand et al[61] is innovative in that it explored a way to make treatment more readily accessible and affordable. However, this trial had an exceptionally high drop-out rate. More research is needed to better understand how to engage and retain patients in internet-based interventions. Pragmatic trials such as the one recently published by Avis et al[52] simulated clinical practice by allowing patients and acupuncturists to negotiate the number and frequency of acupuncture treatments and allowing the acupuncturists to design and administer treatments as they would outside the context of the clinical trial. Nonpharmacologic clinical interventions such as acupuncture, yoga, meditation, and exercise lend themselves well to these types of pragmatic trials, with the caveat that pragmatic, comparative effectiveness trials are not designed to determine mechanisms of actions or to estimate the extent to which observed clinical benefit associated with a given intervention may be attributable to specific versus nonspecific effects.

None of the RCTs included in this umbrella review specifically involved Veterans. Additional research is needed to evaluate the acceptability, feasibility, and comparative effectiveness of nonpharmacologic approaches to managing menopausal symptoms for women Veterans in VA primary care clinics, as well as other settings and patient populations such as medically underserved populations.

CONCLUSIONS

Compared with waitlist controls, evidence from RCTs support acupuncture and yoga for reducing VMS and the impact of such symptoms on women's activities and health-related quality life. The strength of evidence, however, is low to moderate. Moderately good evidence shows no benefit from structured exercise for VMS, but engaging in exercise is known to be important for other reasons. The evidence in support of the effectiveness of mindfulness or relaxation is mixed, with some promising evidence that needs replication for hypnosis. There is insufficient evidence to draw conclusions about the effectiveness of these nonpharmacologic therapies for improving sleep, depression, or anxiety. The safety of the nonpharmacologic, nonhormonal approaches evaluated in this report has not been rigorously examined, but there is no clear signal for a significant potential for harm. Overall, most of the data included in this report comes from smaller studies with homogenous participant populations. Larger trials of populations more reflective of the diversity of women experiencing VMS will be necessary to discern the effectiveness of nonpharmacologic interventions in symptomatic menopausal women.

REFERENCES

1. Woods NF, Mitchell ES. Symptoms during the perimenopause: prevalence, severity, trajectory, and significance in women's lives. *Am J Med.* 2005;118 Suppl 12B:14-24.

2. Santoro N. Perimenopause: From Research to Practice. *J Womens Health (Larchmt).* 2015.

3. Avis NE, Crawford SL, Greendale G, et al. Duration of menopausal vasomotor symptoms over the menopause transition. *JAMA Intern Med.* 2015;175(4):531-539.

4. Thurston RC, Sowers MR, Sternfeld B, et al. Gains in body fat and vasomotor symptom reporting over the menopausal transition: the study of women's health across the nation. *Am J Epidemiol.* 2009;170(6):766-774.

5. Guthrie JR, Dennerstein L, Taffe JR, Donnelly V. Health care-seeking for menopausal problems. *Climacteric.* 2003;6(2):112-117.

6. Avis NE, Colvin A, Bromberger JT, et al. Change in health-related quality of life over the menopausal transition in a multiethnic cohort of middle-aged women: Study of Women's Health Across the Nation. *Menopause.* 2009;16(5):860-869.

7. Ayers B, Hunter MS. Health-related quality of life of women with menopausal hot flushes and night sweats. *Climacteric.* 2013;16(2):235-239.

8. Bhattacharya SM, Jha A. A comparison of health-related quality of life (HRQOL) after natural and surgical menopause. *Maturitas.* 2010;66(4):431-434.

9. Thurston RC, Bromberger JT, Joffe H, et al. Beyond frequency: who is most bothered by vasomotor symptoms? *Menopause.* 2008;15(5):841-847.

10. North American Menopause Society. 2012 hormone therapy position statement. *Menopause.* 2012;19(3):257-271.

11. Menopause: diagnosis and management. Vol NICE Guidelines [NG23]. November 2015. Available at: https://www.nice.org.uk/guidance/ng23. Accessed January 25, 2016.

12. The North American Menopause Society Statement on Continuing Use of Systemic Hormone Therapy After Age 65. *Menopause.* 2015;22(7):693.

13. U.S. Food and Drug Administration. Guidance for Industry: Estrogen and Estrogen/Progestin Drug Products to Treat Vasomotor Symptoms and Vulvar and Vaginal Atrophy Symptoms—Recommendations for Clinical Evaluation. (Draft guidance from the Center for Drug Evaluation and Research (CDER), January 2003). Available at: http://www.fda.gov/downloads/Drugs/DrugSafety/InformationbyDrugClass/UCM135338.pdf. Accessed January 27, 2016.

14. ACOG Practice Bulletin No. 141: management of menopausal symptoms. *Obstet Gynecol.* 2014;123(1):202-216.

15. 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.

16. Das SR, Kinsinger LS, Yancy WS, Jr., et al. Obesity prevalence among veterans at Veterans Affairs medical facilities. *Am J Prev Med.* 2005;28(3):291-294.

17. Farmer MM, Rose DE, Riopelle D, Lanto AB, Yano EM. Gender differences in smoking and smoking cessation treatment: an examination of the organizational features related to care. *Womens Health Issues.* 2011;21(4 Suppl):S182-189.

18. Lehavot K, Hoerster KD, Nelson KM, Jakupcak M, Simpson TL. Health indicators for military, veteran, and civilian women. *Am J Prev Med.* 2012;42(5):473-480.

19. Gerber MR, King MW, Pineles SL, et al. Hormone therapy use in women veterans accessing veterans health administration care: a national cross-sectional study. *J Gen Intern Med.* 2015;30(2):169-175.

20. Taylor M. Complementary and Alternative Approaches to Menopause. *Endocrinol Metab Clin North Am.* 2015;44(3):619-648.

21. Gartoulla P, Davis SR, Worsley R, Bell RJ. Use of complementary and alternative medicines for menopausal symptoms in Australian women aged 40-65 years. *Med J Aust.* 2015;203(3):146, 146e.141-146.

22. Gentry-Maharaj A, Karpinskyj C, Glazer C, et al. Use and perceived efficacy of complementary and alternative medicines after discontinuation of hormone therapy: a nested United Kingdom Collaborative Trial of Ovarian Cancer Screening cohort study. *Menopause.* 2015;22(4):384-390.

23. National Center for Complementary and Integrative Health. Complementary, Alternative, or Integrative Health: What's In a Name? Available at: https://nccih.nih.gov/health/integrative-health. Accessed March 10, 2016.

24. Grant MD, Marbella A, Wang AT, et al. 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.

25. The North American Menopause Society. Nonhormonal management of menopause-associated vasomotor symptoms: 2015 position statement of The North American Menopause Society. *Menopause.* 2015;22(11):1155-1174.

26. Becker LA, Oxman AD. Chapter 22: Overviews of reviews. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available at: www.cochrane-handbook.org. Accessed October 21, 2015.

27. Agency for Healthcare Research and Quality. Methods Guide for Effectiveness and Comparative Effectiveness Reviews: Rockville, MD: Agency for Healthcare Research and Quality. Available at: http://www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction = displayproduct&productid = 318. Accessed July 16, 2015.

28. Goldstein KM, McDuffie JR, Shepherd-Banigan M, et al. Nonpharmacologic, nonherbal management of menopause-associated vasomotor symptoms: an umbrella systematic review (protocol). *Syst Rev.* 2016;5(1):56.

29. PROSPERO: International prospective register of systematic reviews. #CRD42016029335: Available at: http://www.crd.york.ac.uk/prospero/. Accessed January 25, 2016.

30. Whiting P, Savovic J, Higgins JP, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol.* 2015.

31. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol.* 2011;64(4):383-394.

32. Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol.* 2011;64(4):401-406.

33. Higgins JPT, Altman DG, Gøtzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ.* 2011;343.

34. Robinson KA, Chou R, Berkman ND, et al. Twelve recommendations for integrating existing systematic reviews into new reviews: EPC guidance. *J Clin Epidemiol.* 2015.

35. Engauge Digitizer (software package). Available at: http://markummitchell.github.io/engauge-digitizer/. Accessed February 1, 2016.

36. Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package (2010). Available at: https://www.jstatsoft.org/article/view/v036i03. Accessed April 6, 2016.

37. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials.* 1986;7(3):177-188.

38. Knapp G, Hartung J. Improved tests for a random effects meta-regression with a single covariate. *Stat Med.* 2003;22(17):2693-2710.

39. Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, N.J.: L. Erlbaum Associates; 1988.

40. Dodin S, Blanchet C, Marc I, et al. Acupuncture for menopausal hot flushes. *Cochrane Database Syst Rev.* 2013;7:CD007410.

41. Garcia MK, Graham-Getty L, Haddad R, et al. Systematic review of acupuncture to control hot flashes in cancer patients. *Cancer.* 2015;121(22):3948-3958.

42. Garcia MK, McQuade J, Lee R, Haddad R, Spano M, Cohen L. Acupuncture for symptom management in cancer care: an update. *Curr Oncol Rep.* 2014;16(12):418.

43. Cramer H, Lauche R, Langhorst J, Dobos G. Effectiveness of yoga for menopausal symptoms: a systematic review and meta-analysis of randomized controlled trials. *Evid Based Complement Alternat Med.* 2012;2012:863905.

44. Daley A, Stokes-Lampard H, Thomas A, MacArthur C. Exercise for vasomotor menopausal symptoms. *Cochrane Database Syst Rev.* 2014;11:Cd006108.

