Menopause: The Journal of The North American Menopause Society
Vol. 26, No. 8, pp. 850-855
DOI: 10.1097/GME.0000000000001343

Confirmatory factor analysis of the Insomnia Severity Index (ISI) and invariance across race: a pooled analysis of MsFLASH data

Julie L. Otte, PhD, RN,[1] Giorgos Bakoyannis, PhD,[1,2] Kevin L. Rand, PhD,[3] Kristine E. Ensrud, MD, MPH,[4] Katherine A. Guthrie, PhD,[5] Hadine Joffe, MD, MSc,[6] Susan M. McCurry, PhD,[7] Kathrine M. Newton, PhD,[8] and Janet S. Carpenter, PhD, RN, FAAN[1]

Abstract

Objective: Women's sleep at menopause is widely reported to be problematic. The Insomnia Severity Index (ISI) is a commonly used tool for quantifying sleep problems in clinical and research settings, but psychometric properties in postmenopausal women have not been reported. Our study aim was to examine the factor structure of the ISI in a large and diverse sample of midlife women with hot flashes.

Methods: Baseline data were from 899 women enrolled in one of the three clinical trials using similar entry criteria conducted by the Menopause Strategies Finding Lasting Answers to Symptoms and Health research network. We conducted confirmatory factor analyses for the total sample and within strata defined by race/ethnicity (black and white women).

Results: The ISI had two factors in the total sample. The two-factor structure was consistent across black and white women, with the exception of one item ''difficulty falling asleep.''

Conclusions: The ISI in midlife women with hot flashes is composed of two factors that capture dimensions of the insomnia severity and daytime impact. The instrument is a psychometrically sound scale appropriate for use in research and clinical practice to capture the severity and daytime impact of insomnia symptoms in diverse samples of midlife women with hot flashes. An abbreviated screening of two items could be considered to determine if further evaluation is needed of sleep complaints.

Key Words: Insomnia – Instrumentation – Menopause – Race – Sleep.

Extensive research shows that midlife women report poor sleep during various stages of the menopausal transition and postmenopause.[1-6] For example, at least 48% to 50% of women with hot flashes experience problems with nighttime sleep during late menopause and early postmenopause.[3,5-7] Self-reported sleep problems are often characterized as symptoms of insomnia such as the inability to fall asleep, remain asleep, and feel rested during the day.[8-11] Insomnia symptoms are highly associated with daytime fatigue and overall poor quality of life. [10-14] Thus, it is imperative that clinicians and researchers have well-validated measures of insomnia symptoms to clinically diagnose and treat primary and secondary insomnia disorders.[13]

The Insomnia Severity Index (ISI) is widely used to quantify self-reported insomnia symptoms in clinical practice and research. It is a seven-item scale that requires less than 10 minutes to complete and score. Use in clinical practice can, however, be limited due to limited resources, competing demands, and time constraints. Scores indicate no, mild, moderate, or severe insomnia over the past 2 weeks, with

Received November 6, 2018; revised and accepted March 6, 2019.

From the [1]School of Nursing, Indiana University, Indianapolis, IN; [2]Department of Biostatistics, Fairbanks School of Public Health and School of Medicine, Indiana University, Indianapolis, IN; [3]Department of Psychology, School of Science, Indiana University-Purdue University Indianapolis, IN; [4]Medicine and Epidemiology and Community Health, University of Minnesota and Minneapolis VA Health Care System, Minneapolis, MN; [5]MsFLASH Data Coordinating Center, Fred Hutchinson Cancer Research Center, Seattle, WA; [6]Connors Center for Women's Health and Gender Biology, Department of Psychiatry, Brigham and Women's Hospital, Harvard Medical School, Boston, MA; [7]Department of Psychosocial and Community Health, School of Nursing, University of Washington, Seattle, WA; and [8]Kaiser Permanente Washington Health Research Institute, Seattle, WA.

Funding/support: The MsFLASH studies were supported by a cooperative agreement issued by the National Institute of Aging (NIA), in collaboration with the Eunice Kennedy Shriver National Institute of Child Health and Development (NICHD), the National Center for Complementary and Alternative Medicine (NCCAM), and the Office of Research and Women's Health (ORWH), and NIA grants U01AG032659, U01AG032669, U01AG032682, U01AG032699, and U01AG032700. In Indiana, this project was supported by the Indiana Clinical and Translational Sciences Institute, funded in part by grant UL1 RR025761 from the National Institutes of Health, National Center for Research Resources, Clinical, and Translational Sciences Award. The content of the manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Financial disclosure/conflicts of interest: JSC and KLR are funded under an Independent Grant for Learning and Change from Pfizer, Inc (2017-2018). JSC consults for QUE Oncology and Astellas Pharma, Inc. HJ receives grant support from NIH, Merck, Pfizer, Pfizer, and QUE Oncology and consults for KaNDy and Sojournix; and her spouse is a Merck employee. The rest of the authors have no conflicts to report.

