Pharmacological interventions and hormonal therapies for depressive symptoms in peri- and post-menopausal women: a network meta-analysis of randomized controlled trials

Ping-Tao Tseng, Hsien-Jane Chi, Mein-Woei Sue, Bing-Syuan Zen, Ming-Kung Wu, Yu-Kang Tu, Kuo-Chuan Hung, Yi-Cheng W, Kuan-Pin Su, Dian-Jeng Li, Tien-Yu Chen, Brendon Stubbs, Andre F Carvalho, Marco Solmi, Trevor Thompson, Maria Gabriella Caruso, Yutaka J Matsuoka, Yen-Wen Chen, Pao-Yen Lin, Cheuk-Kwan Sun, Yu-Shian Cheng, Yow-Ling Shiue

Abstract

Although significant portion of women experience depressive symptoms during or after menopausal transition, there has been considerable controversy over the benefits of hormone replacement therapy (HRT) and antidepressants due to insufficient evidence supporting the superiority of either treatment. This frequentist model based network meta-analysis (NMA) included randomized controlled trials (RCTs) of menopausal depression symptoms management in menopausal women. Seventy RCTs involving a total of 18,530 women (mean age 62.5) were analyzed. The results demonstrated that fluoxetine plus oral HRT [standardized mean difference (SMD)=-1.59, 95% confidence interval (95%CIs)=-2.69 to -0.50] were associated with the largest improvement in depressive symptoms than placebos in overall menopausal women. Similar findings were also noted in the subgroup of participants with a definite diagnosis of depression, while no pharmacological or hormone replacement therapy was better than placebo in the subgroup of post-menopausal women (amenorrhea > 1 year) or in patients without diagnosis of depression. This NMA presented evidence that fluoxetine plus HRT may be beneficial to menopausal women with a definite diagnosis of depression but not to those without depression or post-menopausal women. Trial registration: PROSPERO (CRD42020167459).

1. Introduction

Menopausal transition is a known risk factor of depression in women (Kessler, 2003; Michaud et al., 2001; Mishra and Kuh, 2012), and up to two thirds of women experienced mild to moderate psychological symptoms on encountering menopause (Cohen et al., 2006). Previous literature about depressive symptoms during the menopausal transition and post-menopausal periods is somewhat contentious (Soares, 2013); some authors have suggested that menopausal transition may be a 'window of vulnerability' for women due to sensitivity to premenstrual changes in the hormonal milieu (Maki et al., 2019), while depression in women after menopause with a declining estrogen level may resemble geriatric depression that bears a prevalence comparable to that of their male counterparts (Bebbington et al., 2003; Soares, 2013). Hence, different treatment approaches may be required for women at different stages of menopause. However, despite the huge impact of menopausal depression, there has not been conclusive evidence regarding its management. A well-conducted network meta-analysis (NMA) of randomized controlled trials would resolve this issue.

For the treatment of perimenopausal depression, a recent review suggested that currently approved therapeutic options for depression such as antidepressants and psychotherapy should first be considered (Maki et al., 2019). In contrast, evidence regarding estrogen therapy is more controversial despite its potential antidepressant effects in perimenopausal women, particularly those with concomitant vasomotor symptoms (Maki et al., 2019). However, evidence supporting endocrinologic and pharmacologic treatments (i.e., antidepressants, hormones, or both) against perimenopausal depression is far from tangible due to the limited number of trials on this issue. Current guidelines on the treatment of menopausal depression mostly focus on women during menopausal transition (Sarri et al., 2015; Stuenkel et al., 2015), while there is no consensus on postmenopausal women in this clinical setting. Moreover, the downsides of long-term hormone replacement therapy, such as an increased risk of cancers (Rossouw et al., 2002), remain unignorable safety concerns that limit its liberal use.

Therefore, the present study aimed at comparing the therapeutic benefits of different hormonal therapies and pharmacological interventions targeting the depressive symptoms of menopausal women by systematically reviewing the available clinical trials and integrating the available evidence with NMA.

2. Methods

2.1. General guidelines applied in the current study

The current NMA followed the latest preferred reporting items systematic reviews and meta-analyses (PRISMA) 2020 guidelines (Appendix 1) (Page et al., 2021) and the AMSTAR2 (Shea et al., 2017). This study is registered with PROSPERO (CRD42020167459) and was approved by the Institutional Review Board of the Tri-Service General Hospital, National Defense Medical Center, Taiwan (TSGHIRB No. B-109-29).

2.2. Search strategy and selection criteria

We conducted a systematic review of publications retrieved from the PubMed, Cochrane Library, and ScienceDirect databases from inception to April 14, 2021 with the relevant keywords (Appendix 2). No language restriction was applied. We also identified unpublished studies from other sources such as registered clinical trials which were not included in the above-mentioned databases and conducted manual searches for potentially eligible articles from the reference lists of review articles and pairwise meta-analysis (Kornstein et al., 2015; Maki et al., 2019; Whedon et al., 2017; Zweifel and O'Brien, 1997). The overall flowchart was followed our previous three NMAs (Chen et al., 2021; Chuang et al., 2022; Lei et al., 2022; Tseng et al., 2020; Tseng et al., 2022a; Wu et al., 2022; Zeng et al., 2022).

2.3. Inclusion and exclusion criteria

The PICO of the current study was: (1) Patient or Problem: female undergoing or after menopausal transition; (2) Intervention: hormonal and pharmacologic intervention; (3) Comparator: waiting-list, placebo-control, or active-control; and (4) Outcome: changes in depressive symptoms. To improve the quality of the included articles, we only included randomized clinical trials (RCTs) adopting either placebo-controlled or active-controlled design in humans. The targets of comparison arms were set to be pharmacologic and hormonal interventions applied in patients during or after menopausal transition including the peri-menopausal stage (e.g., irregular menstruation) and after the menopausal transition (i.e., absence of menstruation > 1 year).

The exclusion criteria included: (1) studies that were not clinical trials, (2) those in which changes in depressive symptoms were not reported, or (3) those not related to effects of either pharmacologic or hormonal interventions against depressive symptoms. In the situation of duplicated usage of data (i.e., different articles based on the same sample sources), we only included the reports with more information and larger sample sizes.

2.4. Data extraction

Two authors (Cheng YS and Sun CK) independently screened the studies, extracted relevant data from the articles, and assessed the risk of bias among the included studies. In the situation of discrepancy, another author (PT Tseng) was involved. If there was a lack of available data from the manuscripts [for example, the necessary standard deviation (SD) of the change means despite of the existed interquartile], we contacted the authors to obtain the original data. We followed the flowchart reported in previous NMAs (Hsieh et al., 2019; Tseng et al., 2023; Tseng et al., 2022b; Wu et al., 2019).

2.5. Outcomes

The primary outcomes were (1) changes in depressive symptoms after treatment with hormones or pharmacologic agents in women during or after menopausal transition and (2) acceptability in aspect of dropout rate. To be specific, the dropout rate was defined as leaving study before the end of the trial due to “any” reason. We conducted subgroup analysis on studies involving women after menopausal transition (i.e., absence of menstruation > 1 year) only, studies involving women during menopausal transition only, those recruiting participants with baseline diagnosis of depression in accordance with the diagnostic criteria from International Classification of Diseases tenth edition, the Diagnostic and Statistical Manual of Mental Disorders (third edition-revision, fourth edition and text revision), the Montgomery–Åsberg Depression Rating Scale higher than 13, and those recruiting participants without baseline diagnosis of depression. For those without baseline diagnosis of depression, some studies used definite criteria for excluding depression such as Beck Depression Inventory > 10 or Depression Anxiety Stress Scale >20 but most did not use any specific criteria (e.g., merely stating that participants with diagnosis of depression were excluded). Since different criteria were used for defining post-menopausal period among the included studies, Stages of Reproductive Aging Workshop criteria (Soules et al., 2001) was used in our subgroup analyses to divide participants into those during (i.e., perimenopausal stage) and those after (i.e., amenorrhea for at least one year) menopausal transition. The secondary outcome was about the safety profiles (i.e. rate of study discontinuation due to adverse events). The rate of discontinuation due to adverse events referred to leaving study before the end of the trial due to “intolerance to adverse events”.