45. Saensak S, Vutyavanich T, Somboonporn W, Srisurapanont M. Relaxation for perimenopausal and postmenopausal symptoms. *Cochrane Database Syst Rev.* 2014;7:CD008582.

46. Woods NF, Mitchell ES, Schnall JG, et al. Effects of mind-body therapies on symptom clusters during the menopausal transition. *Climacteric.* 2014;17(1):10-22.

47. Innes KE, Selfe TK, Vishnu A. Mind-body therapies for menopausal symptoms: a systematic review. *Maturitas.* 2010;66(2):135-149.

48. Rada G, Capurro D, Pantoja T, et al. Non-hormonal interventions for hot flushes in women with a history of breast cancer. *Cochrane Database Syst Rev.* 2010(9):Cd004923.

49. Cramer H, Lauche R, Paul A, Langhorst J, Kummel S, Dobos GJ. Hypnosis in breast cancer care: a systematic review of randomized controlled trials. *Integr Cancer Ther.* 2015;14(1):5-15.

50. Ee C, Xue C, Chondros P, et al. Acupuncture for Menopausal Hot Flashes: A Randomized Trial. *Ann Intern Med.* 2016;164(3):146-154.

51. Mao JJ, Bowman MA, Xie SX, Bruner D, DeMichele A, Farrar JT. Electroacupuncture Versus Gabapentin for Hot Flashes Among Breast Cancer Survivors: A Randomized Placebo-Controlled Trial. *J Clin Oncol.* 2015;33(31):3615-3620.

52. Avis NE, Coeytaux RR, Isom S, Prevette K, Morgan T. Acupuncture in Menopause (AIM) study: a pragmatic, randomized controlled trial. *Menopause.* 2016.

53. Nedeljkovic M, Tian L, Ji P, et al. Effects of acupuncture and Chinese herbal medicine (Zhi Mu 14) on hot flushes and quality of life in postmenopausal women: results of a four-arm randomized controlled pilot trial. *Menopause.* 2014;21(1):15-24.

54. Avis NE, Legault C, Russell G, Weaver K, Danhauer SC. Pilot study of integral yoga for menopausal hot flashes. *Menopause.* 2014;21(8):846-854.

55. Ngowsiri K, Tanmahasamut P, Sukonthasab S. Rusie Dutton traditional Thai exercise promotes health related physical fitness and quality of life in menopausal women. *Complement Ther Clin Pract.* 2014;20(3):164-171.

56. Carpenter JS, Burns DS, Wu J, et al. Paced respiration for vasomotor and other menopausal symptoms: a randomized, controlled trial. *J Gen Intern Med.* 2013;28(2):193-200.

57. Huang AJ, Phillips S, Schembri M, Vittinghoff E, Grady D. Device-guided slow-paced respiration for menopausal hot flushes: a randomized controlled trial. *Obstet Gynecol.* 2015;125(5):1130-1138.

58. Lindh-Astrand L, Nedstrand E. Effects of applied relaxation on vasomotor symptoms in postmenopausal women: a randomized controlled trial. *Menopause.* 2013;20(4):401-408.

59. Saensak S, Vutyavanich T, Somboonporn W, Srisurapanont M. Effectiveness of a modified version of the applied relaxation technique in treatment of perimenopausal and postmenopausal symptoms. *Int J Womens Health.* 2013;5:765-771.

60. Sood R, Sood A, Wolf SL, et al. Paced breathing compared with usual breathing for hot flashes. *Menopause.* 2013;20(2):179-184.

61. Lindh-Astrand L, Holm ACS, Sydsjö G, Andersson G, Carlbring P, Nedstrand E. Internet-delivered applied relaxation for vasomotor symptoms in postmenopausal women: Lessons from a failed trial. *Maturitas.* 2015;80(4):432-434.

62. Daley AJ, Thomas A, Roalfe AK, et al. The effectiveness of exercise as treatment for vasomotor menopausal symptoms: randomised controlled trial. *BJOG.* 2015;122(4):565-575.

63. Duijts SF, van Beurden M, Oldenburg HS, et al. Efficacy of cognitive behavioral therapy and physical exercise in alleviating treatment-induced menopausal symptoms in patients with breast cancer: results of a randomized, controlled, multicenter trial. *J Clin Oncol.* 2012;30(33):4124-4133.

64. Garcia MK, McQuade J, Haddad R, et al. Systematic review of acupuncture in cancer care: a synthesis of the evidence. *J Clin Oncol.* 2013;31(7):952-960.

65. Walker EM, Rodriguez AI, Kohn B, et al. Acupuncture versus venlafaxine for the management of vasomotor symptoms in patients with hormone receptor-positive breast cancer: a randomized controlled trial. *J Clin Oncol.* 2010;28(4):634-640.

66. Liljegren A, Gunnarsson P, Landgren BM, Robeus N, Johansson H, Rotstein S. Reducing vasomotor symptoms with acupuncture in breast cancer patients treated with adjuvant tamoxifen: a randomized controlled trial. *Breast Cancer Res Treat.* 2012;135(3):791-798.

67. Bao T, Cai L, Snyder C, et al. Patient-reported outcomes in women with breast cancer enrolled in a dual-center, double-blind, randomized controlled trial assessing the effect of acupuncture in reducing aromatase inhibitor-induced musculoskeletal symptoms. *Cancer.* 2014;120(3):381-389.

68. Vincent A, Barton DL, Mandrekar JN, et al. Acupuncture for hot flashes: a randomized, sham-controlled clinical study. *Menopause.* 2007;14(1):45-52.

69. Venzke L, Calvert JF, Jr., Gilbertson B. A randomized trial of acupuncture for vasomotor symptoms in post-menopausal women. *Complement Ther Med.* 2010;18(2):59-66.

70. Park JE, Lee MS, Jung S, et al. Moxibustion for treating menopausal hot flashes: a randomized clinical trial. *Menopause.* 2009;16(4):660-665.

71. Painovich JM, Shufelt CL, Azziz R, et al. A pilot randomized, single-blind, placebo-controlled trial of traditional acupuncture for vasomotor symptoms and mechanistic pathways of menopause. *Menopause.* 2012;19(1):54-61.

72. Nir Y, Huang MI, Schnyer R, Chen B, Manber R. Acupuncture for postmenopausal hot flashes. *Maturitas.* 2007;56(4):383-395.

73. Nedstrand E, Wyon Y, Hammar M, Wijma K. Psychological well-being improves in women with breast cancer after treatment with applied relaxation or electro-acupuncture for vasomotor symptom. *J Psychosom Obstet Gynaecol.* 2006;27(4):193-199.

74. Kim DI, Jeong JC, Kim KH, et al. Acupuncture for hot flushes in perimenopausal and postmenopausal women: a randomised, sham-controlled trial. *Acupunct Med.* 2011;29(4):249-256.

75. Kim KH, Kang KW, Kim DI, et al. Effects of acupuncture on hot flashes in perimenopausal and postmenopausal women--a multicenter randomized clinical trial. *Menopause.* 2010;17(2):269-280.

76. Hervik J, Mjaland O. Acupuncture for the treatment of hot flashes in breast cancer patients, a randomized, controlled trial. *Breast Cancer Res Treat.* 2009;116(2):311-316.

77. Deng G, Vickers A, Yeung S, et al. Randomized, controlled trial of acupuncture for the treatment of hot flashes in breast cancer patients. *J Clin Oncol.* 2007;25(35):5584-5590.

78. Borud EK, Alraek T, White A, et al. The Acupuncture on Hot Flushes Among Menopausal Women (ACUFLASH) study, a randomized controlled trial. *Menopause.* 2009;16(3):484-493.

79. Bokmand S, Flyger H. Acupuncture relieves menopausal discomfort in breast cancer patients: a prospective, double blinded, randomized study. *Breast.* 2013;22(3):320-323.

80. Avis NE, Legault C, Coeytaux RR, et al. A randomized, controlled pilot study of acupuncture treatment for menopausal hot flashes. *Menopause.* 2008;15(6):1070-1078.

81. Zhou J, Qu F, Sang X, Wang X, Nan R. Acupuncture and auricular acupressure in relieving menopausal hot flashes of bilaterally ovariectomized chinese women: a randomized controlled trial. *Evid Based Complement Alternat Med.* 2011;2011:713274.

82. Frisk J, Carlhall S, Kallstrom AC, Lindh-Astrand L, Malmstrom A, Hammar M. Long-term follow-up of acupuncture and hormone therapy on hot flushes in women with breast cancer: a prospective, randomized, controlled multicenter trial. *Climacteric.* 2008;11(2):166-174.

83. Wyon Y, Wijma K, Nedstrand E, Hammar M. A comparison of acupuncture and oral estradiol treatment of vasomotor symptoms in postmenopausal women. *Climacteric.* 2004;7(2):153-164.

84. Cornell JE, Mulrow CD, Localio R, et al. Random-effects meta-analysis of inconsistent effects: a time for change. *Ann Intern Med.* 2014;160(4):267-270.

85. Elavsky S, McAuley E. Physical activity and mental health outcomes during menopause: a randomized controlled trial. *Ann Behav Med.* 2007;33(2):132-142.

86. Joshi S, Khandwe R, Bapat D, Deshmukh U. Effect of yoga on menopausal symptoms. *Menopause Int.* 2011;17(3):78-81.

87. Afonso RF, Hachul H, Kozasa EH, et al. Yoga decreases insomnia in postmenopausal women: a randomized clinical trial. *Menopause.* 2012;19(2):186-193.

