CrossMark

META-ANALYSIS

Efficacy of menopausal hormone therapy on sleep quality: systematic review and meta-analysis

Dahima Cintron [1] · Melissa Lipford [2] · Laura Larrea-Mantilla [3] ·
Gabriela Spencer-Bonilla [1,3] · Robin Lloyd [2] · Michael R. Gionfriddo [1,3] ·
Shalak Gunjal [4] · Ann M. Farrell [5] · Virginia M. Miller [6] ·
Mohammad Hassan Murad [3,4,7]

Abstract Sleep complaints are reported by 40–60 % of menopausal women. Poor sleep is a risk factor for cardiovascular disease, diabetes, and obesity. The effect of menopausal hormone therapy on sleep quality is unclear. A systematic review and meta-analysis were conducted to summarize the efficacy of menopausal hormone therapy on self-reported sleep quality. Electronic databases (PubMed, Scopus, Ovid MEDLINE, EMBASE, EBM Reviews CENTRAL, and PsycInfo) were searched from 2002 to October 2015. Randomized trials assessing the effect of menopausal hormone therapy with a minimum follow up of 8 weeks were included. Titles, abstracts, and full texts were screened independently and in duplicate. Primary outcome included sleep items within a questionnaire, scale or diary. Standardized mean differences across trials were pooled using random-effects models. The search identified 424 articles, from which 42 trials were included. Seven trials at a moderate to high risk of bias enrolling 15,468 women were pooled in meta-analysis. Menopausal hormone therapy improved sleep quality in women who had vasomotor symptoms at baseline [standardized mean difference –0.54 (−0.91 to −0.18), moderate quality evidence]. No difference was noted when women without such symptoms were analyzed separately or combined. Across 31 sleep quality questionnaires, daytime dysfunction was the most evaluated sleep domain. Menopausal hormone therapy improves sleep in women with concomitant vasomotor symptoms. Heterogeneity of trials regarding study population, formulations, and sleep scales; limit overall certainty in the evidence. Future menopausal hormone therapy trials should include assessment of self-reported sleep quality using standardized scales and adhere to reporting guidelines.

Keywords Menopause · Sleep disturbances · Vasomotor symptoms · Estrogen replacement

✉ Mohammad Hassan Murad
  murad.mohammad@mayo.edu

1 Mayo Graduate School, Mayo Clinic, Rochester, MN, USA

2 Center for Sleep Medicine, Mayo Clinic, Rochester, MN, USA

3 Knowledge and Evaluation Research Unit, Mayo Clinic, Rochester, MN, USA

4 Evidence-Based Practice Center, Mayo Clinic, Rochester, MN, USA

5 Mayo Medical Library, Mayo Clinic, Rochester, MN, USA

6 Departments of Surgery and Physiology and Biomedical Engineering, Mayo Clinic, Rochester, MN, USA

7 Division of Preventive Medicine, Mayo Clinic, Rochester, MN, USA

Introduction

From 2000 to 2010 the number of women at the age range expected to transition into menopause (50–54 years of age) increased by 26.6 % [1].Worldwide, it is estimated that by 2025, the number of postmenopausal women will be 1.1 billion. With increased life expectancies, women live a third of their life after menopause, with some having the decline of menopause symptoms take many years [2]. The burden associated with untreated menopausal symptoms results in more frequent outpatient visits and incremental health care costs [3]. Menopausal hormone therapy (MHT) is one of the most common treatments used to counteract these symptoms.

The applicability of the evidence for the use of MHT is complicated by the heterogeneity of available trials in terms of age at which MHT is initiated, dose and type of estrogen, contraindications, and adjunct therapies. Additionally, concerns from the Women's Health Initiative (WHI) report in 2002, led guidelines to advise shorter exposure to MHT. Yet, a decade after the WHI, the use of low-dose MHT has remained constant [4], gynecologist continue to favor MHT [5], and guidelines recommend MHT as the most effective treatment for menopause symptoms, including sleep disturbances [6].

Approximately 40–60 % of menopausal women report sleep related symptoms, with the most common complaint nighttime awakenings [7]. The mechanism by which sleep disturbances arise during menopause is still unclear, and studies characterizing how other menopausal symptoms are associated to sleep alterations are conflicting. An inverse relationship between sleep quality and vasomotor symptoms (VMS) has been reported. Sleep difficulties, however, could present independently [8]. Poor sleep is a risk factor for cardiovascular disease, diabetes, obesity, and neurobehavioral dysfunction [9]. Therefore, reducing the burden of emerging sleep symptoms during menopause will result in an improvement in quality of life and overall health.

Sleep symptoms can be measured objectively (e.g., polysomnography) or subjectively (e.g., questionnaire, severity scale or diary). A previous systematic review has shown that patient reported measurements are highly predictive of quality of sleep [10], and a guideline has emphasized that such measures are important for diagnosing and monitoring response to treatment in many sleep disorders, including insomnia [11]. Further, they both empower patients and aid clinicians in recognizing and valuing the patient's perspective in response to treatment [12, 13]. These previous publications, however, addressed all adults, and did not tailor their conclusions to postmenopausal women. Therefore, understanding the effects of MHT on subjective sleep quality is important in helping patients and their clinicians manage the symptoms of menopause. However, synthesized evidence is scarce in regards to MHT effects on sleep quality leading to clinical uncertainty when choosing the best treatment.