2.6. Cochrane risk of bias tool

Two independent authors (YS Cheng and PT Tseng) evaluated the risk of bias (interrater reliability, 0.85) for each domain described in the Cochrane risk of bias tool (Tremolada et al., 2019). In addition, we evaluated the quality of evidence with the GRADE (Grading of Recommendations, Assessment, Development and Evaluations) guidelines as previously described (Cipriani et al., 2018; Puhan et al., 2014).

2.7. Statistical analysis

Our NMA, based on the frequentist model, was performed using STATA (version 16.0; StataCorp Statistics/Data Analysis, StataCorp LLC, College Station, TX, USA) with mvmeta command (White, 2015). For continuous data, we computed summary standardized mean difference (SMD) with 95% confidence intervals (95%CIs). For categorical data, we estimated summary odds ratio (OR) with 95%CIs. In the situation of categorical data, we applied a 0.5 zero-cell correction during the procedure of meta-analysis. However, if, in one study, there were both zero in intervention and control arms, we did not apply such a correction procedure because of the risk of increasing bias (Brockhaus et al., 2014; Cheng et al., 2016). The standard continuity correction method has been considered to result in statistical bias in both-zero-event studies and may result in counterintuitive statistical results (Hsu et al., 2022). Because of the presumed heterogeneity among the included studies in the aspect of sample sources, study methods, or depressive rating scales, we used the random-effects model in pairwise meta-analysis to compare the effect sizes (ES) among studies with the same interventions. The NMA process was based on and frequentist models. All comparisons were made with a two-tailed test with a *p*-value cut-off point set at 0.05 to denote statistical significance. The heterogeneity among the included studies was evaluated by the tau value, which is the estimated standard deviation of the effect across the included studies.

Regarding the procedure of meta-analysis applied in the current study, we used a mixed comparison with generalized linear mixed models to analyze the direct and indirect comparisons for NMA (Tu, 2014). To be specific, indirect comparisons could be calculated by transitivity, through which the differences between treatment A and B could be calculated from their comparisons with the third treatment, C. For comparison among multiple treatment arms, we combined the direct and indirect evidence from the included studies (Lu and Ades, 2004). For the current NMA, a suite of Stata programs using mvmeta for data manipulation was utilized (White, 2015). We used the restricted maximum likelihood method to evaluate the between-study variance (Kontopantelis et al., 2013).

To provide more clinical application, we calculated the relative ranking probabilities between the preventive effects of all treatments for the target outcomes. In brief, the surface under the cumulative ranking curve (SUCRA) is the percentage of the mean rank of each medication relative to an imaginary intervention that is the best without uncertainty (Salanti et al., 2011). The larger the area under the curve, the lower the rank of the treatment in relieving depressive symptoms.

Finally, we evaluated the potential inconsistency between the direct and indirect evidence within the network with the loop-specific approach and local inconsistency with the node-splitting method. Further, we used the design-by-treatment model to evaluate the global inconsistency among the whole NMA (Higgins et al., 2014). We used the comparison-adjusted funnel plot (Chaimani et al., 2013) and Egger regression to evaluate the potentially small study effects and publication bias. In situation of statistically significant heterogeneity or inconsistency so that they might violate the transitive hypothesis, we selected studies to perform subgroup analysis to evaluate the potential effects of confounding factors. In particular, we performed subgroup analysis of RCTs on subjects with definite depression, those excluding definite depression, those recruiting post-menopausal women only, and those recruiting peri-menopausal women only.

3. Results

After initial screening, 237 articles in total were considered eligible for full-text review (Fig. 1). However, 162 were excluded for various reasons (Fig. 1 and Appendix 4). Finally, 75 articles enrolling a total of 19,048 participants (mean age, 59.5 years; range, 44.3 to 77.5 years) were included in the current study (Appendix 5). Seventy-five articles with twenty-seven individual treatment arms and one placebo arm were investigated in the current NMA. The whole geometric distribution of the treatment arms is provided in Fig. 2**AB**.

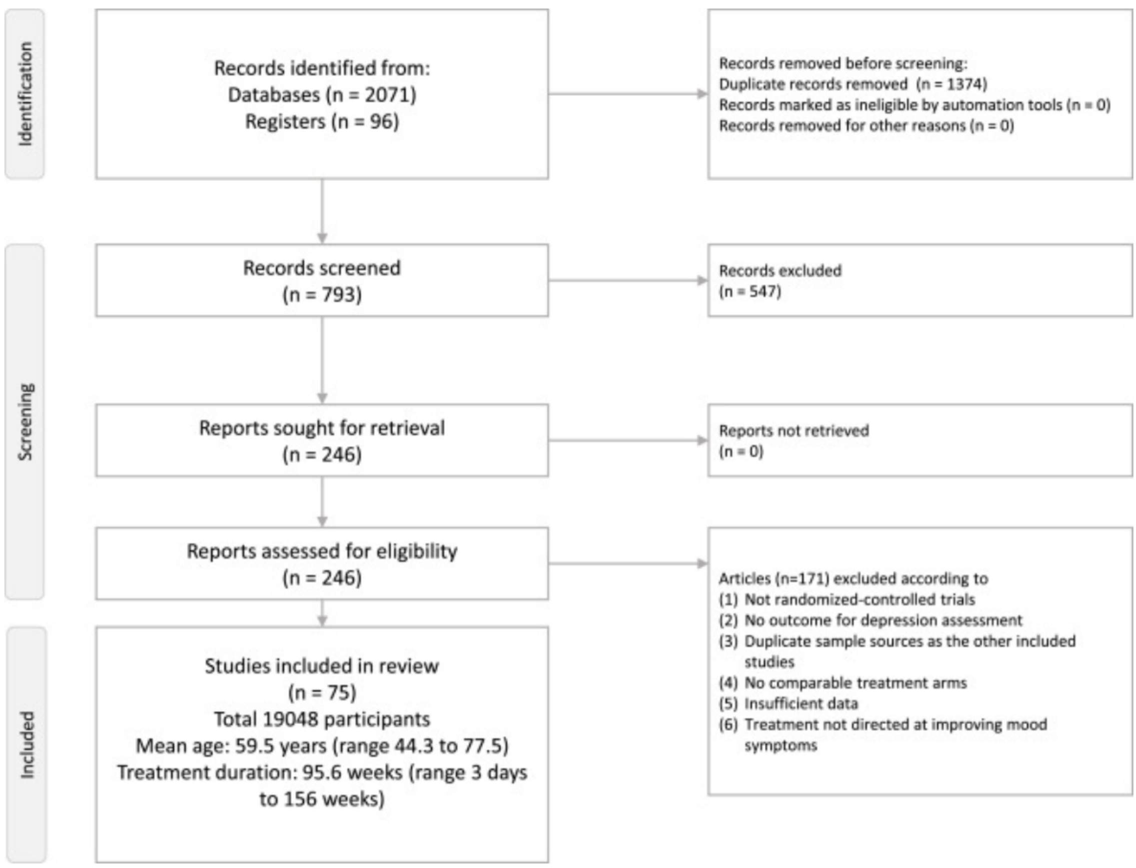

Fig. 1. PRISMA flow diagram of included studies.