88. Carson JW, Carson KM, Porter LS, Keefe FJ, Seewaldt VL. Yoga of Awareness program for menopausal symptoms in breast cancer survivors: results from a randomized trial. *Support Care Cancer.* 2009;17(10):1301-1309.

89. Chattha R, Raghuram N, Venkatram P, Hongasandra NR. Treating the climacteric symptoms in Indian women with an integrated approach to yoga therapy: a randomized control study. *Menopause.* 2008;15(5):862-870.

90. Luoto R, Moilanen J, Heinonen R, et al. Effect of aerobic training on hot flushes and quality of life--a randomized controlled trial. *Ann Med.* 2012;44(6):616-626.

91. Sternfeld B, Guthrie KA, Ensrud KE, et al. Efficacy of exercise for menopausal symptoms: a randomized controlled trial. *Menopause.* 2014;21(4):330-338.

92. Lindh-Astrand L, Nedstrand E, Wyon Y, Hammar M. Vasomotor symptoms and quality of life in previously sedentary postmenopausal women randomised to physical activity or estrogen therapy. *Maturitas.* 2004;48(2):97-105.

93. Hanachi P, Golkho S. Assessment of soy phytoestrogens and exercise on lipid profiles and menopause symptoms in menopausal women. *Journal of Biological Sciences.* 2008;8(4):789-793.

94. Fenlon D. Relaxation therapy as an intervention for hot flushes in women with breast cancer. *Eur J Oncol Nurs.* 1999;3(4):223-231.

95. Carmody J, Crawford S, Churchill L. A pilot study of mindfulness-based stress reduction for hot flashes. *Menopause.* 2006;13(5):760-769.

96. Fenlon DR, Corner JL, Haviland JS. A randomized controlled trial of relaxation training to reduce hot flashes in women with primary breast cancer. *J Pain Symptom Manage.* 2008;35(4):397-405.

97. Freedman RR, Woodward S. Behavioral treatment of menopausal hot flushes: evaluation by ambulatory monitoring. *Am J Obstet Gynecol.* 1992;167(2):436-439.

98. Nedstrand E, Wijma K, Wyon Y, Hammar M. Vasomotor symptoms decrease in women with breast cancer randomized to treatment with applied relaxation or electro-acupuncture: a preliminary study. *Climacteric.* 2005;8(3):243-250.

99. Zaborowska E, Brynhildsen J, Damberg S, et al. Effects of acupuncture, applied relaxation, estrogens and placebo on hot flushes in postmenopausal women: an analysis of two prospective, parallel, randomized studies. *Climacteric.* 2007;10(1):38-45.

100. Elkins GR, Fisher WI, Johnson AK, Carpenter JS, Keith TZ. Clinical hypnosis in the treatment of postmenopausal hot flashes: a randomized controlled trial. *Menopause.* 2013;20(3):291-298.

101. Elkins G, Marcus J, Stearns V, et al. Randomized trial of a hypnosis intervention for treatment of hot flashes among breast cancer survivors. *J Clin Oncol.* 2008;26(31):5022-5026.

102. Maclaughlan David S, Salzillo S, Bowe P, et al. Randomised controlled trial comparing hypnotherapy versus gabapentin for the treatment of hot flashes in breast cancer survivors: a pilot study. *BMJ Open.* 2013;3(9):e003138.

103. Mansikkamaki K, Raitanen J, Nygard CH, Tomas E, Rutanen R, Luoto R. Long-term effect of physical activity on health-related quality of life among menopausal women: a 4-year follow-up study to a randomised controlled trial. *BMJ Open.* 2015;5(9):e008232.

104. Mansikkamaki K, Raitanen J, Nygard CH, et al. Sleep quality and aerobic training among menopausal women--a randomized controlled trial. *Maturitas.* 2012;72(4):339-345.

105. Kinser PA, Robins JL. Control group design: enhancing rigor in research of mind-body therapies for depression. *Evid Based Complement Alternat Med.* 2013;2013:140467.

106. McCambridge J, Sorhaindo A, Quirk A, Nanchahal K. Patient preferences and performance bias in a weight loss trial with a usual care arm. *Patient Educ Couns.* 2014;95(2):243-247.

107. Chiu HY, Pan CH, Shyu YK, Han BC, Tsai PS. Effects of acupuncture on menopause-related symptoms and quality of life in women in natural menopause: a meta-analysis of randomized controlled trials. *Menopause.* 2015;22(2):234-244.

108. Lund I, Lundeberg T. Are minimal, superficial or sham acupuncture procedures acceptable as inert placebo controls? *Acupunct Med.* 2006;24(1):13-15.

109. Lesi G, Razzini G, Musti MA, et al. Acupuncture as an integrative approach for the treatment of hot flashes in women with breast cancer: a prospective multicenter randomized controlled trial (AcCliMaT). *J Clin Oncol.* 2016;34(15):1795-1802.

110. Carpenter JS, Woods NF, Otte JL, et al. MsFLASH participants' priorities for alleviating menopausal symptoms. *Climacteric.* 2015:1-8.

APPENDIX A. SEARCH STRATEGY

SYSTEMATIC REVIEWS

PubMed: November 9, 2015

Search	Query	Items found
#1	Search "Menopause"[Mesh] OR menopaus*[tiab] OR "Climacteric"[Mesh:NoExp] OR "Hot Flashes"[Mesh] OR peri-menopaus*[tiab] OR perimenopaus*[tiab] OR postmenopaus*[tiab] OR post-menopaus*[tiab] OR climacteric*[tiab] OR hot-flash*[tiab] OR hot flash*[tiab] OR hot-flush*[tiab] OR hot flush*[tiab] OR night sweat*[tiab] OR vasomotor symptom*[tiab]	**91908**
#2	Search systematic[sb] OR "Systematic Review"[tiab] OR "Umbrella Review"[tiab] OR meta-analysis[tiab] OR "meta analysis"[tiab]	**280172**
#3	Search #1 AND #2	**2771**
#4	Search "Acupuncture Therapy"[Mesh] OR "Acupuncture"[Mesh] OR "Acupressure"[Mesh] OR "acupuncture"[tiab] OR "acupressure"[tiab] OR "electroacupuncture"[tiab]	**23260**
#5	Search #3 AND #4	**45**
#6	Search "Mind-Body Therapies"[Mesh: NoExp] OR "Mind-Body Therapy"[tiab] OR "Mind Body Therapy"[tiab] OR "Mind-Body Therapies"[tiab] OR "Mind Body Therapies"[tiab] OR "Mind Body Medicine"[tiab] OR "Breathing Exercises"[Mesh] OR "Breathing Exercise"[tiab] OR "Breathing Exercises"[tiab] OR "Respiratory Muscle Training"[tiab] OR "Imagery (Psychotherapy)"[Mesh] OR "Guided Imagery"[tiab] OR "Meditation"[Mesh] OR "Meditation"[tiab] OR "Relaxation Therapy"[Mesh] OR "Relaxation Therapy"[tiab] OR "Relaxation Techniques"[tiab] OR "Relaxation Technique"[tiab] OR "Alexander Technique"[tiab] OR "Mind-Body Relations, Metaphysical"[Mesh] OR "Mindfulness"[Mesh] OR "Mindfulness-based Stress Reduction"[tiab] OR "MBSR"[tiab] OR "paced respiration"[tiab] OR "Hypnosis"[Mesh] OR Hypnoses[tiab] OR Hypnotism[tiab] OR Hypnotherapy[tiab] OR Hypnotherapies[tiab] OR Mesmerism[tiab]	**16823**
#7	Search #3 AND #6	**16**
#8	"Yoga"[Mesh] OR "Yoga"[tiab] OR "Tai Ji"[Mesh] OR "Tai Ji"[tiab] OR "Tai-ji"[tiab] OR "Tai Chi"[tiab] OR "T'ai Chi"[tiab] OR "Taiji"[tiab] OR "Taijiquan"[tiab] OR "Qi Gong"[tiab] OR "Qigong"[tiab] OR "Ch'I Kung"[tiab]	**4440**
#9	Search #3 AND #8	**22**
#10	Search "Exercise"[Mesh:NoExp] OR "Exercise"[Majr] OR "Circuit-Based Exercise"[Mesh] OR "Muscle Stretching Exercises"[Mesh] OR "Physical Conditioning, Human"[Mesh] OR "Resistance Training"[Mesh] OR "Resistance Training"[tiab] OR "Running"[Mesh] OR "Jogging"[Mesh] OR "Swimming"[Mesh] OR "Walking"[Mesh] OR "Exercise"[tiab] OR "Exercises"[tiab] OR "physical activity"[tiab] OR "aerobic activity"[tiab] OR "Exercise Movement Techniques"[Mesh] OR "Sports"[Mesh]	**378950**
#11	Search #3 AND #10	**204**
#12	Search #5 OR #7 OR #9 OR #11	**247**
#13	Search (#12) AND ("2009/01/01"[Date – Publication] : "3000"[Date – Publication])	**132**
#14	Search #13 AND "English"[lang]	**123**