This systematic review and meta-analysis aims to (1) evaluate the effects of MHT on self-reported sleep outcomes when compared to placebo in postmenopausal women and (2) explore the use of a multi-domain assessment of sleep quality across trials.

Methods

This review adheres to the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA)

statement [14] and was guided by a registered protocol (PROSPERO CRD42015027189). Screening and extraction was performed using online software (https://www.covidence.org/).

Eligibility criteria

Randomized clinical trials (RCTs) that compared the effects of MHT, to each other or placebo, on self-reported outcomes within a sleep questionnaire, symptom scale, or quality of life assessment tool were included. Selection was not restricted by blinding scheme, type or dose of MHT, whether sleep was a primary or secondary outcome, or type of self-reported sleep measurement tool. Minimal intervention length was 8 weeks. This timing was chosen arbitrarily as there is no current agreement on duration for which MHT changes in sleep quality would be anticipated. However, MHT alleviation of other menopausal symptoms have shown benefit as early as 8 weeks [15]. Trials where MHT was combined with compounds other than progesterone derived or selective estrogen receptor modulators were excluded. Women at any stage of natural or surgical menopause above 40 years old were included [16].

Identification and selection of trials

An experienced librarian developed search strategies, using methods recommended by the Institute of Medicine [17], in the following databases: PubMed, Scopus, Ovid MEDLINE, Ovid EMBASE, Ovid EBM Reviews CENTRAL, and Ovid PsycInfo (for search strategy, eAppendix in electronic supplementary). Search included MESH headings and keywords such as menopause, estrogen, and sleep. Databases were searched from 2002 to October 2015, aiming to gather evidence produced or published during and after the WHI reports. Electronic search was supplemented by hand searching eligible articles. There were no language restrictions with non-English articles translated by fluent bilingual speakers. Full texts of included trials were screened in duplicate and independently ($\kappa = 0.74$) [18], and disagreements were resolved by arbitration.

Data collection and study appraisal

Data were extracted using an electronic form designed by the reviewers; which was tested and piloted and contained information on patient characteristics, intervention descriptions, methodological quality indications, and outcomes of interest. Data extraction and risk-of-bias assessment were performed independently by two reviewers. Mean and standard deviation at baseline and longest follow up were extracted for the outcome. RCTs were assessed for methodological quality using the Cochrane Risk of Bias

Tool [19]. Available study protocols were searched in trial registries. If the blinding of study participants or personnel was rated to be at a high or unclear risk of bias, the trial was considered to be at high risk of bias overall. If all domains were judged to be at low risk of bias, the trial was considered at a low risk of bias. Otherwise, the trial was considered to be at a moderate risk of bias. The quality of evidence was evaluated using the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach [20].

Author contact

When scores for sleep items within questionnaires were not available or when clarification was necessary the corresponding author was contacted by e-mail. If there was no response, a second, final e-mail was sent after 2 weeks. Authors were given 6 weeks to answer and send requested information.

Meta-analysis

Standardized mean differences (SMD) were pooled using random-effects models. This approach was preferred given the construct of sleep quality was evaluated using different scales, thus, the results were standardized and expressed using standard deviation units to allow meta-analysis. SMD results can be interpreted as $0.2 = $ small effect, $0.5 = $ moderate effect, and $0.8 = $ large effect [19]. For all trials, lower scores indicated better sleep quality (the direction was reversed for one trial to be consistent with the rest). In a trial with more than one active MHT arm, the weighted SMD between groups was compared to placebo. To explain possible inconsistencies across trials, a sensitivity analysis was used to assess the effect of the WHI on the pooled estimate effect. Inconsistency of effects across trials was assessed using forest plots and the I^2 statistic with values over 50 % indicative of moderate to high heterogeneity [21]. Statistical analyses, including overall and subgroup effect estimates, were done using Review Manager v5.3 [22].

Subgroup analysis

Pre-specified subgroup analyses were performed to explore heterogeneity. Trials with inclusion criteria restricted to women with presence of VMS (hot flashes and night sweats) were compared to trials with no VMS criteria. To address duration of MHT and risk of bias, subgroup analysis by duration of intervention (8 weeks vs. >8 weeks) and by overall risk of bias (moderate vs. high) were performed.

Outcome assessment

Two board-certified sleep specialists (M.L.; R.L.) classified sleep items across multiple measurement tools using the seven sleep domains of the Pittsburgh Sleep Quality Index (PSQI). This allowed standardizing results to seven sleep characteristics routinely assessed in clinical interviews of patients with sleep complaints [23]. The PSQI was found to cover most domains of relevance to researchers when studying sleep disorders [10]. For each questionnaire, with at least one self-reported sleep item, each sleep specialist reviewed questionnaire items and dichotomized each under primary and, if applicable, secondary domains of sleep quality. Conflicts were resolved by consensus.