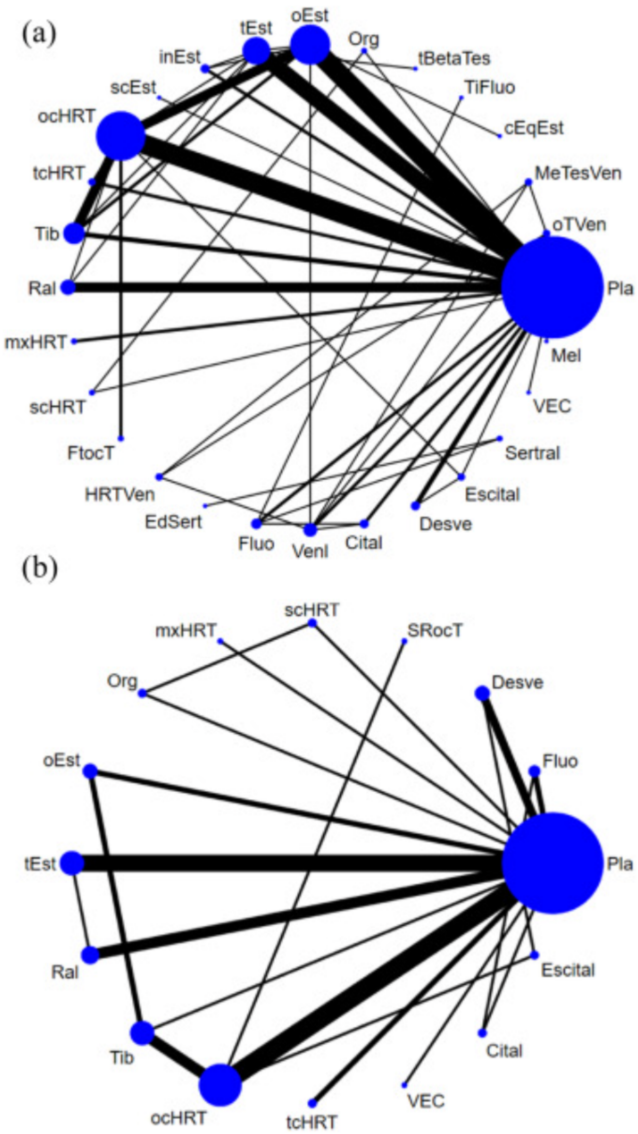

Fig. 2. ***A:*** **Network structure of the current network meta-analysis in changes in depressive severity;** ***B:*** **Network structure of the current network meta-analysis in acceptability in aspect of dropout rate**.

The lines between nodes represent direct comparisons in various trials, and the size of each circle is proportional to the size of the population involved in each specific treatment; the thickness of the lines is proportional to the number of trials connected to the network.

Abbreviation: cEqEst=conjugated equine estrogens + testosterone cream. Cital=citalopram. Desve=desvenlafaxine. EdSert=transdermal estradiol + sertraline. Escital=escitalopram.

Fluo=fluoxetine. FtocT=fluoxetine + oral combined therapy (conjugated equine estrogens + cyclic medroxyprogesterone acetate). RTVen=venlafaxine + estrogen + medroxyprogesterone acetate + methyltestosterone. inEst=intranasal estradiol. Mel=melatonin. MeTesVen=venlafaxine + methyltestosterone. mxHRT=mixed hormone replacement therapy (estrogen in women with hysterectomy; estrogen + progesterone in those without hysterectomy). ocHRT=oral combined hormone replacement therapy (estrogen + progesterone). oEst=oral estrogen. Org=estradiol implants and testosterone or (norethisterone for those without hysterectomy). oTVen=venlafaxine + oral combined therapy (conjugated equine estrogens + cyclic medroxyprogesterone acetate). Pla=placebo. scEst=subcutanenous implant of estradiol. scHRT=subcutaneous combined hormone replacement therapy (estradiol implants + norethisterone). Sertral=sertraline. tBetaTes=transdermal 17β-estradiol + testosterone cream. tcHRT=transdermal combined hormone replacement therapy (estrogen + progesterone). tEst=transdermal estradiol. Tib=tibolone. TiFluo=fluoxetine + tibolone. VEC=vaginal estrogen cream. Venl=venlafaxine

3.1. Primary outcome: (1) changes in depressive severity

With respect to primary outcome, the regimen of fluoxetine *plus* oral combined hormone therapy (conjugated equine estrogens + cyclic medroxyprogesterone acetate) [SMD=-1.58 (95%CIs=-2.64 to -0.52)], transdermal estradiol [SMD=-0.64 (95%CIs=-1.08 to -0.20)], tibolone only [SMD=-0.62 (95%CIs=-1.11 to -0.12)], and oral estrogen therapy [SMD=-0.54 (95%CIs=-0.90 to -0.10)] demonstrated significantly better therapeutic effects than those in placebo/control groups (Table 1A, Figs. 2**A**, and 3**A**). The heterogeneity had been depicted in Appendix 15. In brief, there had been some heterogeneity detected in some comparison. Therefore, we arranged subgroup analysis as below.

Table 1A. League table of overall improvements in depressive symptoms during and after menopausal transition.