EMBASE: November 9, 2015

Search	Query	Items found
#1	'menopause and climacterium'/exp OR 'menopause':ti,ab OR 'menopausal':ti,ab OR 'peri-menopause':ti,ab OR 'peri-menopausal':ti,ab OR 'perimenopause':ti,ab OR 'perimenopausal':ti,ab OR postmenopause:ti,ab OR postmenopausal:ti,ab OR post-menopause:ti,ab OR post-menopausal:ti,ab OR climacteric:ti,ab OR 'hot flash':ti,ab OR 'hot flashes':ti,ab OR 'hot flush':ti,ab OR 'hot flushes':ti,ab OR 'night sweat':ti,ab OR 'night sweats':ti,ab OR 'vasomotor symptoms':ti,ab	**140,141**
#2	'SR'/exp OR 'meta analysis'/exp OR 'Systematic Review':ti,ab OR 'Umbrella Review':ti,ab OR 'meta analysis':ti,ab	**200,849**
#3	#1 AND #2	**2,962**
#4	'acupuncture'/exp OR 'acupuncture':ab,ti OR 'acupressure':ab,ti OR 'electroacupuncture':ab,ti	**37,866**
#5	#3 AND #4	**55**
#6	'alternative medicine'/exp OR 'Mind Body Therapy':ti,ab OR 'Mind Body Therapies':ti,ab OR 'Mind Body Medicine':ti,ab OR 'breathing exercise'/exp OR 'Breathing Exercise':ti,ab OR 'Breathing Exercises':ti,ab OR 'Respiratory Muscle Training':ti,ab OR 'guided imagery'/exp OR 'Guided Imagery':ti,ab OR 'meditation'/exp OR 'Meditation':ti,ab OR 'transcendental meditation'/exp OR 'relaxation training'/exp OR 'relaxation training':ti,ab OR 'Relaxation Therapy':ti,ab OR 'Relaxation Techniques':ti,ab OR 'Relaxation Technique':ti,ab OR ' Alexander Technique':ti,ab OR 'mindfulness'/exp OR 'Mindfulness-based Stress Reduction':ti,ab OR 'MBSR':ti,ab OR 'paced respiration':ti,ab OR 'hypnosis'/exp OR hypnogenesis:ti,ab OR hypnosis:ti,ab OR hypnotism:ti,ab OR mesmerism:ti	**59,899**
#7	#3 AND #6	**65**
#8	'Qi Gong':ti,ab OR 'Qigong':ti,ab OR 'Chi Kung':ti,ab OR 'Tai Chi'/exp OR 'Tai Ji':ti,ab OR 'Tai Chi':ti,ab OR 'Taiji':ti,ab OR 'Taijiquan':ti,ab OR 'Yoga':ti,ab OR 'kinesiotherapy'/exp	**56,974**
#9	#3 AND #8	**60**
#10	'exercise'/exp OR 'sport'/exp OR 'Resistance Training':ti,ab OR 'Exercise':ti,ab OR 'Exercises':ti,ab OR 'physical activity':ti,ab OR 'aerobic activity':ti,ab	**510,515**
#11	#3 AND #10	**249**
#12	#5 OR #7 OR #9 OR #11	**319**
#13	#12 AND [2009-2015]/py	**180**
#14	#13 AND [english]/lim	**173**

Cochrane: November 9, 2015

Search	Query	Items found
#1	menopause OR menopausal OR peri-menopause OR peri-menopausal OR perimenopause OR perimenopausal OR postmenopause OR postmenopausal OR post-menopause OR post-menopausal OR climacteric OR hot flash OR hot flashes OR hot flush OR hot flushes OR night sweat OR night sweats OR vasomotor symptoms:ti,ab,kw (Word variations have been searched)	**15991**
#2	(MH "Acupuncture+") OR (MH "Alternative Therapies+") OR (MH "Mind Body Techniques+") OR (MH "Breathing Exercises+") OR (MH "Qigong") OR (MH "Guided Imagery") OR (MH "Meditation") OR (MH "Hypnosis+") OR (MH "Relaxation Techniques+") OR (MH "Alexander Technique") OR (MH "Tai Chi") OR (MH "Mindfulness") OR (MH "Resistance Training") OR (MH "Muscle Strengthening+") OR (MH "Therapeutic Exercise+") OR (MH "Exercise+") OR (MH "Sports+") OR (MH "Physical Activity") OR TI (acupuncture OR acupressure OR electroacupuncture OR "alternative medicine" OR "Mind Body Therapy" OR "Mind Body Therapies" OR "Mind Body Medicine" OR "Breathing Exercise" OR "Breathing Exercises" OR "Qi Gong" OR Qigong OR "Chi Kung" OR "Respiratory Muscle Training" OR "Guided Imagery" OR Meditation OR "relaxation training" OR "Relaxation Therapy" OR "Relaxation Techniques" OR "Relaxation Technique" OR hypnosis OR "Alexander Technique" OR "Tai Ji" OR "Tai Chi" OR Taiji OR Taijiquan OR Yoga OR "Mindfulness-based Stress Reduction" OR MBSR OR "paced respiration" OR sport OR "Resistance Training" OR Exercise OR Exercises OR "physical activity" OR "aerobic activity") OR AB (acupuncture OR acupressure OR electroacupuncture OR "alternative medicine" OR "Mind Body Therapy" OR "Mind Body Therapies" OR "Mind Body Medicine" OR "Breathing Exercise" OR "Breathing Exercises" OR "Qi Gong" OR Qigong OR "Chi Kung" OR "Respiratory Muscle Training" OR "Guided Imagery" OR Meditation OR "relaxation training" OR "Relaxation Therapy" OR "Relaxation Techniques" OR "Relaxation Technique" OR hypnosis OR "Alexander Technique" OR "Tai Ji" OR "Tai Chi" OR Taiji OR Taijiquan OR Yoga OR "Mindfulness-based Stress Reduction" OR MBSR OR "paced respiration" OR sport OR "Resistance Training" OR Exercise OR Exercises OR "physical activity" OR "aerobic activity") (Word variations have been searched)	**72084**
#3	#1 AND #2 Publication Year from 2009 to 2015	**806**
#4	#3 Cochrane Reviews: 15 Other Reviews: 23	**38**

RANDOMIZED CONTROLLED TRIALS

PubMed: February 12, 2016

Search	Query	Items found
#1	Search "Menopause"[Mesh] OR menopaus*[tiab] OR "Climacteric"[Mesh:NoExp] OR "Hot Flashes"[Mesh] OR peri-menopaus*[tiab] OR perimenopaus*[tiab] OR postmenopaus*[tiab] OR post-menopaus*[tiab] OR climacteric*[tiab] OR hot-flash*[tiab] OR hot flash*[tiab] OR hot-flush*[tiab] OR hot flush*[tiab] OR night sweat*[tiab] OR vasomotor symptom*[tiab]	92983
#2	Search (randomized controlled trial[pt] OR controlled clinical trial[pt] OR randomized[tiab] OR randomised[tiab] OR randomization[tiab] OR randomisation[tiab] OR placebo[tiab] OR drug therapy[sh] OR randomly[tiab] OR trial[tiab] OR groups[tiab]) NOT (animals[mh] NOT humans[mh]) NOT (Editorial[ptyp] OR Letter[ptyp] OR Case Reports[ptyp] OR Comment[ptyp])	2838660
#3	Search #1 AND #2	33136
#4	Search "Acupuncture Therapy"[Mesh] OR "Acupuncture"[Mesh] OR "Acupressure"[Mesh] OR "acupuncture"[tiab] OR "acupressure"[tiab] OR "electroacupuncture"[tiab]	23994
#5	Search #3 AND #4	186
#6	Search "Mind-Body Therapies"[Mesh: NoExp] OR "Mind-Body Therapy"[tiab] OR "Mind Body Therapy"[tiab] OR "Mind-Body Therapies"[tiab] OR "Mind Body Therapies"[tiab] OR "Mind Body Medicine"[tiab] OR "Breathing Exercises"[Mesh] OR "Breathing Exercise"[tiab] OR "Breathing Exercises"[tiab] OR "Respiratory Muscle Training"[tiab] OR "Imagery (Psychotherapy)"[Mesh] OR "Guided Imagery"[tiab] OR "Meditation"[Mesh] OR "Meditation"[tiab] OR "Relaxation Therapy"[Mesh] OR "Relaxation Therapy"[tiab] OR "Relaxation Techniques"[tiab] OR "Relaxation Technique"[tiab] OR "Alexander Technique"[tiab] OR "Mind-Body Relations, Metaphysical"[Mesh] OR "Mindfulness"[Mesh] OR "Mindfulness-based Stress Reduction"[tiab] OR "MBSR"[tiab] OR "paced respiration"[tiab] OR "Hypnosis"[Mesh] OR Hypnoses[tiab] OR Hypnotism[tiab] OR Hypnotherapy[tiab] OR Hypnotherapies[tiab] OR Mesmerism[tiab]	28100
#7	Search #3 AND #6	79
#8	Search "Yoga"[Mesh] OR "Yoga"[tiab] OR "Tai Ji"[Mesh] OR "Tai Ji"[tiab] OR "Tai-ji"[tiab] OR "Tai Chi"[tiab] OR "T'ai Chi"[tiab] OR "Taiji"[tiab] OR "Taijiquan"[tiab] OR "Qi Gong"[tiab] OR "Qigong"[tiab] OR "Ch'I Kung"[tiab]	4576
#9	Search #3 AND #8	84
#10	Search "Exercise"[Mesh:NoExp] OR "Exercise"[Majr] OR "Circuit-Based Exercise"[Mesh] OR "Muscle Stretching Exercises"[Mesh] OR "Physical Conditioning, Human"[Mesh] OR "Resistance Training"[Mesh] OR "Resistance Training"[tiab] OR "Running"[Mesh] OR "Jogging"[Mesh] OR "Swimming"[Mesh] OR "Walking"[Mesh] OR "Exercise"[tiab] OR "Exercises"[tiab] OR "physical activity"[tiab] OR "aerobic activity"[tiab] OR "Exercise Movement Techniques"[Mesh] OR "Sports"[Mesh]	387613
#11	Search #3 AND #10	2129
#12	Search #5 OR #7 OR #9 OR #11	2338
#13	Search #12 AND ("2012/01/01"[Date – Publication] : "3000"[Date – Publication])	653
#14	Search #13 AND "English"[lang] Sort by: Author	612