Address correspondence to: Julie L. Otte, PhD, RN, OCN, 600 Barnhill Drive, NU W401, Indianapolis, IN 46202.
E-mail: jlelam@iu.edu

higher scores indicating more severe insomnia.[15] The scale has been used in adults[16,17] and teenagers,[18] and has been translated to over 50 languages such as Chinese,[19] Spanish,[20] and Arabic.[21]

The research team that created the ISI did not conduct an exploratory factor analysis in the initial psychometric evaluation studies that limited the depth of understanding regarding the theoretical structure of insomnia. There are only a small number of studies that have examined ISI psychometrics and/or factor structure with the original English-language version of the scale, and none of these studies were specific to postmenopausal women.[15,18,22-24] Factor structures reported contained models of two- or three-factor structures. Two-factor models included a factor of insomnia symptoms and factor of the impact of insomnia symptoms.[18,22,23] Three-factor models include a factor structure of severity of symptoms, impact of poor sleep, and satisfaction with sleep.[15,24] Because the factor structures of validated questionnaires can vary by population (eg, age, sex, race), the purpose of this study was to examine the factor structure of the ISI using data from midlife women with hot flashes enrolled in the Menopause Strategies Finding Lasting Answers to Symptoms and Health (MsFLASH) menopausal symptom treatment trials to determine the variable dimensions on insomnia in this population. We sought to evaluate the factor structure in the overall sample as well as whether it varied between women according to race (white vs black). Although women were not selected for the first three MsFLASH trials based on a priori self-reports of sleep problems, 70% of the sample reported sleep problems at baseline, further supporting the relationship between hot flashes and sleep complaints. Based on the factor structure we will examine if an abbreviated scale is possible to use in clinical practice when assessing insomnia symptoms in this population.

METHODS

Design

We collated baseline data from 899 women enrolled in the MsFLASH 01, 02, and 03 trials.[25-32] Common inclusion criteria across studies were 14 or more hot flashes per week, peri- or postmenopausal, age 40 to 62, and in good general health. Women were recruited from July 2009 through October 2012. Recruitment methods included mass mailings using purchased mailing lists and health plan enrollment files. Interested women maintained a daily hot flash diary for 2 weeks; if they continued to be eligible, they took part in two baseline assessments scheduled 1 week apart. All trials were reviewed by institutional review boards, and informed consent was obtained for all participants. The details of these trials have been extensively described in other reports.[25-32]

Measures

Women reported basic demographic information at baseline. These data included race, ethnicity, age, education, and smoking history. Study staff measured women's height and weight in the clinics for calculations of body mass index. Women completed the ISI as part of a larger packet of baseline questionnaires.

Vasomotor frequency, severity, and bother were assessed using daily diaries. Participants completed the paper-based diaries upon awakening and before falling asleep. Severity of daytime and nighttime hot flashes was rated from 0 (mild) to 2 (severe). Bother of daytime and nighttime hot flashes was rated from 0 (not at all) to 3 (a lot).

The seven-item ISI evaluates perceived severity of insomnia over the past 2 weeks. Items are scored using a Likert-type scale with five response options (0 = not at all to 4 = very much). Total scores are obtained by summing items, with total scores ranging from 0 to 28. Higher scores indicate greater perception of insomnia severity. Reliability and validity have been established in healthy individuals ($n = 145$, Cronbach's alpha = 0.74) and those with chronic illnesses ($n = 1{,}634$, Cronbach's alpha = 0.90).[15,23]

Statistical analysis

For demographic variables, we used absolute and relative frequencies to describe categorical variable and median and interquartile range for continuous variables. To compare demographic variables between black and white participants, we used Pearson's chi-squared test for categorical variables and the Mann–Whitney test for continuous variables.