PtocT																			*-1.33 (-2.08,-0.57)								
-0.25 (-2.29, 1.80)	MeTesVen					-0.49 (-1.15, 0.17)	-0.61 (-1.28, 0.05)			*-0.83 (-1.54, -0.12)																	
-0.80 (-2.60, 1.01)	-0.55 (-2.83, 1.73)	cEqEst						-0.24 (-0.80, 0.31)																			
-0.78 (-2.66, 1.10)	-0.53 (-2.88, 1.81)	0.02 (-2.13, 2.17)	tBetaTes	-0.17 (-0.85, 0.52)																							
-0.94 (-2.08, 0.19)	-0.70 (-2.50, 1.11)	-0.15 (-1.69, 1.39)	-0.17 (-1.66, 1.33)	tEst	-0.22 (-0.81, 0.37)			*0.99 (0.50, 1.49)			-0.37 (-0.86, 0.12)								-0.25 (-0.87,0.38)		-0.52 (-1.21,0.18)					*-0.86 (-1.58, -0.13)	
-0.96 (-2.08, 0.15)	-0.72 (-2.53, 1.10)	-0.17 (-1.70, 1.37)	-0.19 (-1.81, 1.43)	-0.02 (-0.65, 0.61)	Tib			-0.29 (-1.39, 0.82)											-0.19 (-0.78,0.40)							-1.01 (-2.31, 0.29)	
-0.74 (-2.76, 1.29)	-0.49 (-1.97, 1.00)	0.06 (-2.21, 2.33)	0.04 (-2.29, 2.37)	0.21 (-1.58, 2.00)	0.23 (-1.57, 2.03)	HRTVen	-0.12 (-0.74, 0.50)			-0.34 (-1.00, 0.32)																	
-0.86 (-2.89, 1.17)	-0.61 (-2.10, 0.87)	-0.06 (-2.33, 2.20)	-0.08 (-2.41, 2.25)	0.08 (-1.70, 1.87)	0.10 (-1.69, 1.90)	-0.12 (-1.59, 1.34)	oTVen			-0.22 (-0.87, 0.44)																	
-1.04 (-2.12, 0.05)	-0.79 (-2.56, 0.98)	-0.24 (-1.68, 1.20)	-0.26 (-1.85, 1.33)	-0.09 (-0.64, 0.45)	-0.07 (-0.60, 0.46)	-0.30 (-2.05, 1.45)	-0.18 (-1.93, 1.57)	oEst		-0.09 (-0.38, 0.21)	*-1.36 (-1.87, -0.86)								-0.57 (-1.44,0.31)							*-0.43 (-0.85, -0.01)	
-1.03 (-2.46, 0.40)	-0.79 (-2.76, 1.18)	-0.24 (-2.00, 1.53)	-0.26 (-2.09, 1.58)	-0.09 (-1.15, 0.97)	-0.07 (-1.15, 1.01)	-0.30 (-2.25, 1.66)	-0.17 (-2.13, 1.78)	0.01 (-1.02, 1.03)	Fluo						0.12 (-0.38, 0.61)		-0.42 (-1.31, 0.47)									*-0.63 (-1.16, -0.11)	*-1.00 (-1.75, -0.25)
-1.07 (-2.46, 0.31)	-0.83 (-2.33, 0.68)	-0.28 (-1.99, 1.44)	-0.30 (-2.09, 1.50)	-0.13 (-1.13, 0.87)	-0.11 (-1.12, 0.90)	-0.34 (-1.82, 1.15)	-0.22 (-1.70, 1.27)	-0.04 (-0.96, 0.89)	-0.04 (-1.31, 1.23)	Venl					-0.51 (-1.14, 0.11)											-0.34 (-0.99, 0.32)	
-1.15 (-2.52, 0.23)	-0.90 (-2.86, 1.06)	-0.35 (-2.07, 1.36)	-0.37 (-2.14, 1.40)	-0.21 (-1.15, 0.74)	-0.19 (-1.19, 0.82)	-0.41 (-2.36, 1.53)	-0.29 (-2.23, 1.65)	-0.11 (-1.04, 0.81)	-0.12 (-1.42, 1.19)	-0.08 (-1.33, 1.18)	inEst															-1.31 (-3.52, 0.91)	
-1.20 (-2.47, 0.06)	-0.96 (-2.85, 0.93)	-0.41 (-2.05, 1.23)	-0.43 (-2.14, 1.28)	-0.26 (-1.09, 0.57)	-0.24 (-1.10, 0.62)	-0.47 (-2.34, 1.41)	-0.34 (-2.22, 1.53)	-0.17 (-0.95, 0.62)	-0.17 (-1.36, 1.02)	-0.13 (-1.27, 1.01)	-0.05 (-1.19, 1.08)	Desve				0.12 (-0.08, 0.32)										*-0.43 (-0.62, -0.25)	
-1.23 (-2.55, 0.09)	-0.98 (-2.92, 0.96)	-0.43 (-2.13, 1.26)	-0.45 (-2.21, 1.31)	-0.28 (-1.22, 0.65)	-0.26 (-1.21, 0.68)	-0.49 (-2.41, 1.43)	-0.37 (-2.29, 1.55)	-0.19 (-1.08, 0.70)	-0.20 (-1.46, 1.07)	-0.15 (-1.37, 1.06)	-0.08 (-1.29, 1.13)	-0.02 (-1.11, 1.06)	tcHRT						-0.58 (-1.34, 0.18)							-0.19 (-0.43, 0.05)	
-1.26 (-3.06, 0.55)	-1.01 (-3.29, 1.27)	-0.46 (-2.54, 1.62)	-0.48 (-2.61, 1.65)	-0.31 (-1.83, 1.21)	-0.29 (-1.83, 1.25)	-0.52 (-2.78, 1.74)	-0.40 (-2.66, 1.87)	-0.22 (-1.71, 1.28)	-0.22 (-1.97, 1.52)	-0.18 (-1.89, 1.53)	-0.11 (-1.81, 1.60)	-0.05 (-1.67, 1.57)	-0.03 (-1.70, 1.65)	Org										-0.34 (-0.91, 0.22)		-0.33 (-0.91, 0.26)	
-1.24 (-2.66, 0.18)	-0.99 (-2.87, 0.90)	-0.44 (-2.20, 1.32)	-0.46 (-2.28, 1.36)	-0.29 (-1.33, 0.75)	-0.27 (-1.34, 0.79)	-0.50 (-2.37, 1.37)	-0.38 (-2.24, 1.49)	-0.20 (-1.20, 0.80)	-0.20 (-1.35, 0.94)	-0.16 (-1.29, 0.97)	-0.09 (-1.38, 1.21)	-0.03 (-1.21, 1.15)	-0.01 (-1.26, 1.25)	0.02 (-1.72, 1.75)	Cital											-0.37 (-0.76, 0.02)	
-1.30 (-2.64, 0.03)	-1.06 (-3.01, 0.89)	-0.51 (-2.21, 1.20)	-0.53 (-2.30, 1.25)	-0.36 (-1.31, 0.59)	-0.34 (-1.30, 0.62)	-0.57 (-2.50, 1.36)	-0.44 (-2.38, 1.49)	-0.27 (-1.18, 0.64)	-0.27 (-1.56, 1.01)	-0.23 (-1.46, 1.01)	-0.15 (-1.38, 1.07)	-0.10 (-1.03, 0.83)	-0.08 (-1.25, 1.10)	-0.05 (-1.73, 1.64)	-0.07 (-1.34, 1.20)	Escital			0.11 (-0.58, 0.81)							-0.18 (-0.46, 0.10)	
-1.45 (-3.60, 0.69)	-1.21 (-3.74, 1.33)	-0.66 (-3.04, 1.73)	-0.68 (-3.11, 1.76)	-0.51 (-2.43, 1.41)	-0.49 (-2.42, 1.44)	-0.72 (-3.24, 1.81)	-0.59 (-3.12, 1.93)	-0.41 (-2.31, 1.48)	-0.42 (-2.02, 1.18)	-0.38 (-2.42, 1.66)	-0.30 (-2.37, 1.76)	-0.25 (-2.25, 1.75)	-0.22 (-2.27, 1.82)	-0.20 (-2.56, 2.17)	-0.22 (-2.18, 1.75)	-0.15 (-2.20, 1.90)	Sertral			-0.13 (-0.97, 0.71)							
-1.43 (-3.42, 0.56)	-1.19 (-3.62, 1.24)	-0.64 (-2.88, 1.61)	-0.66 (-2.95, 1.64)	-0.49 (-2.23, 1.25)	-0.47 (-2.22, 1.29)	-0.70 (-3.12, 1.72)	-0.57 (-2.99, 1.84)	-0.40 (-2.12, 1.32)	-0.40 (-2.34, 1.54)	-0.36 (-2.27, 1.55)	-0.28 (-2.19, 1.62)	-0.23 (-2.06, 1.60)	-0.21 (-2.08, 1.67)	-0.18 (-2.40, 2.05)	-0.20 (-2.13, 1.73)	-0.13 (-2.02, 1.76)	0.02 (-2.50, 2.53)	Mel								-0.15 (-1.18, 0.89)	