EMBASE: February 12, 2016

Search	Query	Items found
#1	'menopause and climacterium'/exp OR 'menopause':ti,ab OR 'menopausal':ti,ab OR 'peri-menopause':ti,ab OR 'peri-menopausal':ti,ab OR 'perimenopause':ti,ab OR 'perimenopausal':ti,ab OR postmenopause:ti,ab OR postmenopausal:ti,ab OR post-menopause:ti,ab OR post-menopausal:ti,ab OR climacteric:ti,ab OR 'hot flash':ti,ab OR 'hot flashes':ti,ab OR 'hot flush':ti,ab OR 'hot flushes':ti,ab OR 'night sweat':ti,ab OR 'night sweats':ti,ab OR 'vasomotor symptoms':ti,ab	142,022
#2	(('randomized controlled trial'/exp OR 'crossover procedure'/exp OR 'double blind procedure'/exp OR 'single blind procedure'/exp OR random* OR factorial* OR crossover* OR cross NEAR/1 over* OR placebo* OR doubl* NEAR/1 blind* OR singl* NEAR/1 blind* OR assign* OR allocat* OR volunteer*) NOT ('case report'/exp OR 'case study'/exp OR 'editorial'/exp OR 'letter'/exp OR 'note'/exp OR 'conference paper'/exp)) AND ([humans]/lim AND [english]/lim AND [2012-2016]/py)	428,259
#3	#1 AND #2	6,358
#4	'acupuncture'/exp OR 'acupuncture':ab,ti OR 'acupressure':ab,ti OR 'electroacupuncture':ab,ti OR 'alternative medicine'/exp OR 'Mind Body Therapy':ti,ab OR 'Mind Body Therapies':ti,ab OR 'Mind Body Medicine':ti,ab OR 'breathing exercise'/exp OR 'Breathing Exercise':ti,ab OR 'Breathing Exercises':ti,ab OR 'Respiratory Muscle Training':ti,ab OR 'guided imagery'/exp OR 'Guided Imagery':ti,ab OR 'meditation'/exp OR 'Meditation':ti,ab OR 'transcendental meditation'/exp OR 'relaxation training'/exp OR 'relaxation training':ti,ab OR 'Relaxation Therapy':ti,ab OR 'Relaxation Techniques':ti,ab OR 'Relaxation Technique':ti,ab OR ' Alexander Technique':ti,ab OR 'mindfulness'/exp OR 'Mindfulness-based Stress Reduction':ti,ab OR 'MBSR':ti,ab OR 'paced respiration':ti,ab OR 'hypnosis'/exp OR hypnogenesis:ti,ab OR hypnosis:ti,ab OR hypnotism:ti,ab OR mesmerism:ti,ab OR 'Qi Gong':ti,ab OR 'Qigong':ti,ab OR 'Chi Kung':ti,ab OR 'Tai Chi'/exp OR 'Tai Ji':ti,ab OR 'Tai Chi':ti,ab OR 'Taiji':ti,ab OR 'Taijiquan':ti,ab OR 'Yoga':ti,ab OR 'kinesiotherapy'/exp OR 'exercise'/exp OR 'sport'/exp OR 'Resistance Training':ti,ab OR 'Exercise':ti,ab OR 'Exercises':ti,ab OR 'physical activity':ti,ab OR 'aerobic activity':ti,ab	630,791
#4	#3 AND #4	849

CINAHL: February 12, 2016

Search	Query	Items found
S1	(MH "Menopause+") OR TI (menopause OR menopausal OR "peri-menopause" OR "peri-menopausal" OR perimenopause OR perimenopausal OR postmenopause OR postmenopausal OR "post-menopause" OR "post-menopausal" OR climacteric OR "hot flash" OR "hot flashes" OR "hot flush" OR "hot flushes" OR "night sweat" OR "night sweats" OR "vasomotor symptoms") OR AB (menopause OR menopausal OR "peri-menopause" OR "peri-menopausal" OR perimenopause OR perimenopausal OR postmenopause OR postmenopausal OR "post-menopause" OR "post-menopausal" OR climacteric OR "hot flash" OR "hot flashes" OR "hot flush" OR "hot flushes" OR "night sweat" OR "night sweats" OR "vasomotor symptoms")	22,202
S2	(MH "Acupuncture+") OR (MH "Alternative Therapies+") OR (MH "Mind Body Techniques+") OR (MH "Breathing Exercises+") OR (MH "Qigong") OR (MH "Guided Imagery") OR (MH "Meditation") OR (MH "Hypnosis+") OR (MH "Relaxation Techniques+") OR (MH "Alexander Technique") OR (MH "Tai Chi") OR (MH "Mindfulness") OR (MH "Resistance Training") OR (MH "Muscle Strengthening+") OR (MH "Therapeutic Exercise+") OR (MH "Exercise+") OR (MH "Sports+") OR (MH	341,974

Search	Query	Items found
	"Physical Activity") OR TI (acupuncture OR acupressure OR electroacupuncture OR "alternative medicine" OR "Mind Body Therapy" OR "Mind Body Therapies" OR "Mind Body Medicine" OR "Breathing Exercise" OR "Breathing Exercises" OR "Qi Gong" OR Qigong OR "Chi Kung" OR "Respiratory Muscle Training" OR "Guided Imagery" OR Meditation OR "relaxation training" OR "Relaxation Therapy" OR "Relaxation Techniques" OR "Relaxation Technique" OR hypnosis OR "Alexander Technique" OR "Tai Ji" OR "Tai Chi" OR Taiji OR Taijiquan OR Yoga OR "Mindfulness-based Stress Reduction" OR MBSR OR "paced respiration" OR sport OR "Resistance Training" OR Exercise OR Exercises OR "physical activity" OR "aerobic activity") OR AB (acupuncture OR acupressure OR electroacupuncture OR "alternative medicine" OR "Mind Body Therapy" OR "Mind Body Therapies" OR "Mind Body Medicine" OR "Breathing Exercise" OR "Breathing Exercises" OR "Qi Gong" OR Qigong OR "Chi Kung" OR "Respiratory Muscle Training" OR "Guided Imagery" OR Meditation OR "relaxation training" OR "Relaxation Therapy" OR "Relaxation Techniques" OR "Relaxation Technique" OR hypnosis OR "Alexander Technique" OR "Tai Ji" OR "Tai Chi" OR Taiji OR Taijiquan OR Yoga OR "Mindfulness-based Stress Reduction" OR MBSR OR "paced respiration" OR sport OR "Resistance Training" OR Exercise OR Exercises OR "physical activity" OR "aerobic activity")	
S3	S1 AND S2 Published Date: 20120101-20160231	**1,029**
S4	TI ("randomized controlled trial" OR "controlled clinical trial" OR "randomized" OR "randomized" OR "randomization" OR "randomization" OR "placebo" OR "randomly" OR "trial" OR "groups" OR AB ("randomized controlled trial" OR "controlled clinical trial" OR "randomized" OR "randomized" OR "randomization" OR "randomization" OR "placebo" OR "randomly" OR "trial" OR "groups") OR (MH "Randomized Controlled Trials")	**391,791**
S5	S3 AND S4	**345**
S6	S5 NOT PT (Book OR Book Chapter OR Book Review OR Case Study OR Commentary OR Doctoral Dissertation OR Editorial OR Letter OR Masters Thesis OR Pamphlet OR Pamphlet Chapter OR Poetry)	**339**

APPENDIX B. CRITERIA USED IN QUALITY ASSESSMENT

SYSTEMATIC REVIEWS

General instructions: The purpose of this rating tool is to evaluate the scientific quality of SRs. It is not intended to measure the literary quality, importance, relevance, originality, or other attributes of SRs.

Step 1: First determine whether it is a SR (SR). Systematic reviews are studies that: 1) include an explicit and adequate search, 2) apply prespecified eligibility criteria, and 3) consider quality of included studies or risk of bias assessment, and/or describe plans to synthesize or attempt to synthesize findings quantitatively and/or qualitatively.

Step 2: For SRs, grade each of the criteria listed below as "Yes," "No," "Can't tell," or "Not Applicable." Factors to consider when making an assessment are listed under each criterion. For each domain, summarize key methods, level of concern overall (low, high, unclear), and rationale for concerns.

STUDY ELIGIBILITY CRITERIA

1. **Was an 'a priori' design provided?**
   The research question and inclusion criteria should be established before the conduct of the review and the review should adhere to pre-defined objectives and eligibility criteria.

   Note: *Need to refer to a protocol, ethics approval, or pre-determined/a priori published research objectives to score a "yes."*

   [] Yes [] No [] Can't tell [] N/A
   Comment: No published protocol; however, their approach is reasonable

2. **Were study eligibility criteria clearly specified?**
   Note: Criteria should be sufficiently detailed to allow replication of study
   [] Yes [] No [] Can't tell [] N/A
   Comment:

3. **Were restrictions in eligibility criteria appropriate?** Restrictions based on study characteristics (*eg,* date, sample size, study quality, outcomes measured) and based on sources of information (*eg,* publication status or format, language) should be appropriate?
   [] Yes [] No [] Can't tell [] N/A
   Comment: No language or country restrictions given

Summarize key methods related to Eligibility:

Summarize concerns (low/high/unclear) and your rationale regarding specification of study eligibility criteria:

IDENTIFICATION AND SELECTION OF STUDIES

4. **Was a comprehensive literature search performed?**
   At least 2 electronic sources should be searched and electronic searches should be supplemented by consulting: reference lists from prior reviews, textbooks, or included studies; specialized registries (*eg*, Cochrane registries); or queries to experts in the field.