We performed a series of confirmatory factor analyses based on unweighted least squares (ULS) estimation of the model parameters. We chose ULS estimation because the items of the ISI scale were ordinal variables, and thus the usual maximal likelihood estimation was not applicable. Other popular approaches to deal with ordinal items are weighted least squares (WLS) and diagonally weighted least squares (DWLS) estimation methods. ULS performance has, however, been shown to be superior in finite samples to that of WLS and DWLS estimation.[33] Good model fit was defined as values of both Tucker–Lewis fit Index (TLI) and Comparative Fit Index (CFI) greater than 0.95, a Root Mean Square Error of Approximation (RMSEA) value smaller than 0.06, and a Standardized Root Mean Square Residual (SRMR) value smaller than 0.08.[34]

The assessment of measurement invariance between black and white women was based on a sequence of nested model comparisons using scale items. After the recommendation by Cheung and Rensvold (2002),[35] we used the CFI measure for comparing nested models between black and white women. Unlike Cheung and Rensvold, we, however, performed bootstrap hypothesis tests to statistically evaluate the potential differences between the CFI measures for pairs of nested models that were considered in the invariance analysis.[36] This choice was crucial because, to our knowledge, the asymptotic distribution of the CFI estimator has not been studied, and therefore there is no available parametric approach for hypothesis testing. To compare differences in the intercepts of individual items between black and white women, we performed formal Wald tests for the differences between each item's intercepts, using the appropriate variance components

TABLE 1. *Sample demographics*

	Overall ($N = 899$)	Group of interest		
		Black ($N = 288$)	White ($N = 530$)	
	N (%)	*N* (%)	*N* (%)	*P*
Age at screening				<0.001
<50	73 (8.1)	39 (13.5)	29 (5.5)	
50-54	404 (44.9)	139 (48.3)	234 (44.2)	
55-59	319 (35.5)	88 (30.6)	199 (37.5)	
≥60	103 (11.5)	22 (7.6)	68 (12.8)	
Education				<0.001
≤High school diploma or general education degree	114 (12.7)	66 (22.9)	37 (7.0)	
Post-high school	310 (34.6)	128 (44.4)	150 (28.4)	
College graduate	473 (52.7)	94 (32.6)	341 (64.6)	
Smoking				<0.001
Never	507 (56.6)	136 (47.2)	325 (61.7)	
Past	255 (28.5)	66 (22.9)	161 (30.6)	
Current	134 (15.0)	86 (29.9)	41 (7.8)	
BMI, kg/m^2				<0.001
<25	295 (33.1)	45 (15.7)	219 (41.8)	
25-29.9	323 (36.3)	96 (33.6)	198 (37.8)	
≥30	273 (30.6)	145 (50.7)	107 (20.4)	
ISI sleep score				0.006
None (0-7)	251 (28.4)	84 (30.2)	140 (26.7)	
Subthreshold (8-14)	358 (40.5)	101 (36.3)	240 (45.7)	
Moderate (15-21)	235 (26.6)	73 (26.3)	130 (24.8)	
Severe (22-28)	40 (4.5)	20 (7.2)	15 (2.9)	
	Median (IQR)	Median (IQR)	Median (IQR)	*P*
Age at screening	54.0 (52.0-57.0)	53.0 (51.0-56.0)	55.0 (52.0-58.0)	<0.001
BMI, kg/m^2	27.1 (23.7-31.1)	30.1 (26.8-33.4)	25.9 (22.9-28.9)	<0.001
VMS frequency	7.1 (5.2-10.1)	7.4 (5.4-10.5)	7.1 (5.0-9.9)	0.032
VMS severity	1.0 (0.7-1.3)	1.2 (0.9-1.6)	0.9 (0.7-1.2)	<0.001
VMS bother	2.0 (1.6-2.3)	2.2 (1.9-2.6)	1.9 (1.5-2.2)	<0.001
ISI sleep score	11.0 (7.0-16.0)	12.0 (6.0-16.0)	11.0 (7.0-15.0)	0.468

BMI, body mass index; ISI, Insomnia Severity Index; IQR, interquartile range; VMS, vasomotor symptoms (hot flashes, night sweats); VMS severity rated 0 mild to 2 severe; VMS bother rated 0 not at all to 3 a lot.

of the estimated variance covariance matrix for the fitted model for the strong/scalar invariance and applied the Bonferroni correction for type I error inflation due to multiple comparisons.

All statistical analyses except for the confirmatory factor analysis were performed using STATA 14 for Windows (STATACorp LLC, College Station, TX). The free statistical software package called R was used to perform the confirmatory factor analysis and the evaluation of construct invariance.[37]

RESULTS

In total, 899 participants from 3 MsFLASH studies were included in the analysis. Of them, 530 (59.0%) were white, 288 (32.0%) were black, 78 (8.7%) were American Indian, Asia/Pacific Islanders, or other, while ethnicity information for 3 (0.3%) participants was missing. For this overall sample, most participants were educated beyond high school, were never smokers, were overweight, and had subthreshold or higher insomnia (Table 1). When comparing only the subgroups of white and black women, black women were younger, had higher BMI, higher VMS symptomatology, and higher ISI severity (Table 1).