*-1.35 (-2.35, -0.34)	-1.10 (-2.88, 0.68)	-0.55 (-2.05, 0.95)	-0.57 (-2.15, 1.02)	-0.40 (-0.93, 0.12)	-0.38 (-0.85, 0.09)	-0.61 (-2.37, 1.15)	-0.49 (-2.25, 1.27)	-0.31 (-0.71, 0.09)	-0.31 (-1.33, 0.70)	-0.27 (-1.22, 0.68)	-0.20 (-1.13, 0.74)	-0.14 (-0.91, 0.63)	-0.12 (-0.97, 0.74)	-0.09 (-1.58, 1.40)	-0.11 (-1.11, 0.89)	-0.04 (-0.92, 0.83)	0.11 (-1.79, 2.00)	0.09 (-1.63, 1.80)	ocHRT		0.16 (-0.88, 1.20)					*-0.23 (-0.38, -0.08)	
-1.58 (-4.24, 1.08)	-1.34 (-4.32, 1.65)	-0.79 (-3.64, 2.07)	-0.81 (-3.70, 2.09)	-0.64 (-3.12, 1.84)	-0.62 (-3.11, 1.87)	-0.85 (-3.82, 2.13)	-0.72 (-3.70, 2.25)	-0.55 (-3.01, 1.92)	-0.55 (-2.79, 1.69)	-0.51 (-3.09, 2.07)	-0.43 (-3.03, 2.16)	-0.38 (-2.92, 2.16)	-0.36 (-2.93, 2.22)	-0.33 (-3.17, 2.51)	-0.35 (-2.87, 2.17)	-0.28 (-2.86, 2.30)	-0.13 (-1.70, 1.44)	-0.15 (-3.11, 2.81)	-0.24 (-2.70, 2.22)	EdSert							
*-1.41 (-2.59, -0.23)	-1.16 (-2.99, 0.67)	-0.61 (-2.19, 0.97)	-0.63 (-2.27, 1.01)	-0.46 (-1.13, 0.21)	-0.44 (-1.16, 0.28)	-0.67 (-2.49, 1.15)	-0.55 (-2.36, 1.27)	-0.37 (-1.01, 0.27)	-0.37 (-1.48, 0.73)	-0.33 (-1.38, 0.71)	-0.26 (-1.29, 0.78)	-0.20 (-1.09, 0.68)	-0.18 (-1.16, 0.80)	-0.15 (-1.70, 1.40)	-0.17 (-1.26, 0.92)	-0.10 (-1.11, 0.90)	0.05 (-1.90, 1.99)	0.03 (-1.74, 1.80)	-0.06 (-0.67, 0.55)	0.18 (-2.32, 2.68)	Ral					-0.13 (-0.41, 0.15)	
-1.58 (-3.31, 0.14)	-1.33 (-3.55, 0.89)	-0.78 (-2.80, 1.23)	-0.80 (-2.87, 1.27)	-0.64 (-2.07, 0.79)	-0.62 (-2.06, 0.83)	-0.84 (-3.05, 1.36)	-0.72 (-2.92, 1.48)	-0.54 (-1.95, 0.86)	-0.55 (-2.21, 1.12)	-0.51 (-2.14, 1.12)	-0.43 (-2.05, 1.19)	-0.38 (-1.91, 1.16)	-0.35 (-1.94, 1.24)	-0.33 (-2.32, 1.67)	-0.34 (-2.00, 1.31)	-0.28 (-1.88, 1.33)	-0.13 (-2.44, 2.18)	-0.15 (-2.31, 2.02)	-0.24 (-1.63, 1.16)	0.00 (-2.79, 2.80)	-0.17 (-1.64, 1.29)	VEC				0.00 (-0.28, 0.28)	
-1.62 (-3.43, 0.19)	-1.37 (-3.66, 0.91)	-0.82 (-2.91, 1.26)	-0.84 (-2.98, 1.30)	-0.68 (-2.21, 0.86)	-0.66 (-2.20, 0.89)	-0.88 (-3.16, 1.39)	-0.76 (-3.03, 1.51)	-0.58 (-2.09, 0.93)	-0.59 (-2.34, 1.17)	-0.55 (-2.27, 1.18)	-0.47 (-2.18, 1.24)	-0.42 (-2.04, 1.21)	-0.39 (-2.08,1.29)	-0.36 (-2.43, 1.70)	-0.38 (-2.13, 1.36)	-0.32 (-2.01, 1.38)	-0.17 (-2.54, 2.21)	-0.19 (-2.42, 2.05)	-0.27 (-1.78, 1.23)	-0.04 (-2.88, 2.81)	-0.21 (-1.78, 1.35)	-0.04 (-2.04, 1.96)	scEst			0.04 (-0.58, 0.66)	
-1.60 (-3.40, 0.20)	-1.35 (-3.63, 0.93)	-0.80 (-2.88, 1.27)	-0.82 (-2.95, 1.31)	-0.65 (-2.17, 0.86)	-0.63 (-2.17, 0.90)	-0.86 (-3.12, 1.40)	-0.74 (-3.00, 1.52)	-0.56 (-2.05, 0.93)	-0.56 (-2.31, 1.18)	-0.52 (-2.23, 1.18)	-0.45 (-2.15, 1.25)	-0.39 (-2.01, 1.22)	-0.37 (-2.04, 1.30)	-0.34 (-1.79, 1.10)	-0.36 (-2.09, 1.37)	-0.29 (-1.98, 1.39)	-0.14 (-2.51, 2.22)	-0.16 (-2.39, 2.06)	-0.25 (-1.74, 1.24)	-0.01 (-2.85, 2.83)	-0.19 (-1.74, 1.36)	-0.02 (-2.01, 1.97)	0.02 (-2.04, 2.09)	scHRT		0.02 (-0.56, 0.60)	
*-1.66 (-3.12, -0.20)	-1.41 (-3.43, 0.61)	-0.86 (-2.65, 0.93)	-0.88 (-2.73, 0.97)	-0.72 (-1.81, 0.38)	-0.70 (-1.81, 0.42)	-0.92 (-2.93, 1.08)	-0.80 (-2.80, 1.20)	-0.62 (-1.68, 0.44)	-0.63 (-2.01, 0.76)	-0.59 (-1.93, 0.76)	-0.51 (-1.85, 0.83)	-0.46 (-1.68, 0.77)	-0.43 (-1.73, 0.86)	-0.40 (-2.17, 1.36)	-0.42 (-1.80, 0.95)	-0.36 (-1.67, 0.96)	-0.21 (-2.32, 1.91)	-0.23 (-2.18, 1.73)	-0.31 (-1.37, 0.74)	-0.08 (-2.71, 2.56)	-0.25 (-1.39, 0.88)	-0.08 (-1.77, 1.61)	-0.04 (-1.82, 1.74)	-0.06 (-1.82, 1.70)	maHRT	0.18 (-0.13, 0.49)	
*-1.58 (-2.64, -0.52)	-1.33 (-3.09, 0.42)	-0.78 (-2.27, 0.70)	-0.80 (-2.36, 0.76)	*-0.64 (-1.08, -0.20)	*-0.62 (-1.11, -0.12)	-0.84 (-2.58, 0.89)	-0.72 (-2.46, 1.01)	*-0.54 (-0.90, -0.19)	-0.55 (-1.51, 0.41)	-0.51 (-1.40, 0.39)	-0.43 (-1.32, 0.45)	-0.38 (-1.08, 0.33)	-0.35 (-1.18, 0.47)	-0.33 (-1.78, 1.13)	-0.34 (-1.29, 0.60)	-0.28 (-1.13, 0.58)	-0.13 (-1.99, 1.74)	-0.15 (-1.83, 1.54)	-0.24 (-0.57, 0.09)	0.00 (-2.44, 2.44)	-0.17 (-0.71, 0.37)	0.00 (-1.36, 1.36)	0.04 (-1.43, 1.51)	0.02 (-1.43, 1.47)	0.08 (-0.92, 1.08)	Pla	
-2.03 (-4.13, 0.06)	-1.78 (-4.28, 0.71)	-1.24 (-3.57, 1.10)	-1.25 (-3.64, 1.13)	-1.09 (-2.95, 0.77)	-1.07 (-2.94, 0.81)	-1.30 (-3.78, 1.18)	-1.17 (-3.65, 1.31)	-0.99 (-2.83, 0.85)	-1.00 (-2.53, 0.53)	-0.96 (-2.94, 1.03)	-0.88 (-2.89, 1.13)	-0.83 (-2.77, 1.11)	-0.80 (-2.79, 1.18)	-0.78 (-3.10, 1.54)	-0.80 (-2.71, 1.11)	-0.73 (-2.72, 1.27)	-0.58 (-2.79, 1.63)	-0.60 (-3.07, 1.87)	-0.69 (-2.52, 1.15)	-0.45 (-3.16, 2.27)	-0.62 (-2.51, 1.26)	-0.45 (-2.71, 1.81)	-0.41 (-2.74, 1.92)	-0.43 (-2.75, 1.88)	-0.37 (-2.44, 1.69)	-0.45 (-2.26, 1.36)	TlFluo