Results

Search strategy and contact of authors

The search identified 424 articles of these, 234 were excluded at title and abstract screening. The full text assessment of the remaining 190 articles resulted in 64 meeting eligibility criteria. After merging multiple publications of a same trial, a total of 42 RCTs were included (Fig. 1). Thirty-five trials were missing data necessary for appraisal. Authors were contacted for 23 trials; the remaining 10 did not provide contact information. Two of 23 contacted authors provided the requested data [24, 25]. Thereafter, a total of nine trials had complete report of sleep quality.

Description of trials

The data from 42 trials was used for qualitative assessment and is summarized in supplementary eTable1. Across all RCTs, 21 where comparing MHT interventions to each other and nine comparing MHT to placebo. The most commonly administered formulation was oral conjugated equine estrogen (o-CEE) at a dose of 0.625 mg/day (12/42 trials).

The definition of menopause was variable. Most trials used self-report of last menstrual period (LMP) as definition with intervals post-LMP ranging from 6 months to 10 years. Seven of the 42 trials included sleep quality as a primary outcome measure. Across trials significant variability was found on reports of MHT effect on sleeping problems.

One trial was judged to be at low risk of bias, 23 (55 %) at moderate risk and 18 (43 %) at high risk (supplementary eTable 2). Sequence generation and blinding of outcome assessors were the domains least reported. The 18 trials rated at high risk of bias had either not blinded participants or not clearly reported blinding methodology.

Fig. 1 PRISMA flow diagram for study selection process * denotes number of included articles does not match number of included RCTs, as some trials had multiple publications
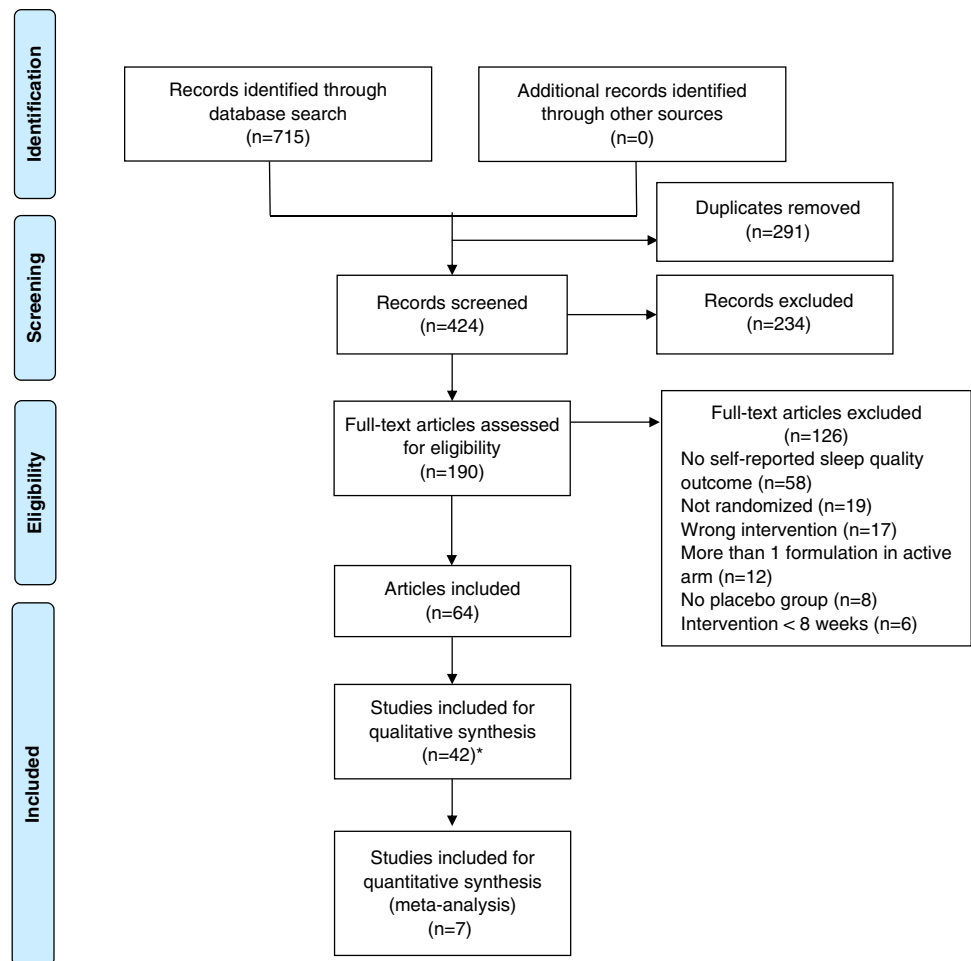
Effects of MHT on sleep quality

From the 42 trials, nine trials had mean and standard deviation reports at baseline and longest follow up. Seven of the nine trials included a placebo treatment arm [15, 24–29], while 2 had parallel comparisons of MHT formulations [30, 31]. Therefore, the seven RCTs with placebo arm as comparator had similar interventions and reported sufficient quantitative data to allow for statistical pooling (Table 1). The trials were at moderate to high risk of bias (Table 2).