Results of the factor analysis in the overall sample are shown in Table 2 and Figure 1. Fit indices shown in Table 2 indicate that a one-factor model did not provide a good fit according to the RMSEA. In contrast, a two-factor model met all the criteria of a good fit. The two-factor model (severity and impact of symptoms) is depicted along with the estimated factor loadings, correlation between the factors, and the residual standard errors in Figure 1.

Results of the ISI measurement invariance are shown in Table 3. The CFI for configural invariance was excellent for a two-group, two-factor model with all model parameters allowed to freely vary between groups. When the factor loadings were constrained to be equal for black and white

TABLE 2. *Fit indices of different Insomnia Severity Index factor models*

Model	CFI	TLI	RMSEA	RMSEA 90% CI	SRMR
One-factor	0.990	0.986	0.076	(0.061-0.091)	0.050
Two-factor	0.997	0.994	0.047	(0.030-0.064)	0.033

CFI, Comparative Fit Index; TLI, Tucker–Lewis fit Index; RMSEA, Root Mean Square Error of Approximation; SRMR, Standardized Root Mean Square Residual.

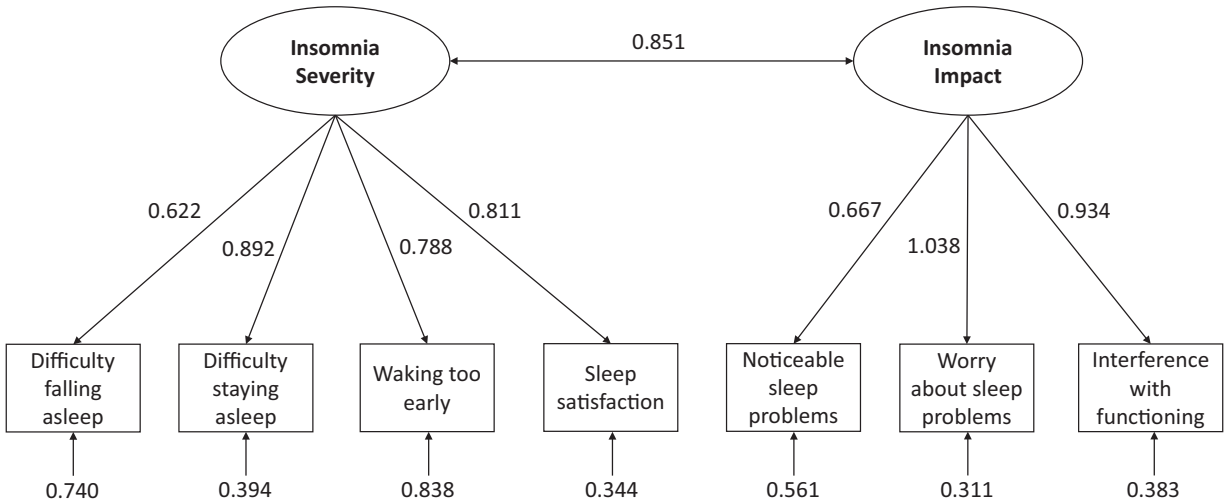

FIG. 1. The two-factor model of the Insomnia Severity Index in midlife women.

women, the CFI did not vary significantly. These results show the ISI factor structure did not differ by race.

When the intercepts were constrained to be equal across the two groups, however, the model fit was significantly inferior, suggesting a difference in at least one item intercept. Additional analyses not shown in the table showed one difference between racial groups. When we compared differences of the intercepts of the individual items between black and white women, the item ''difficulty falling asleep'' ($P < 0.001$) showed a partial strong/scalar noninvariance of the ISI construct across race. We further considered the fit and the invariance properties of the two-factor model for the reduced six-item ISI scale, with the item ''difficulty falling asleep'' omitted. The fit of the two-factor model based on the reduced scale was good (CFI = 0.996, TLI = 0.993, RMSEA = 0.055, SRMR = 0.034). More importantly, there was no evidence for noninvariance of that model between black and white women, based on the reduced six-item scale meaning the scale performed the same in both groups.

DISCUSSION

The study results show that a two-factor model (symptom severity and daytime impact) is a good fit for this sample of midlife women with hot flashes. Although this is the first analysis to solely focus on the ISI factor structure in women, our findings are similar to previous research studies that found similar two-factor models that included mixed sex samples.[19,22,23] This is the first paper to focus solely on women in this specific age group with insomnia symptoms and hot flashes with more extensive analyses to determine factor structure.