   Note: If at least 2 sources + one supplementary strategy used, select "yes"; grey literature search counts as supplementary

[] Yes [] No [] Can't tell [] N/A
Comment:

5. Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?

Note: Search methods should describe: search date, databases used, and search terms (Key words and/or MESH terms must be stated and where feasible the search strategy should be provided).

[] Yes [] No [] Can't tell [] N/A
Comment:

6. Were restrictions based on date, publication format, or language appropriate?
[] Yes [] No [] Can't tell [] N/A
Comment:

7. Was selection bias avoided?
The review reports the number of studies identified through searches, the numbers excluded, and gives appropriate reasons for excluding – based on explicit inclusion/exclusion criteria. Two or more raters made inclusion/exclusion decisions.

[] Yes [] No [] Can't tell [] N/A
Comment:

Summarize key methods related to Study Selection:

Summarize concerns (low, high, unclear) and rationale regarding Study Selection

DATA COLLECTION AND STUDY APPRAISAL

8. Was there duplicate study selection and data extraction?
Did two or more investigators abstract data, and assess study quality – either independently or with one rater over-reading the first raters result?
Was an appropriate method used to resolve disagreements (*eg,* a consensus procedure)?
[] Yes [] No [] Can't tell [] N/A
Comment: last paragraph, page 136.

9. Were the characteristics of the included studies provided?
In an aggregated form such as a table, data from the original studies should be provided on the participants, interventions and outcomes. The ranges of characteristics in all the studies analyzed (*eg,* age, race, sex, relevant socioeconomic data, disease status, duration, severity or other diseases) should be reported in sufficient detail to allow the review authors and readers to interpret the results.

Note: Acceptable if not in table format as long as they are described as above.

[] Yes [] No [] Can't tell [] N/A
Comment:

10. Was the scientific quality of the included studies assessed and documented?
A priori methods of assessment should be provided and criteria used to assess study quality specified in enough detail to permit replication.

Note: Can include use of a quality scoring tool or checklist, eg, Jadad scale, risk of bias, sensitivity analysis, etc., or a description of quality items, with some kind of result for EACH study ("low" or

"high" is fine, as long as it is clear which studies scored "low" and which scored "high"; a summary score/range for all studies is not acceptable).

[] Yes [] No [] Can't tell [] N/A
Comment:

Summarize key methods related to data collection/ROB:
Summarize concerns regarding data collection/ROB:

SYNTHESIS AND FINDINGS

11. Was the scientific quality of the included studies used appropriately in formulating conclusions?
The results of the methodological rigor and scientific quality should be considered in the analysis (*eg*, subgroup analyses) and the conclusions of the review, and explicitly stated in formulating recommendations.

Note: Might say something such as "the results should be interpreted with caution due to poor quality of included studies." Cannot score "yes" for this question if scored "no" for question 10.

[] Yes [] No [] Can't tell [] N/A
Comment:

12. Were the methods used to combine the findings of studies appropriate?
For pooled results, the synthesis should be appropriate given the nature and similarity of included studies (conceptual homogeneity), and an accepted quantitative method of pooling should be used (*ie*, more than simple addition; *eg,* random-effects or fixed-effect model). If only qualitative analyses are completed, the study should describe the reasons that quantitative analyses were not completed.
[] Yes [] No [] Can't tell [] N/A@
Comment:.

13. Was between-study variation (heterogeneity) minimal or addressed in the synthesis?
For pooled results, a qualitative and quantitative assessment of homogeneity (Cochran's Q and/or I^2) should be performed.

Note: Indicate "yes" if they explain that they cannot pool because of heterogeneity/variability between interventions.

[] Yes [] No [] Can't tell [] N/A@
Comment:

14. Was the likelihood of publication bias assessed?
Publication bias tested using funnel plots, test statistics (*eg,* Egger's regression test), and/or search of trials registry for unpublished studies.

Note: If no test values, clinical trials.gov search, or funnel plot included, score "no". Score "yes" if mentions that publication bias could not be assessed because there were fewer than 10 included studies.

[] Yes [] No [] Can't tell [] N/A
Comment:

15. Are the stated conclusions supported by the data presented?
Were the conclusions made by the author(s) supported by the data and/or analyses reported in the SR? Conclusions should address limitations of the SR and limitations of the primary studies.

Conclusions should consider relevance of the included studies to the research question.
[] Yes [] No [] Can't tell [] N/A

Summarize key methods related to synthesis:
Summarize concerns regarding synthesis:

OTHER

16. Was the conflict of interest stated?
Potential sources of support should be clearly acknowledged in both the SR and the included studies.
Note: To get a "yes," must indicate source of funding or support for the SR AND for each of the included studies
[] Yes [] No [] Can't tell [] N/A
Comment: Competing interest for SR authors given (none) but not for primary studies

Step 3: Rate the overall quality of the SR as "Good," "Fair," or "Poor" using the guidance below and summarize major reasons for rating in the comments box.

Good = After considering items 1-15, item 15 is rated "Yes" with no important limitations. This means that few of the items 1-14 are rated "No," and none of the limitations are thought to decrease the validity of the conclusions. If items 3, 4, 7, 9, 10, 11 or 12 are rated "no", then the review is likely to have major flaws

Fair = After considering items 1-15, item 15 is rated "Yes," but with at least some important limitations. This means that enough of the items 1-15 are rated "No" to introduce some uncertainty about the validity of the conclusions.

Poor = After considering items 1-15, item 15 is rated "No." This means that several of items 1-15 are rated "No," introducing serious uncertainty about the validity of the conclusions.

Overall rating comments:

Concerns regarding specifications of eligibility criteria

Low concern	Considerable effort has been made to clearly specify the review question and objectives, and to pre-specify and justify appropriate and detailed eligibility criteria that have been adhered to during the review
High concern	Studies that would have been important and relevant to answering the review question are likely to have been excluded from the review, either due to the lack of pre-specified objectives and eligibility criteria, or because inappropriate restrictions were imposed or studies that are not appropriate for addressing the review question have been included.
Unclear concern	Insufficient information is reported to make a judgement about risk of bias.

Concerns regarding methods used to identify and/or select studies

Low concern	Given the review question and eligibility criteria as assessed in Domain 1, a substantial effort has been made to identify as many relevant studies as possible through a variety of search methods using a sensitive and appropriate search strategy and steps were taken to minimise bias and errors when selecting studies for inclusion.
High concern	Some eligible studies are likely to be missing from the review.
Unclear concern	There is insufficient information reported to make a judgement on risk of bias.

Concerns regarding methods used to collect data and appraise studies

Low concern	Given the studies included in the review as assessed in domain 2, risk of bias was assessed using appropriate criteria, data extraction and risk of bias assessment involved two reviewers, and relevant study characteristics and results were extracted
High concern	Some bias may have been introduced through the data collection or risk of bias assessment processes.
Unclear concerns	There is insufficient information reported to inform a judgement on risk of bias.

Concerns regarding methods used to synthesize results

Low concern	The synthesis is unlikely to produce biased results, because any limitations in the data were overcome, or the findings were so convincing that the limitations would have little impact
High concern	The synthesis is likely to produce biased results, because (i) potential biases were ignored (within and/or across studies), (ii) important between-study variation was not accounted for; (iii) there were important inadequacies in the methodology; or (iv) findings are incompletely reported in a way that raises concerns.
Unclear concerns	There is insufficient information reported to make a judgement on risk of bias.

References

Marinopoulos SS, Dorman T, Ratanawongsa N, et al. Effectiveness of continuing medical education. Evid Rep Technol Assess (Full Rep). 2007(149):1-69.

Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. Lancet. 1999;354(9193):1896-1900.

Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of SRs. BMC Medical Research Methodology. 2007;7(1):1-7.

Whiting P, Savovic J, Higgins JP, et al. ROBIS: A new tool to assess risk of bias in SRs was developed. J Clin Epidemiol. 2015.

RANDOMIZED CONTROLLED TRIALS

General instructions: Rate each risk of bias item listed below as “Low,” “High,” or “Unclear.”

Rating of individual items:
1.Selection bias:

Domain: Random sequence generation

(***Support for judgement:*** *Describe the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups.)*

Was the allocation sequence adequately generated?

Low risk High risk Unclear risk

Comment

Domain: Allocation concealment?

***(Support for judgement**: Describe the method used to conceal the allocation sequence in sufficient detail to determine whether intervention allocations could have been foreseen in advance of, or during, enrolment)*

Was allocation adequately concealed?

Low risk High risk Unclear risk

Comment

2. Performance bias

Domain: Blinding of participants and "treating" personnel (ie, the person(s) delivering the intervention).

***(Support for judgement:** Describe all measures used, if any, to blind study participants and personnel from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.)*

Was knowledge of the allocated intervention adequately prevented during the study?

Low risk High risk Unclear risk Outcome NR

Comment

3a. Detection bias (VMS symptoms):

Domain: Blinding of outcome assessment

***(Support for judgement:** Describe all measures used, if any, to blind outcome assessors from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.)*

Was knowledge of the allocated intervention adequately prevented from outcome assessors?