Subgroup analysis showed that MHT improved sleep quality among women who had concomitant VMS [SMD −0.54 (−0.91 to −0.18, $I^2 = 0$ %), moderate quality evidence]; test for subgroup difference $p = <0.007$. No significant difference was noted in trials that included women without VMS criteria [SMD −0.04 (−0.15 to 0.24, $I^2 = 43$ %)], or when both groups were combined

(with and without VMS), [SMD −0.12 (−0.37 to 0.13, $I^2 = 66$ %)]. Results are depicted as a forest plot in Fig. 2.

A sensitivity analysis performed to examine whether the WHI affected effect estimate showed no significant difference [SMD −0.17 (0.35–0.02, $I^2 = 53$ %)]. Subgroup analysis comparing duration of MHT and risk of bias did not show significant differences as shown in supplementary eFig. 1 and eFig. 2.

Outcome assessment

Across 31 self-report sleep tools, the most frequently assessed domains of sleep quality were daytime dysfunction followed by sleep quality and sleep disturbances (Fig. 3). Prior medication use for aid in sleep was only assessed in two scales. Three scales were not accessible for item dichotomizing, two were independently created by authors'

Table 1 Characteristics of included eligible randomized placebo controlled trials with available data for meta-analysis

Reference	Country	Sample size	Menopause definition	Inclusion criteria	Type of MHT (mg/day)	Duration (weeks)	Sleep quality scale
Hays et al. [26]	USA	166,008	50–54 y/o + LMP > 12 m; >54 y/o + LMP > 6 m	—	o-CEE (0.625)+ MPA (2.5)	156	WHIIRS
Saletu-Zyhlarz et al. [27]	Austria	55 +	LMP > 24 m + E < 55 pg/mL + FSH > 19 mIU/mL + KI > 15	Insomnia+VMS	o-EV(2) + dienogest (3); o-EV (2)	8	PSQI
Sherman et al. [25]	USA	249	>55 y/o + LMP > 5 years	Coronary artery disease	o-CEE (0.625); o-CEE (0.625) + MPA (2.5)	156	WHIIRS
Heinrich et al. [28]	Germany	51	Hysterectomy	VMS	o-EV (2); o-EV (2) + mP (100)	24	ADSK Sleep item + Sleep item from Menopausal Index (combined)
Kalleinen et al. [29]	Finland	18 +	By age and amenorrhea (no further details)	—	o-EV (2) + MDP (5)	24	BNSQ
LeBlanc et al. [15]	USA	37	6 m > LMP < 36 m + FSH > 920 IU/L	—	o-EV (2) + MDP (5)	8	OHSU SL sleep diary
Savolainen-Peltonen et al. [24]	Finland	128	6 m > LMP > 36 m + FSH > 30 U/L	½ women with severe VMS, ½ symptom free	o-E2; o-E2 + MPA (5); t-E2	24	WHQ sleep item

LMP last menstrual period, *m* months, *VMS* vasomotor symptoms, *WHIIRS* Women's Health Initiative Insomnia Rating Scale, *ADSK* German version of Center for Epidemiological Studies Depression Scale, authors used only sleep item for sleep quality assessment, *BNSQ* Basic Nordic Sleep Questionnaire, *OHSU SL* Oregon Health and Science University Sleep Laboratory, *PSQI* Pittsburgh Sleep Quality Index, *WHQ* Women's Health Questionnaire, *o-* oral, *t-* transdermal CEE, conjugated equine estrogen, *MPA* medroxyprogesterone acetate, *EV* estradiol valerate, *mP* micronized-progesterone, *E2* 17 beta estradiol, *E2* 17 beta estradiol or insomnia

Table 2 Cochrane risk of bias quality assessment

Reference	SG	AC	BP	BOA	IOD	SOR	Other	Overall
Hays et al. [26]	Unclear	High[a]	Low	Unclear	Unclear	Unclear	High[b]	High
Saletu-Zyhlarz et al. [27]	Unclear	Unclear	Low	Unclear	Unclear	Unclear	High[c]	Moderate
Sherman et al. [25]	Low	Unclear	Low	Low	Unclear	Low	Unclear	Moderate
Heinrich et al. [28]	Unclear	Unclear	Unclear	Unclear	High[d]	Unclear	Low	High
Kalleinen et al. [29]	Low	Low	Unclear	Low	Low	Unclear	High[e]	High
LeBlanc et al. [15]	High[f]	Unclear	Low	Unclear	Low	Unclear	High[g]	High
Savolainen-Peltonen et al. [24]	Low	Low	Low	Unclear	Low	Unclear	Low	Moderate

SG sequence generation, *AC* allocation concealment, *BP* blinding of participants, *BOA* blinding of outcome assessors, *IOD* I incomplete outcome data, *SOR* selective outcome reporting

[a] Women on estrogen alone group were un-blinded and reassigned to estrogen + progestin group

[b] Early termination of trial, significant conflict of interest of authors with private sponsor

[c] One of the authors is private sponsor employee and no conflict disclosure was provided

[d] High proportion of patients withdrew from study, unclear to which group participants belonged too