Pairwise (upper-right portion) and network (lower-left portion) meta-analysis results are presented as estimate effect sizes for the outcome of improvement of depressive symptoms in menopausal women. Interventions are reported in order of mean ranking of depressive symptoms improvement, and outcomes are expressed as standardised mean difference (SMD) [95% confidence intervals (CIs)]. For pairwise meta-analyses, SMD of less than 0 indicates that the treatment specified in the row got more improvement than that specified in the column. For network meta-analysis (NMA), SMD of less than 0 indicates that the treatment specified in the column got more improvement than that specified in the row. Bold figures marked with * indicate statistical significance.

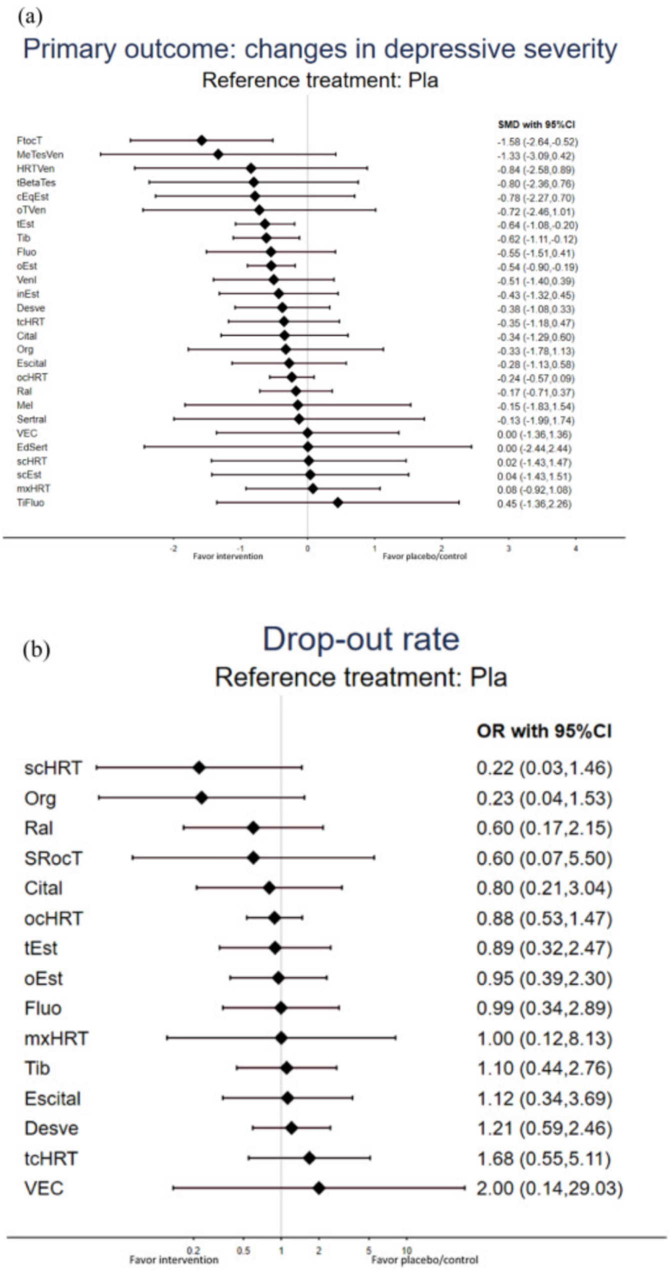

Fig. 3. ***A:*** **Forest plot of current network meta-analysis of changes in depressive severity;** *Fig. 3B:* **Forest plot of current network meta-analysis of acceptability in aspect of dropout rate**.

In Fig. 3A, effect size (presented as SMD) < 0 meant better improvement in depressive symptoms through specified interventions than that in placebo/control groups; whereas, in Fig. 3B, effect size (presented as OR) < 1 meant better acceptability in aspect of dropout rate through specified interventions than that in placebo/control groups.

Abbreviation: 95%CI=95% confidence interval. cEqEst=conjugated equine estrogens + testosterone cream. Cital=citalopram. Desve=desvenlafaxine. EdSert=transdermal estradiol + sertraline. Escital=escitalopram. Fluo=fluoxetine. FtocT=fluoxetine + oral combined therapy (conjugated equine estrogens + cyclic medroxyprogesterone acetate). HRTVen=venlafaxine + estrogen + medroxyprogesterone acetate + methyltestosterone. inEst=intranasal estradiol. Mel=melatonin. MeTesVen=venlafaxine + methyltestosterone. mxHRT=mixed hormone replacement therapy (estrogen in women with hysterectomy. estrogen + progesterone in those without hysterectomy). ocHRT=oral combined hormone replacement therapy (estrogen + progesterone). oEst=oral estrogen. OR=odds ratio; Org=estradiol implants and testosterone or (norethisterone for those without hysterectomy). oTVen=venlafaxine + oral combined therapy (conjugated equine estrogens + cyclic medroxyprogesterone acetate). Ral=raloxifene. scEst=subcutanenous implant of estradiol. scHRT=subcutaneous combined hormone replacement therapy (estradiol implants + norethisterone). Sertral=sertraline. SMD=standardised mean difference. tBetaTes=transdermal 17β-estradiol + testosterone cream. tcHRT=transdermal combined hormone replacement therapy (estrogen + progesterone). tEst=transdermal estradiol. Tib=tibolone. TiFluo=fluoxetine + tibolone. VEC=vaginal estrogen cream. Venl=venlafaxine.

3.1.1. Subgroup of baseline with definite diagnosis of depression

Of the 18 articles involving eleven individual treatment arms and one placebo arm being analyzed in the current subgroup NMA for participants with a definite diagnosis of depression (Appendix 6.1), one was removed because it could not be linked to the other treatment arms within the network structure (Dias et al., 2006). Only fluoxetine *plus* oral combined hormone therapy (conjugated equine estrogens + cyclic medroxyprogesterone acetate) [SMD=-1.87 (95%CIs=-3.13 to -0.61)] was associated with significantly better therapeutic effect than that in placebos/controls in the subgroup of participants with a definite diagnosis of depression (Appendix 7.1, and Appendix 8.1).

3.1.2. Subgroup of baseline without definite diagnosis of depression

For subgroup analysis of participants excluding the diagnosis of depression, 20 articles with fourteen individual treatment arms and one placebo arm were studied in the current subgroup NMA (Appendix 6.2). None of the investigated interventions showed significantly better improvement in depressive symptoms than that in placebo/control groups (Appendix 7.2 and Appendix 8.2).

3.1.3. Subgroup of post-menopausal period only

Of the 26 articles with a total of 13 individual treatment arms and one placebo arm included in the current subgroup NMA for women after menopausal transition (i.e., amenorrhea >1 year) (Appendix 6.3), two were removed because it could not be linked to the other treatment arms within the network structure (Berlanga et al., 2003; Rasgon et al., 2007). Only the improvement associated with the venlafaxine-methyltestosterone regimen achieved statistical significance [SMD=-1.57 (95%CIs=-2.90 to -0.24)] (Appendix 7.3 and Appendix 8.3).

3.1.4. Subgroup of peri-menopausal period only

Of the 6 articles with a total of 5 individual treatment arms and one placebo arm included in the current subgroup NMA for peri-menopausal women (Appendix 6.4), Only the improvement associated with the transdermal estradiol (tEst) achieved statistical significance [SMD=-0.95 (95%CIs=-1.77 to -0.14)] (Appendix 7.4 and Appendix 8.4).

3.2. Primary outcome: (2) acceptability in aspect of dropout rate

With respect to dropout rate, none of the 36 trials with 15 individual treatment arms and one placebo arm showed higher dropout rates than those in placebo/control groups (Table 1B, Figs. 2**B**, and 3**B**). Ranking of the acceptability of different interventions with surface under the cumulative ranking curve (SURCA) for overall analyses are provided in Appendix 9.

Table 1B. League table of acceptability in aspect of drop-out rate in menopausal women receiving hormonal/pharmacologic treatments.