Low risk High risk Unclear risk Outcome NR

Comment

3b. Detection bias (outcomes measured by self-report, *ie*, depression, anxiety, sleep, QoL, *etc*):

Domain: Blinding of outcome assessment

***(Support for judgement:** Describe all measures used, if any, to blind outcome assessors from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.)*

Was knowledge of the allocated intervention adequately prevented from outcome assessors?

Low risk High risk Unclear risk Outcome NR

Comment

3c. Detection bias (Adverse effects):

Domain: Blinding of outcome assessment

***(Support for judgement:** Describe all measures used, if any, to blind outcome assessors from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.)*

Was knowledge of the allocated intervention adequately prevented from outcome assessors?

Low risk High risk Unclear risk Outcome NR

Comment

4. Attrition bias:

Domain: Incomplete outcome data

***(Support for judgement:** Describe the completeness of outcome data for each main outcome, including attrition and exclusions from the analysis. State whether attrition and exclusions were reported, the numbers in each intervention group (compared with total randomized participants), reasons for attrition/exclusions where reported, and any re-inclusions in analyses performed by the review authors.)*

Were incomplete outcome data adequately addressed?

Low risk High risk Unclear risk

Comment

5. Reporting bias:

Domain: Selective outcomes reporting

*(**Support for judgement:** State how the possibility of selective outcome reporting was examined by the review authors, and what was found.)*

Are reports of the study free of suggestion of selective outcome reporting? (i.e., the author states they will measure an outcome but do not report it)

Low risk High risk Unclear risk

Comment

6. Other

Domain: Other sources of bias

***(Support for judgement:** State any important concerns about bias not addressed in the other domains in the tool.*
If particular questions/entries were pre-specified in the review's protocol, responses should be provided for each question/entry.)

Are reports of the study free from other bias due to problems not covered above?

Low risk High risk Unclear risk

Comment

Overall risk of bias rating

Low Unclear High

Narrative:

Risk of Bias	Interpretation	Criteria
Low risk of bias	Bias, if present, is unlikely to alter the results seriously.	Adequacy of random sequence generation, allocation concealment, and blinding scored as “low risk of bias” and no important concerns related to the other domains.
Unclear risk of bias	A risk of bias that raises some doubts about the results	One or two domains are scored “not clear” or not done.
High risk of bias	Bias may alter the results seriously	More than 2 domains are scored as “not clear” or not done

* Items contained in Cochrane Risk of Bias Tool

APPENDIX C. QUALITY ASSESSMENT RATINGS

References cited in this appendix appear in the reference list of the main report.

SYSTEMATIC REVIEWS

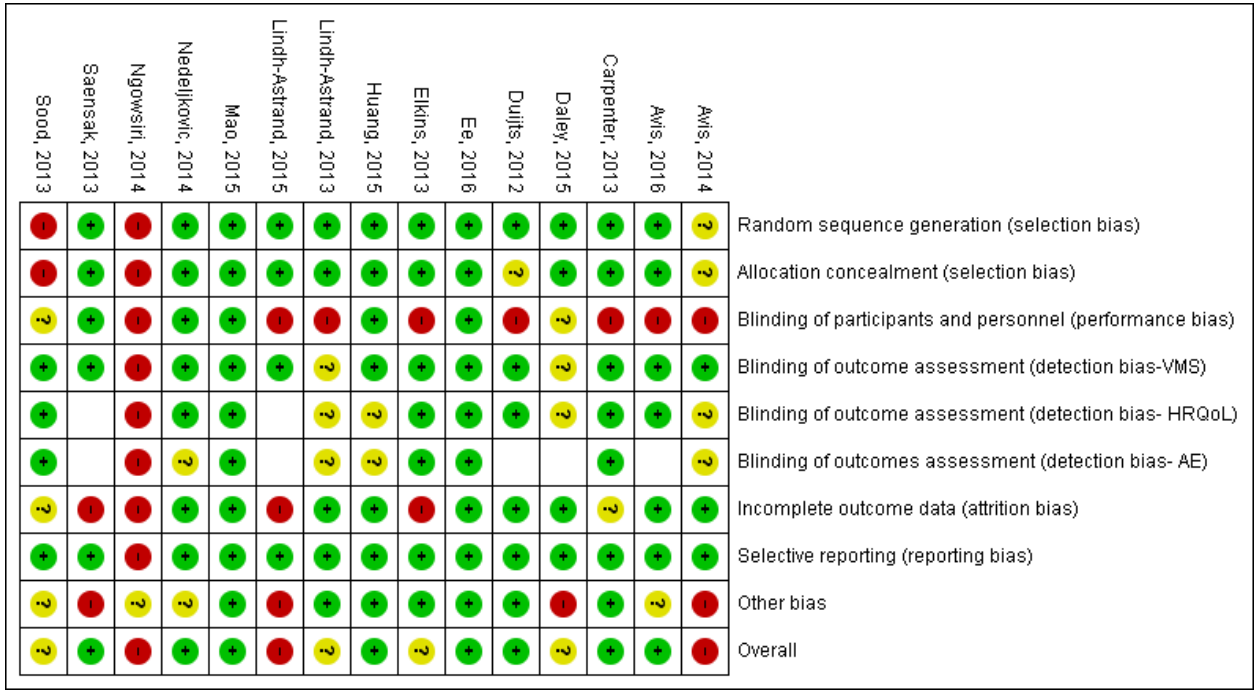
Publication	Eligibility Criteria	ID/Select Studies	Data Collection	Synthesis	Overall
Acupuncture					
Dodin 2013[40]	Low	Low	Low	Low	**Good**
Garcia 2015[41]	Low	High	Low	Low	**Fair**
Garcia 2013[64]	Low	High	Low	Low	**Fair**
Yoga					
Cramer 2012[43]	Low	Low	Low	Low	**Good**
Exercise					
Daley 2014[44]	Low	Low	Low	Low	**Good**
Meditation/Mindfulness/ Hypnosis/ Relaxation/Mixed					
Saensak 2014[45]	Low	Low	Low	Low	**Good**
Woods 2014[46]	Low	Low	High	High	**Fair**
Cramer 2015[49]	Low	Low	Low	High	**Fair**
Innes 2010[47]	Low	Low	Low	High	**Fair**
Rada 2010[48]	Low	Low	Low	Low	**Good**

RANDOMIZED CONTROLLED TRIALS

Publication	1	2	3	4	5	6	7	8	9	Overall
Acupuncture										
Avis 2016[52]	LR	LR	HR	LR	LR	Outcome NR	LR	LR	UR	**Low**
Ee 2016[50]	LR	LR	LR	LR	LR	LR	LR	LR	LR	**Low**
Mao 2015[51]	LR	LR	LR	LR	LR	LR	LR	LR	LR	**Low**
Nedeljkovic 2014[53]	LR	LR	LR	LR	LR	UR	LR	LR	UR	**Low**
Yoga										
Avis 2014 [54]	UR	UR	HR	LR	UR	UR	LR	LR	HR	**High**
Ngowsiri 2014[55]	HR	HR	HR	HR	HR	HR	HR	HR	UR	**High**
Exercise										
Daley 2015[62]	LR	LR	UR	UR	UR	Outcome NR	LR	LR	HR	**Unclear**
Duijts 2012[63]	LR	UR	HR	LR	LR	Outcome NR	LR	LR	LR	**Low**
Meditation/Mindfulness/ Hypnosis/Relaxation/Mixed										
Carpenter 2013[56]	LR	LR	HR	LR	LR	LR	UR	LR	LR	**Low**
Huang 2015[57]	LR	LR	LR	LR	UR	UR	LR	LR	LR	**Low**
Lindh-Astrand 2013[58]	LR	LR	HR	UR	UR	UR	LR	LR	LR	**Unclear**
Lindh-Astrand 2015[61]	LR	LR	HR	LR	Outcome NR	Outcome NR	HR	LR	HR	**High**
Saensak 2013[59]	LR	LR	LR	LR	Outcome NR	Outcome NR	LR	LR	HR	**Low**
Sood 2013[60]	HR	HR	UR	LR	LR	LR	UR	LR	UR	**Unclear**

Risk of bias abbreviations: HR = High risk; LR = Low risk; UR = Unclear risk
Column headings:
1. Randomization adequate
2. Allocation concealment
3. Performance bias
4. Detection bias: VMS
5. Detection bias: Health-related quality of life
6. Detection bias: Adverse effects
7. Incomplete outcome
8. Selective outcomes reporting
9. Other bias

Risk of Bias Summary for New RCTs



APPENDIX D. MEASUREMENT SCALES FOR MENOPAUSE

VASOMOTOR SYMPTOMS (VMS)

- Daily Hot Flash Diary (Sloan et al, 2001). A measure of self-reported hot flashes, night sweats and/or severity of hot flashes typically using a dairy to record the frequency and severity of hot flashes using a 4-point scale: mild, moderate, severe, very severe to provide a hot flash index (sum of the number of hot flashes multiplied by severity).

- Hot Flush Rating Scale (Hunter et al, 1995). Yields a hot flush frequency score (2 items that each have times/day and days/week summed to give number of hot flushes and night sweats that have caused waking in the past week) and a hot flush problem rating (3 items, each scored 1 to 10).

- Vasomotor subscales. From instruments such as the Green Climacteric Scale, the Kupperman Menopausal Index, the Women's Health Questionnaire (WHQ), and the Menopause-specific Quality of Life (MENQOL), Health-related Quality of Life (HRQOL) Specific to Perimenopausal Symptoms (*ie*, measures used to report menopause-related HRQOL).