[e] Different time-points across patients for sleep quality outcome assessment, some did not complete 6 months (follow up to 3 months but authors report as 6 months)

[f] Five participants were stratified by a third party after computer generated randomization was completed

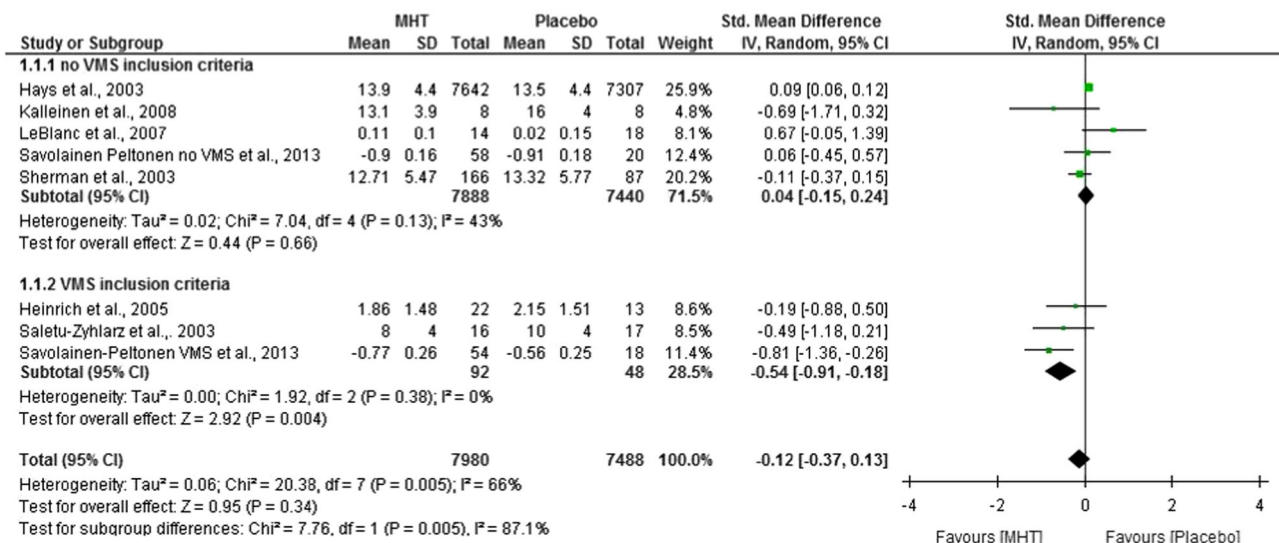[g] Significant baselines sleep outcome imbalances across treatment groups

Fig. 2 SMD for subgroup analysis by VMS, smaller scores indicate better sleep quality. The green square markers indicate standardized mean difference from primary studies, with sizes reflecting the statistical weight of the study using random-effects meta-analysis. The horizontal lines indicate 95 % confidence intervals. The diamond markers represent the subtotal and overall effect estimate and 95 % confidence intervals. SMD interpretation, 0.2 = small effect, 0.5 = moderate effect, >0.8 = large effect

institutions and were not provided, and one was inaccessible through library resources (supplementary eTable 3).

Quality of evidence

The certainty in the estimates following the GRADE approach was moderate confidence in women with VMS and low in women without VMS (Table 3).

Discussion

Summary of evidence

In this systematic review and meta-analysis of the effects of MHT on sleep quality, seven RCTs provided similar interventions and sufficient data for meta-analysis. MHT was associated with modest improved sleep quality in

Fig. 3 Distribution of the PSQI seven domains of sleep quality across 27 self-reported sleep scales used in included studies
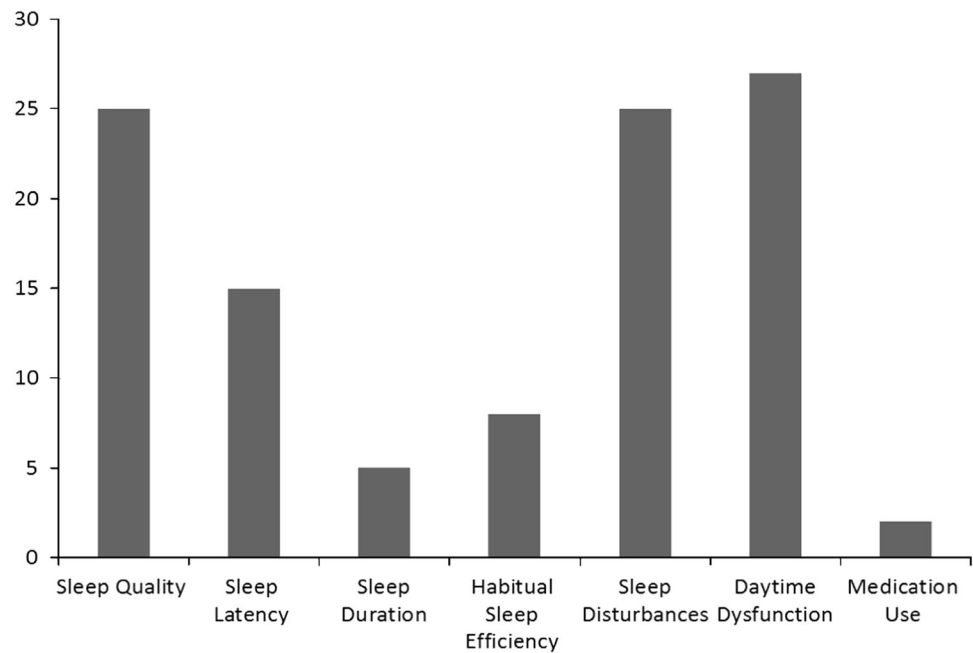
women with concomitant VMS at baseline. The effect of MHT is uncertain in women without VMS.