scHRT	0.95 (0.17, 5.26)									***0.22 (0.05, 0.99)**					
0.95 (0.12, 7.47)	Org									0.23 (0.05, 1.04)					
0.37 (0.04, 3.61)	0.39 (0.04, 3.78)	Ral				0.32 (0.01, 8.27)				0.67 (0.18, 2.44)					
0.37 (0.02, 6.77)	0.39 (0.02, 7.11)	1.00 (0.08, 12.87)	SRocT		0.68 (0.11, 4.22)										
0.28 (0.03, 2.77)	0.29 (0.03, 2.91)	0.75 (0.12, 4.72)	0.75 (0.06, 9.90)	Cital					0.84 (0.37, 1.90)	0.77 (0.34, 1.74)					
0.25 (0.04, 1.76)	0.26 (0.04, 1.85)	0.68 (0.17, 2.67)	0.68 (0.08, 5.86)	0.91 (0.22, 3.78)	ocHRT					0.84 (0.50, 1.41)	1.00 (0.20, 4.96)	0.92 (0.20, 4.34)			
0.25 (0.03, 2.11)	0.26 (0.03, 2.22)	0.67 (0.14, 3.22)	0.67 (0.06, 7.70)	0.90 (0.17, 4.81)	0.99 (0.32, 3.09)	tEst				0.86 (0.31, 2.38)					
0.23 (0.03, 1.87)	0.24 (0.03, 1.96)	0.63 (0.13, 2.98)	0.63 (0.06, 6.73)	0.85 (0.17, 4.18)	0.93 (0.35, 2.46)	0.94 (0.24, 3.62)	oEst			0.86 (0.59, 1.24)		1.23 (0.42, 3.64)			
0.22 (0.01, 3.71)	0.23 (0.01, 3.89)	0.60 (0.05, 6.97)	0.60 (0.03, 12.65)	0.80 (0.07, 9.62)	0.88 (0.10, 7.63)	0.89 (0.09, 9.18)	0.95 (0.10, 9.24)	mxHRT		1.00 (0.17, 5.77)					
0.22 (0.03, 1.94)	0.23 (0.03, 2.04)	0.60 (0.11, 3.19)	0.60 (0.05, 7.06)	0.81 (0.21, 3.07)	0.89 (0.27, 2.90)	0.90 (0.21, 3.93)	0.96 (0.24, 3.82)	1.01 (0.10, 10.56)	Fluo	0.98 (0.50, 1.90)					
0.22 (0.03, 1.46)	0.23 (0.04, 1.53)	0.60 (0.17, 2.15)	0.60 (0.07, 5.50)	0.80 (0.21, 3.04)	0.88 (0.53, 1.47)	0.89 (0.32, 2.47)	0.95 (0.39, 2.30)	1.00 (0.12, 8.13)	0.99 (0.34, 2.89)	Pla		***0.36 (0.13, 0.97)**	0.81 (0.50, 1.34)	0.50 (0.04, 5.62)	0.63 (0.33, 1.22)
0.20 (0.02, 1.83)	0.21 (0.02, 1.92)	0.53 (0.09, 3.04)	0.53 (0.04, 6.36)	0.72 (0.12, 4.27)	0.79 (0.23, 2.66)	0.79 (0.17, 3.81)	0.85 (0.19, 3.68)	0.89 (0.08, 9.90)	0.88 (0.18, 4.37)	0.89 (0.27, 2.92)	Escital		1.03 (0.61, 1.72)		
0.20 (0.02, 1.63)	0.21 (0.03, 1.71)	0.54 (0.11, 2.57)	0.54 (0.05, 5.61)	0.73 (0.14, 3.67)	0.80 (0.33, 1.96)	0.81 (0.20, 3.20)	0.86 (0.30, 2.45)	0.91 (0.09, 8.92)	0.90 (0.22, 3.68)	0.91 (0.36, 2.26)	1.02 (0.24, 4.38)	Tib			
0.18 (0.02, 1.37)	0.19 (0.03, 1.44)	0.50 (0.11, 2.14)	0.50 (0.05, 5.04)	0.67 (0.15, 3.01)	0.73 (0.31, 1.71)	0.74 (0.21, 2.56)	0.79 (0.25, 2.44)	0.83 (0.09, 7.57)	0.82 (0.23, 2.97)	0.83 (0.41, 1.69)	0.93 (0.31, 2.77)	0.91 (0.29, 2.88)	Desve		
0.11 (0.00, 2.92)	0.12 (0.00, 3.06)	0.30 (0.02, 5.81)	0.30 (0.01, 9.67)	0.40 (0.02, 7.98)	0.44 (0.03, 6.73)	0.45 (0.03, 7.82)	0.48 (0.03, 7.95)	0.50 (0.02, 14.95)	0.50 (0.03, 8.85)	0.50 (0.03, 7.26)	0.56 (0.03, 10.50)	0.55 (0.03, 9.33)	0.60 (0.04, 9.62)	VEC	
0.13 (0.01, 1.18)	0.14 (0.02, 1.24)	0.36 (0.07, 1.95)	0.36 (0.03, 4.28)	0.48 (0.08, 2.72)	0.53 (0.15, 1.80)	0.53 (0.12, 2.41)	0.57 (0.14, 2.35)	0.60 (0.06, 6.40)	0.59 (0.13, 2.77)	0.60 (0.20, 1.82)	0.67 (0.13, 3.43)	0.66 (0.15, 2.80)	0.72 (0.19, 2.70)	1.19 (0.07, 21.64)	tcHRT

Pairwise (upper-right portion) and network (lower-left portion) meta-analysis results are presented as estimate effect sizes for the outcome of tolerability in aspect of drop-out rate. Interventions are reported in order of mean ranking of tolerability, and outcomes are expressed as odds ratio (OR) (95% confidence intervals). For the pairwise meta-analyses, OR of less than 1 indicate that the treatment specified in the row got more tolerability than that specified in the column. For the network meta-analysis (NMA), OR of less than 1 indicate that the treatment specified in the column got more tolerability than that specified in the row. Bold figures marked with * indicate statistical significance.

Abbreviation: cEqEst=conjugated equine estrogens + testosterone cream. Cital=citalopram. Desve=desvenlafaxine. EdSert=transdermal estradiol + sertraline. Escital=escitalopram. Fluo=fluoxetine. FtocT=fluoxetine + oral combined therapy (conjugated equine estrogens + cyclic medroxyprogesterone acetate). HRTVen=venlafaxine + estrogen + medroxyprogesterone acetate + methyltestosterone. inEst=intranasal estradiol. Mel=melatonin. MeTesVen=venlafaxine + methyltestosterone. mxHRT=mixed hormone replacement therapy (estrogen in women with hysterectomy; estrogen + progesterone in those without hysterectomy). ocHRT=oral combined hormone replacement therapy (estrogen + progesterone). oEst=oral estrogen. Org=estradiol implants and testosterone or (norethisterone for those without hysterectomy). oTVen=venlafaxine + oral combined therapy (conjugated equine estrogens + cyclic medroxyprogesterone acetate). Ral=raloxifene. scEst=subcutanenous implant of estradiol. scHRT=subcutaneous combined hormone replacement therapy (estradiol implants + norethisterone). Sertral=sertraline. tBetaTes=transdermal 17β-estradiol + testosterone cream. tcHRT=transdermal combined hormone replacement therapy (estrogen + progesterone). tEst=transdermal estradiol. Tib=tibolone. TiFluo=fluoxetine + tibolone. VEC=vaginal estrogen cream. Venl=venlafaxine.

3.3. Safety profile: rate of discontinuation due to adverse events

Regarding the rate of discontinuation due to adverse events, we investigated 23 articles involving a total of 13 individual treatment arms and one placebo arm (Appendix 6.5). One article was removed because it could not be linked to the other treatment arms within the network structure (Dias et al., 2006). Only the desvenlafaxine-only [OR=2.16 (95%CIs=1.13 to 4.15)], oral combined hormone replacement therapy (estrogen *plus* progesterone) [OR=6.81 (95%CIs=1.71 to 27.18)], and tibolone-only [OR=7.86 (95%CIs=1.48 to 41.68)] groups correlated with significantly higher rates of study discontinuation due to adverse events than those in placebo/control groups (Appendix 7.5 and Appendix 8.5). Ranking of safety profile of different interventions with surface under the cumulative ranking curve (SURCA) for overall analyses are provided in Appendix 9.

3.3. Risk of bias, funnel plot, publication bias, and GRADE ratings

Focusing on the risk of bias, we found that 55.3% (303/525), 31.0% (163/525), and 11.2% (59/525) of the included studies had an overall low, unclear, and high risk of bias, respectively. The vague reporting of randomization procedures or blindness of the studies further contributed to the risk of bias. In addition, as many as 22 out of the 75 trials were sponsored by pharmaceutical companies, thereby predisposing the reports to potential bias in favor of the related pharmacologic treatments (Appendix 10).

Funnel plots of publication bias across the included studies showed general symmetry, and the results of Egger's test indicated no significant asymmetry that may suggest publication bias among the articles included in the present NMA (Appendix 11). In general, inconsistency concerning either local inconsistency, as assessed using the loop-specific approach and the node-splitting method, or global inconsistency, as determined using the design-by-treatment method, was not demonstrated in the present NMA except for situations mentioned below. There were significant inconsistencies detected in overall subjects or the subgroup of menopausal women without definite baseline depression by design-by-treatment method (p=0.0003 and p <0.0001, respectively). Besides, statistically significant inconsistencies were also noted when comparing tEst with placebo as well as oral estrogen (oEst) with intranasal estradiol (inEst) by the side-splitting inconsistency model in the subgroup of postmenopausal women (p=0.012 and 0.032, respectively). Appendix 12 and Appendix 13 depicted the inconsistency in the current NMA. Finally, Appendix 14 shows the overall quality of evidence among the current NMA according to the GRADE rating system (Puhan et al., 2014). In brief, the overall quality of evidence ranged between low to medium. Appendix 15 shows the overall heterogeneity among the investigated treatments.