- Greene Climacteric Scale. 21 questions covering 5 domains: anxiety, depression, somatic symptoms, vasomotor symptoms, and sexual function. Each question is answered on a 4-point Likert scale. The answers to all 21 questions are summed to give a total quality-of-life measure; a higher score indicates a worse quality of life.

- Global quality of life. Not a specific measure but a term used for single-item visual analogue scales designed to measure overall quality of life when 0 is the lowest possible and 100 is the highest possible quality of life.

- Hot Flush Behavior Scale (or Hunter menopause scale). A measure of behavioral reactions to VMS.

- Hot Flash Related Daily Interference Scale. A 10-item measure for assessing the impact of vasomotor symptoms on daily activities in nine specific domains within the past week (work, social activities, leisure activities, sleep, mood, concentration, relation with others, sexuality, and enjoyment of life) and overall quality of life.

- Kupperman Menopausal Index. A numerical index that scores 11 menopausal symptoms: hot flushes, paresthesia, insomnia, nervousness, melancholia, vertigo, weakness, arthralgia or myalgia, headache, palpitations, and formication. Each symptom is rated from 0 to 3 according to severity. The scores are weighted and a total sum is calculated with a higher score indicating a worse quality of life.

- Menopause-specific quality of life (MENQOL) (Lewis et al, 2005). 29 questions covering 4 domains: vasomotor, psychosocial, physical, and sexual. The scoring for each question is 1–"No", 2–"Yes, but not at all bothered" through 8–"Yes, extremely

bothered." The scores for each question are summed for a total quality-of-life score where a higher score indicates a worse quality of life.

- Menopause Rating Scale. Scores 11 menopausal symptoms: hot flushes, heart discomfort, sleep problems, depressive mood, irritability, anxiety, physical and mental exhaustion, sexual problems, bladder problems, vaginal dryness, and joint and muscular discomfort. Each item is scored from 0-4. The scores are summed for a total quality-of-life score, in which a higher score indicates a worse quality of life.

- Women's Health Questionnaire (WHQ). Contains menopause-specific quality of life subscales enabling a detailed assessment of dimensions of emotional and physical health, such as depression, anxiety, sleep problems, somatic symptoms, with optional subscales for menstrual problems and sexual difficulties.

- Utian Quality of Life Scale (UQOL). 23 items scored 1 to 5; 4 subscales (occupational, health, emotional, sexual) and total score.

GENERAL HEALTH-RELATED QUALITY OF LIFE (HRQOL)

- SF-36 or Rand-36. Consists of 36 questions covering 8 domains: physical functioning, role limitations caused by physical health problems, role limitations caused by emotional problems, social functioning, emotional well-being, energy/fatigue, pain, and general health perceptions. This questionnaire produces outcomes on total quality of life, sub-scores for each of the domains, a physical health sub-score, or a mental health sub-score. For this scale, the higher the score, the better the quality of life.

- Functional Assessment of Cancer Therapy-Endocrine Subscale (FACT-ES). An 18-item instrument designed to measure the side effects and benefits of hormonal treatments used in women with breast cancer.

- EQ-5D-3L. Evaluates function in 5 domains: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has 3 levels: no problems, some problems, extreme problems.

- EQ-5D-5L. Like the EQ-5D-3L but each dimension has 5 levels: no problems, slight problems, moderate problems, severe problems, and extreme problems.

- EQ VAS. Records the respondent's self-rated health on a 20 cm vertical, visual analogue scale with endpoints labelled "the best health you can imagine" and "the worst health you can imagine."

- Functional Assessment of Cancer Therapy-Breast (FACT-B). Measures overall quality of life and has 5 subscales including physical, social, emotional, functional well-being, and breast concerns.

APPENDIX E. PEER REVIEW COMMENTS/AUTHOR RESPONSES

Reviewer	Comment	Response
Question 1. Are the objectives, scope, and methods for this review clearly described?		
1	Yes	Acknowledged
2	Yes	Acknowledged
3	**No**: I had difficulty with statements like "judged to be of high quality" (p 7, line 48) and "poor quality" (p 7, line 57) that came before the detailed Methods explanation and discussion of how "quality" was judged. This may just be the ESP format but I found it hard to read statements like these w/o knowing more. This is a minor point, but may be best to avoid the statement HT "only for short-term use" as this doesn't encompass newer recommendations to individualize care. These alternative/integrative methods are really a move toward individualization or personalization of health.	We added a statement in the abstract that briefly defines good- and fair-quality reviews. We clarified the statement about current recommendations for duration of HT use to encompass guidance around individualization of care as recommended.
4	Yes	Acknowledged
Question 2. Are there any <u>published</u> or <u>unpublished</u> studies that we may have overlooked?		
1	No	Acknowledged
2	No	Acknowledged
3	No	Acknowledged
4	No	Acknowledged
Question 3. Is there any indication of bias in our synthesis of the evidence?		
1	No	Acknowledged
2	No	Acknowledged
3	No	Acknowledged
4	**Yes**: Lesi et al. J. Clin Onc, 2016 studied acupuncture vs. usual care (included 190 women with breast cancer)	This study was published after our search date and including it in the analysis without completing an updated search for other recently published studies could introduce bias. However, we added a citation to this study in the discussion and note that results were consistent with our review findings.

Reviewer	Comment	Response
Question 4: Please write additional suggestions or comments below. If applicable, please indicate the page and line numbers from the draft report.		
1	This is an incredibly extensive and thorough review. It would be helpful to indicate if at all possible if the results from the meta-analysis that were statistically significant translated to clinically meaningful changes in the outcomes (e.g. frequency of VMS). Some of these look quite small and it is difficult to determine whether these changes would make a meaningful difference to women struggling with VMS.	Thank you for this comment. For daily hot flash frequency, which was measured in a consistent manner, we were able to make this transformation. For hot flash severity and health-related quality-of-life, there was too much variability in measurement to make a reliable transformation back to natural units. Although we wanted to add information on the minimum clinically important difference for hot flash frequency, there is no consensus on this value.
2	Very thorough review of existing SR's and RCTs on topics related to nonpharmacological therapies for VMS. This review was well-written, clear, and described each area very well. Two suggestions: 1. Please add information on the inclusion/exclusion criteria for the SR's and RCTs that were reviewed for this synthesis. 2. Please provide information on the strength of evidence guidelines that were used by the authors--there is a reference to AHRQ's Methods, however, even a short description of what good, fair, etc... entails would be helpful for the reader. Otherwise, well done, comprehensive review.	Eligibility criteria and definitions of the strength of evidence are presented in detail in the Methods section
3	See above, well written, difficult topic to take on. This provides important contributions to Women's Health.	Thank you.

Reviewer	Comment	Response
4	I had the pleasure reviewing the Non-pharmacological treatment of menopause associated vasomotor symptoms prepared for the Department of Veterans Affairs (Lead author: John Williams). This review is thorough, well-written, with appropriately stated conclusions. I think this will have important implications for improving VM symptom management. A few aspects should be addressed to enhance the quality of this manuscript	Thank you
	One new paper just published is highly relevant to this review. Lesi et al. J. Clin Onc, 2016 studied acupuncture vs. usual care (included 190 women with breast cancer)	See response to this comment above.
	I am not sure whether there would be enough trials to do a meaningful comparison for acupuncture against non-hormonal based drugs (Gabapentin, Venlafaxine) for effect size, adverse events, and treatment durability. I am aware of at least two such studies. This may be a particularly useful comparison given that for those women who do not want hormones, these are viable options. The authors advocated for comparative effectiveness trials so these analyses (even a qualitative review) may help the authors to be more explicit on what these comparisons should focus on, main effects on hot flashes, adverse effects, durability.	We considered but did not think it appropriate to compare acupuncture with any drug therapy (i.e., drugs from differing classes with differing mechanisms of effect). Our *a priori* threshold for meta-analysis was 3 trials and comparisons of acupuncture to differing drug classes, and so this did not meet that threshold.
	Another aspect is the durability of treatment effect. Currently, this is buried in the text. For women, they want to know if they choose a therapy (acupuncture, yoga, or drugs), if the reduction of their symptoms are durable or they need to get continued treatments. It will be great to summarize these findings as it will have impact on patient quality of life and health care utilization (getting a defined course of acupuncture/drugs will have very different cost implications for VA than getting continued treatment for a long period of time	Thank you for this comment. We have revised the text to more clearly identify timing of outcome assessments and whether the durability of treatment effect was addressed.

Reviewer	Comment	Response
4 continued	Other minor edits Introduction: P11line10: change “2” to “two”	Per VA report editorial style, all numbers below 10 (except “one”) are to use numerals.
	P11line 22: please add “severity”	This line of the report is in the Data Analysis section of the Methods, which is not concerned with frequency or severity of VMS. Perhaps the reviewer meant a different line. We have carefully reviewed the manuscript to ensure that frequency and severity are used properly and are present where necessary.
	In third paragraph of the introduction, please discuss that hormonal treatment is contraindicated for breast cancer survivors or those with high risk for breast cancer	Text was added to note that HT is contraindicated in women with a history of breast cancer or those at high risk for breast cancer.
	Page 45: “though it found no difference due to hypnosis when compared to treatment with gabapentin”. I would assume gabapentin is an active control since it has been shown to be more beneficial than placebo for hot flashes	This sentence was adjusted to clarify that gabapentin is an active comparator.