The heterogeneity in trial populations and formulations of MHT limit conclusions. The absorption, distribution, and metabolism of MHT differ among women based on genotype, age, distribution of adipose tissue, comorbidities, and use of other medications [32]. These covariates should help guide the design of future comparative effectiveness trials. Following the WHI, the use of low-dose transdermal estrogen increased more than tenfold [4], yet only three trials [24, 33, 34] had a direct comparison between routes of administration. Additionally, there is still a need for a standard definition of menopause, given that both age and years from menopause have shown to be important indicators of the benefit-risk ratio of MHT [35].

Self-reported sleep quality captures different parameters of sleep than objective measurements [36]. Lack of accessibility to polysomnography resources, and the limited utility of this clinical test within a large population setting, supports the need to develop validated self-reported sleep measurements in menopausal women. It is understood that sleep is best characterized across multiple domains including quality, duration, continuity and effects on daytime function [37]. A thorough assessment of these measurable characteristics of sleep quality, results in a detailed sleep scenario that is understood by both health professionals and patients [38].

In the present analysis, daytime dysfunction, sleep disturbances and overall sleep quality were the most commonly assessed domains. The other domains were infrequently incorporated into questionnaires, including sleep duration and latency (the ease of falling asleep), which have both been associated to negative health outcomes such as higher mortality, coronary heart disease, and diabetes [38]. This finding underscores the importance of further work in the area of standardizing sleep assessment tools.

Limitations and strengths

The systematic review faces a number of limitations. First, the majority of trials lacked a baseline screen for sleep disorders. After menopause, there is an increased risk of sleep disordered breathing due to fluctuating hormones and weight gain. Yet, only three trials [27, 39, 35], had exclusion or testing criteria for sleep related breathing disorders, narcolepsy or periodic limb movements. Second, evidence from this review cannot discern the magnitude of effect on sleep quality through indirect reduction of mood disturbances or frequency and severity of VMS, both known to affect sleep. This is due to heterogeneity in enrollment, as most trials do not follow the recommendations of the Food and Drug Administration for studies assessing treatment of moderate to severe VMS, where participants enrolled should have a minimum of 7 to 8 moderate to severe hot flushes per day, or 50 to 60 per week at baseline [40]. Finally, MHT formulations vary in the inclusion of progesterone or selective estrogen receptor modulator compounds. Progesterone has independent effects on sleep, through anxiolytic and respiratory stimulant action [41]. A number of studies used progestin

Table 3 Summary of the evidence quality grading using GRADE

Outcome	Quality assessment							Effect	Quality
	No. of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SMD (95 %CI)	
Sleep quality overall	8	RCT	Serious[a]	Not serious	Not serious	Serious[b]	Serious[c]	−0.12 (−0.36, 0.13)	⊕⊕◯◯ Low
Sleep quality VMS subgroup	3	RCT	Serious[a]	Not serious	Not serious	Not serious	Serious[c]	−0.54 (−0.91, −0.17)	⊕⊕⊕◯ Moderate

The GRADE working group grades of evidence are as follow: 1. High quality, future research is very unlikely to change the group's confidence in the estimate of effect; 2. Moderate quality, further research is likely to have an important impact on the group's confidence in the estimate of effect and may change the estimate; 3. Low quality, further research is very likely to have an important effect on the group's confidence in the estimate of effects and is likely to change the estimate; 4. Very low quality, the group is very uncertain about the estimate

[a] The quality was downgraded because there was overall poor reporting of blinding and allocation concealment

[b] Wide confidence intervals in our estimate effect limit conclusions

[c] Large numbers of studies do not contribute to sleep quality outcome, suggesting possible reporting bias

CI confidence interval, SMD standardized mean difference

synthetic derivatives that may not have the same effects as progesterone. Therefore, the independent effects of estrogen vs. progesterone and progestin compounds require further evaluation.

This systematic review also has several strengths. It provides a comprehensive review of the current evidence guided by an a priori registered protocol, with an extensive search for eligible studies in multiple databases including both published and unpublished work. All of the included studies were assessed in duplicate and registered protocols were searched on online trial protocol databases. We incorporated an in depth assessment of the sleep quality outcome scales, to best provide information for clinicians to understand and apply in practice. This is, to our knowledge, the first use of dichotomizing scheme to bring together self-reported sleep quality scales in menopausal women.

Implications for practice and research

Sleep disturbances are a common indication for MHT. Yet, at present there is insufficient evidence to determine how a woman's self-reported sleep quality during menopause is affected by different routes of administration or formulations of MHT.