4. Discussion

Based on 75 RCTs recruiting a total of 19,048 patients, the current NMA found that only a few of the investigated interventions were associated with significantly better therapeutic effects against depressive symptoms than those in placebo/control groups in a mixed sample of peri- and post-menopausal women, including fluoxetine *plus* oral combined estrogen-progesterone therapy, transdermal estradiol, tibolone only, and oral estrogen therapy**.** However, the superiority of hormone replacement therapies failed to persist in women after menopausal transition (amenorrhea >1 year). The findings were partially consistent with those of previous RCTs (El-Hage et al., 2007; Polisseni et al., 2013) and meta-analyses (Freeman et al., 2004) as well as the recommendations of clinical guidelines (MacQueen et al., 2016). Among all of the 27 active treatment arms, regimens that combined hormone treatment (i.e., estrogen and progesterone) with antidepressants were ranked higher than those of hormone replacement therapy or antidepressants alone despite the lack of statistically significant differences. Besides, in the current NMA, there was no evidence showing that tolerability as reflected by dropout rate was lower in regimens with combined hormone replacement therapy and antidepressants than that in the placebo/control groups. Taking into account the synergistic serotonergic effect of hormone replacement therapy, especially estrogen, against depressive symptoms through enhancing serotonin synthesis and/or reducing serotonin reuptake (Shors and Leuner, 2003), it appears that estrogen may provide additional therapeutic benefit when combined with selective serotonin reuptake inhibitors (SSRIs) in this setting. Nevertheless, such evidence came only from two studies (Liu et al., 2004; Yu et al., 2004); therefore, the results require judicious interpretation.

When focusing on the regimens achieving statistical significance, both transdermal and oral estrogen demonstrated similar effects with moderate effect sizes that were significantly better than those in the placebo/control groups. However, compared with that in a mixed sample of women during and after menopausal transition, women after menopausal transition (i.e., amenorrhea >1 year) showed a notable drop in efficacy of hormone therapies against depressive symptoms to such a degree that it was statistically comparable to that of the placebo. Although venlafaxine-methyltestosterone was the only regimen significantly better than placebo in this subgroup of participants, the result came from only one study with a very high dropout rate (44.45%) in that treatment arm (Dias et al., 2006) and needs to be validated by further studies. We were unable to conduct subgroup analysis for women during menopause transition (i.e., perimenopause) due to a lack of available data for this subgroup. However, there was a trend of a decrease in efficacy of hormone therapies in women after menopausal transition, suggesting that hormone therapies may not be beneficial in the treatment of depressive symptoms for women after menopausal transition. Moreover, the efficacy of hormone therapies was not significantly better than that of placebo in patients with a definite diagnosis of depression.

Among all the antidepressant-related interventions, only the therapeutic benefit of fluoxetine *plus* oral combined estrogen-progesterone treatment achieved statistical significance. Despite the seemingly superior treatment effect of the other antidepressant-related interventions over that of placebo/control, their therapeutic benefits failed to reach statistical significance. The results may be attributable to the wide variation in treatment effects (i.e., the wide 95%CIs) due to the small number of RCTs investigating the efficacy of regimens with antidepressants in menopausal women as stated in the Canadian guideline (MacQueen et al., 2016). Therefore, further large-scale RCTs on the therapeutic benefits of antidepressant-related interventions in menopausal women are warranted.

Regarding the treatment effect of androgen against depression, the current study provided evidence that the androgen-associated regimens, namely tibolone, were superior to placebo/control in relieving depressive symptoms in menopausal women. The finding is similar to that of a previous meta-analysis, which also showed that androgen, alone or in combination with estrogen, was related to a more promising improvement in depressed mood compared with that in the placebo/control groups (Zweifel and O'Brien, 1997). Previous reports also support the rationale of testosterone supplement for depressive symptom management, especially in women with sexual dysfunction (Rohr, 2002; Sherwin et al., 1985). The current NMA showed that although testosterone supplement was associated with an apparently better improvement in depressive symptoms, the finding failed to achieve statistical significance. Moreover, the results mainly came from RCTs recruiting women with sexual dysfunction (El-Hage et al., 2007; Shifren et al., 2000). Therefore, the therapeutic effects of androgen derivatives against depressive symptoms in menopausal women without sexual dysfunction remain to be elucidated.

The current study has its strengths and weaknesses. It represented the first NMA that provided comparisons of efficacy against depressive symptoms and tolerability between hormone therapies with different administrative routes and a variety of pharmacological interventions, especially antidepressants, in women during and after menopausal transition. We also included as many as 75 RCTs with a large sample size of 19,048 patients. The large databank enabled subgroup analysis for participants with different severities of depression. Nevertheless, several limitations of our study need to be considered for accurate interpretation of the results. First, evidence of the current NMA came from a mixed sample of peri- and post-menopausal women. Although we conducted subgroup analysis for women after menopausal transition, similar analysis was not performed for perimenopausal women due to a lack of eligible studies (only 8 out of 75 RCTs enrolled only perimenopausal women). Given that hormonal changes are different between postmenopausal and perimenopausal stages, the present study may not provide sufficient evidence for the perimenopausal population. Second, the evidence derived from the current NMA may be obscured by the heterogeneous characteristics of the participants (e.g., comorbid diseases, baseline depression severity, placebo-controlled or not, ethnical background, and trial duration), small numbers of trials for some treatment arms, not separating natural from bio-identical hormones as well as cyclic from continuous combined estrogen-progesterone treatments, and heterogeneity in psychopathology assessment tools. Despite our effort to reduce the heterogeneity through subgroup analysis, some effects may still persist. Third, there were statistically significant inconsistencies when a comparison was made between tEst and placebo as well as that between oEst and inEst using the side-splitting inconsistency model in the subgroup of postmenopausal women. Finally, since most of the included trials included women with other existing health problems (e.g., vasomotor symptoms or cardiovascular diseases) for which treatments were given, the possibility that the observed improvements in depressive symptoms were the results of alleviation of potential aggravating factors for depression (e.g., vasomotor symptoms, sleep disturbance) cannot be ruled out. Nevertheless, given that accompanying health problems are common among women during or after menopausal transition, it is still valuable to include those trials to reflect the real-world clinical scenario.

The current NMA provided evidence that hormone therapies with/without antidepressants were more effective against depressive symptoms than placebo/control in menopausal women. However, hormone therapies only maintained their efficacies against depressive symptoms in peri-menopausal women but lost the anti-depressive efficacies in post-menopausal women. Fluoxetine *plus* oral combined estrogen-progesterone treatment showed the highest SMD according to the available evidence, but the sample size was small and the finding needs to be interpreted with caution. As some interventions were included in only a few RCTs and most trials recruited both peri- and post-menopausal women, caution should be exercised when considering the prescription indications for certain interventions. To confirm the clinical significance of treatment regimens demonstrating statistically significant efficacy in this NMA, large-scale clinical trials targeting specific subgroups are necessary.

Contributors

Conceptualization: Ming-Kung Wu, Kuo-Chuan Hung, Kuan-Pin Su, and Dian-Jeng Li; Data curation: Ping-Tao Tseng, Hsien-Jane Chiu, Mein-Woei Suen, Yu-Kang Tu, Yi-Cheng Wu, and Bing-Syuan Zeng; Formal analysis: Yu-Kang Tu, Yi-Cheng Wu, Ping-Tao Tseng, and Yu-Shian Cheng; Funding acquisition: None; Investigation: Tien-Yu Chen, Brendon Stubbs, Andre F Carvalho, and Marco Solmi; Methodology: Yu-Kang Tu, Yi-Cheng Wu, Ping-Tao Tseng, and Yu-Shian Cheng; Project administration: Trevor Thompson, Maria Gabriella Caruso, Yutaka J. Matsuoka, Yen-Wen Chen, and Pao-Yen Lin; Resources: Ming-Kung Wu, Kuo-Chuan Hung, Trevor Thompson, Maria Gabriella Caruso, and Yutaka J. Matsuoka; Software: Yu-Kang Tu, Yi-Cheng Wu, Ping-Tao Tseng, and Yu-Shian Cheng; Supervision: Kuan-Pin Su, Cheuk-Kwan Sun, Yu-Shian Cheng, and Yow-Ling Shiue; Validation: Ming-Kung Wu, Kuo-Chuan Hung, Trevor Thompson, Maria Gabriella Caruso, and Yutaka J. Matsuoka; Visualization: Kuo-

Chuan Hung, Kuan-Pin Su, and Dian-Jeng Li; Writing-original draft: Ping-Tao Tseng, Hsien-Jane Chiu, Mein-Woei Suen, and Bing-Syuan Zeng; Writing-review and editing: Cheuk-Kwan Sun, Yu-Shian Cheng, and Yow-Ling Shiue.