Future research focused on determining if MHT is beneficial in improving sleep quality during menopause, will be best addressed by head-to-head RCTs between various formulations and routes of administration. Likewise, sleep quality assessment tools need to be developed with a consensus on a tool with sufficient domains of sleep quality and validated in menopausal women. We suggest use of the seven major sleep domains listed in the PSQI. Yet, important questions specific to menopause must also be addressed in the context of sleep quality. The evidence derived from this systematic review suggests that MHT benefits sleep in women with VMS. This is congruent with current North American Menopause Society recommendations where use of MHT for sleep disturbance is suggested for women with bothersome nighttime hot flashes [42]. Other guidelines lack specifics regarding MHT use, and may benefit from incorporating these results.

Conclusion

It is imperative to endorse meaningful conversations about sleep quality in the clinical setting with female patients nearing menopausal age and to tailor treatment recommendations towards patient-specific complaints. Validated clinical screening tools addressing various domains of sleep quality are needed in order to better describe the current burden of sleep disturbances in this population.

References

1. L. Howden, J. Meyer, Age and sex composition: 2010. 2010 Census Briefs (2011)
2. Z.A. Al-Safi, N. Santoro, Menopausal hormone therapy and menopausal symptoms. Fertil. Steril. **101**, 905–915 (2014)
3. P. Sarrel, D. Portman, P. Lefebvre et al., Incremental direct and indirect costs of untreated vasomotor symptoms. Menopause **22**, 260–266 (2015)
4. B. Ettinger, S.M. Wang, R.S. Leslie et al., Evolution of postmenopausal hormone therapy between 2002 and 2009. Menopause **19**, 610–615 (2012)
5. L. Danckers, J.E. Blumel, S. Witis et al., Personal and professional use of menopausal hormone therapy among gynecologists: a multinational study (REDLINC VII). Maturitas **87**, 67–71 (2016)
6. C.A. Stuenkel, S.R. Davis, A. Gompel et al., Treatment of symptoms of the menopause: an endocrine society clinical practice guideline. J. Clin. Endocrinol. Metab. **100**, 3975–4011 (2015)
7. H.M. Kravitz, P.A. Ganz, J. Bromberger, L.H. Powell, K. Sutton-Tyrrell, P.M. Meyer, Sleep difficulty in women at midlife: a community survey of sleep and the menopausal transition. Menopause **10**, 19–28 (2003)
8. M. Bruyneel, Sleep disturbances in menopausal women: aetiology and practical aspects. Maturitas **81**, 406–409 (2015)
9. F.S. Luyster, P.J. Strollo Jr., P.C. Zee, J.K. Walsh, Sleep: a health imperative. Sleep **35**, 727–734 (2012)
10. E.B. Devine, Z. Hakim, J. Green, A systematic review of patient-reported outcome instruments measuring sleep dysfunction in adults. Pharmacoeconomics **23**, 889–912 (2005)
11. S. Schutte-Rodin, L. Broch, D. Buysse, C. Dorsey, M. Sateia, Clinical guideline for the evaluation and management of chronic insomnia in adults. J. Clin. Sleep Med. **4**, 487–504 (2008)
12. P.R. Deshpande, S. Rajan, B.L. Sudeepthi, C.P. Abdul Nazir, Patient-reported outcomes: a new era in clinical research. Perspect. Clin. Res. **2**, 137–144 (2011)
13. I. Wiklund, Assessment of patient-reported outcomes in clinical trials: the example of health-related quality of life. Fundam. Clin. Pharmacol. **18**, 351–363 (2004)
14. D. Moher, A. Liberati, J. Tetzlaff, D.G. Altman, Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann. Intern. Med. **151**(264–269), W64 (2009)
15. E.S. LeBlanc, M.B. Neiss, P.E. Carello, M.H. Samuels, J.S. Janowsky, Hot flashes and estrogen therapy do not influence cognition in early menopause women. Menopause **14**, 191–202 (2007)
16. S.D. Harlow, M. Gass, J.E. Hall et al., Executive summary of the Stages of Reproductive Aging Workshop+10: addressing the unfinished agenda of staging reproductive aging. Climacteric **15**, 105–114 (2012)
17. J. Eden, L. Levit, A. Berg, S. Morton eds. *Finding What Works in Health Care: Standards for Systematic Reviews.* (National Academy of Sciences, Washington, DC, 2011)
18. T. McGinn, P.C. Wyer, T.B. Newman, S. Keitz, R. Leipzig, G.G. For, Tips for learners of evidence-based medicine: 3. Measures of observer variability (kappa statistic). CMAJ **171**, 1369–1373 (2004)
19. J. Higgins, S. Green, Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011] (The Cochrane Collaboration, 2011).
20. N. Meader, K. King, A. Llewellyn et al., A checklist designed to aid consistency and reproducibility of GRADE assessments: development and pilot validation. Syst. Rev. **3**, 82 (2014)
21. J.P. Higgins, S.G. Thompson, J.J. Deeks, D.G. Altman, Measuring inconsistency in meta-analyses. BMJ **327**, 557–560 (2003)
22. Review Manager (RevMan) [Computer program]. Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014
23. D.J. Buysse, C.F. Reynolds 3rd, T.H. Monk, S.R. Berman, D.J. Kupfer, The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res. **28**, 193–213 (1989)
24. H. Savolainen-Peltonen, H. Hautamaki, P. Tuomikoski, O. Ylikorkala, T.S. Mikkola, Health-related quality of life in women with or without hot flashes: a randomized placebo-controlled trial with hormone therapy. Menopause **21**, 732–739 (2014)
25. A.M. Sherman, S.A. Shumaker, P. Sharp et al., No effect of HRT on health-related quality of life in postmenopausal women with heart disease. Minerva Ginecol. **55**, 511–517 (2003)
26. J. Hays, J.K. Ockene, R.L. Brunner et al., Effects of estrogen plus progestin on health-related quality of life. N. Engl. J. Med. **348**, 1839–1854 (2003)
27. G. Saletu-Zyhlarz, P. Anderer, G. Gruber et al., Insomnia related to postmenopausal syndrome and hormone replacement therapy: sleep laboratory studies on baseline differences between patients and controls and double-blind, placebo-controlled investigations on the effects of a novel estrogen-progestogen combination (Climodien, Lafamme) versus estrogen alone. J. Sleep Res. **12**, 239–254 (2003)
28. A.B. Heinrich, O.T. Wolf, Investigating the effects of estradiol or estradiol/progesterone treatment on mood, depressive symptoms, menopausal symptoms and subjective sleep quality in older healthy hysterectomized women: a questionnaire study. Neuropsychobiology **52**, 17–23 (2005)
29. N. Kalleinen, O. Polo, S.L. Himanen, A. Joutsen, P. Polo-Kantola, The effect of estrogen plus progestin treatment on sleep: a randomized, placebo-controlled, double-blind trial in premenopausal and late postmenopausal women. Climacteric **11**, 233–243 (2008)
30. A.F. Polisseni, A.T. Andrade, L.C. Ribeiro et al., Effects of a continuous-combined regimen of low-dose hormone therapy (oestradiol and norethindrone acetate) and tibolone on the quality of life in symptomatic postmenopausal women: a double-blind, randomised study. Maturitas **74**, 172–178 (2013)
31. A. Elfituri, F. Sherif, M. Elmahaishi, H. Chrystyn, Two hormone replacement therapy (HRT) regimens for middle-eastern postmenopausal women. Maturitas **52**, 52–59 (2005)
32. B.R. Bhavnani, F.Z. Stanczyk, Pharmacology of conjugated equine estrogens: efficacy, safety and mechanism of action. J. Steroid Biochem. Mol. Biol. **142**, 16–29 (2014)
33. J.M. Kling, B.A. Lahr, K.R. Bailey, S.M. Harman, V.M. Miller, S.L. Mulvagh, Endothelial function in women of the Kronos Early Estrogen Prevention Study. Climacteric **18**, 187–197 (2015)
34. D. Serrano, L. Mariani, S. Mora et al., Quality of life assessment in a chemoprevention trial: fenretinide and oral or transdermal HRT. Maturitas **55**, 69–75 (2006)
35. H. Joffe, L.F. Petrillo, A. Koukopoulos et al., Increased estradiol and improved sleep, but not hot flashes, predict enhanced mood during the menopausal transition. J. Clin. Endocrinol. Metab. **96**, E1044–E1054 (2011)

36. D.J. Buysse, S. Ancoli-Israel, J.D. Edinger, K.L. Lichstein, C.M. Morin, Recommendations for a standard research assessment of insomnia. Sleep **29**, 1155–1173 (2006)

37. M.H. Hall, C.E. Kline, S. Nowakowski, Insomnia and sleep apnea in midlife women: prevalence and consequences to health and functioning. F1000Prime Rep. **7**, 63 (2015)

38. D.J. Buysse, Sleep health: can we define it? Does it matter? Sleep **37**, 9–17 (2014)

39. H. Hachul, L.R. Bittencourt, M.L. Andersen, M.A. Haidar, E.C. Baracat, S. Tufik, Effects of hormone therapy with estrogen and/or progesterone on sleep pattern in postmenopausal women. Int. J. Gynaecol. Obstet. **103**, 207–212 (2008)

40. Guidance for Industry: Estrogen and Estrogen/Progesting Drug Products to Treat Vasomotor Symptoms and Vulvar and Vaginal Atrophy Symptoms-Recommendations for Clinical Evaluation http://google2.fda.gov/search?q=estrogen&client=FDAgov&site=FDAgov&lr=&proxystylesheet=FDAgov&requiredfields=-archive%3AYes&output=xml_no_dtd&getfields=* accessed 08/05/16

41. P.S. Eichling, J. Sahni, Menopause related sleep disorders. J. Clin. Sleep Med. **1**, 291–300 (2005)

42. J.L. Shifren, M.L. Gass, The North American Menopause Society recommendations for clinical care of midlife women. Menopause **21**, 1038–1062 (2014)