TABLE 3. *Evaluation of invariance of Insomnia Severity Index between black and white women*

	CFI	ΔCFI	*P*
Invariance			
Configural	0.997	NA	NA
Loadings	0.990	0.007	0.080
Intercepts	0.986	0.004	0.010
Means	0.985	0.001	0.396

CFI, Comparative Fit Index; ΔCFI, Change in Comparative Fit Index.

The ISI was developed by the original authors to provide a reliable and valid brief scale to screen for insomnia and provide a tool for evaluation of treatment outcomes.[16] No factor structure was, however, presented in that original article. Typically, the instrument is presented using the established cutoff scores and clinically significant change scores to determine incidence of insomnia and treatment response. Subsequent researchers have added to the dimensionality of the measure by conducting factor analysis, suggesting that the measure can perform differently in various populations.[22] Specifically, the items from the questionnaire fall into the after two-factor model. Factor 1 includes items 1 to 4 of the questionnaire that focus on the symptoms of insomnia: difficulty falling asleep, staying asleep, waking too early, and satisfaction with current sleep. Factor 2 includes natural content on the impact of the insomnia symptoms. This is an important outcome as clinicians use both symptoms and impact of symptoms to diagnose and recommend treatment.

We found that the two-factor structure and factor intercorrelations were consistent between black and white women, meaning that the scale performed similarly in both groups. Comparisons, however, showed one difference between the models. The ''difficulty falling asleep'' item did not perform the same across two racial groups. This particular item was the weakest loading item on the symptom severity factor, suggesting the other three items within this factor were better indicators of that construct. After removing this item, the resulting model showed good fit and performed consistently across racial groups. This suggests consideration should be given to using the six-item version of the ISI to allow for meaningful comparisons of insomnia symptoms between black and white women. The goal is to have ISI models that measure insomnia symptoms equally well in both racial

groups to enable cross-group quantitative comparisons and ensure proper interpretation of results. Clinically, this item omission is unlikely to affect the accurate assessment of insomnia symptoms. The six-item version of the ISI is flexible and appropriate for use in both groups of women.

Study findings have important implications for scoring the ISI in clinical practice and research. For clinical purposes, it would be acceptable to report a single ISI score. The high correlation between the two underlying factors suggests that items might not need to be scored and reported as two separate scores to be clinically meaningful. Clinically, having an abbreviated scale or single-item questions could be beneficial as it reduces the time of administering and scoring the scale. For research purposes, it is, however, more descriptive to report scores for each of the two factors. In particular, in intervention studies, it may be beneficial to report scores for each factor to determine the differential impact of interventions on each factor. This is especially important when there could be a disconnect between symptom severity and the impact on daytime outcomes such as individuals that did not sleep but do not always rate daytime outcomes to be detrimentally impacted or vice versa.

Lastly, for clinicians looking for an abbreviated scale or single items to use during a clinical visit to assess sleep, two items emerged that could provide an initial screening for further exploration into sleep complaints. Looking at the items within each factor separately, single items emerged within each factor that could provide some indication of insomnia severity and daytime impact. For symptom severity, item 1 ‘‘‘difficulty staying asleep’’ had the highest correlation with the factor score. For impact of insomnia, the item 6 that evaluates how worried/distressed the person is about sleep could be the best indicator of how daytime functioning is impacted. In prior research studies that supported a three-factor structure, the sleep satisfaction item loaded into a different factor.[15,24] In our analyses, the sleep satisfaction loaded onto the nighttime symptoms severity factor. It is unclear why this particular item loads differently compared with previous literature but could be attributed to the particular sample and how that item is interpreted.

CONCLUSIONS

The incidence and prevalence of sleep complaints in women continue to be high, especially in midlife women with vasomotor symptoms, although women continue to be underrepresented in sleep research.[9] Using existing data, we performed the first factor analysis of the ISI in midlife women. Findings suggest the ISI performed equally well in both black and white women. Clinicians can also consider using an abbreviated screening of the two items suggested above in this population of women (item 1 ‘‘difficulty staying asleep’’ and item 6 ‘‘how worried/distressed are you about your current sleep problem’’) to determine if further exploration is needed regarding sleep complaints. For research, we recommend scoring based on a two-factor structure to understand both symptom severity and daytime impact, two important outcomes in descriptive and intervention research. In the clinical setting, we recommend scoring based on a one-factor structure (traditional scoring) or considering screening for insomnia using only the two items most highly correlated with each factor score. Clinically having an overall severity rating using the one factor or total score is optimal for ease of use and continues to provide information on insomnia severity for treatment.

Acknowledgments: The network sites that participated in these studies included Boston, MA (Massachusetts General Hospital and Brigham and Women’s Hospital; Principal Investigators: Lee Cohen, MD and Hadine Joffe, MD, MSc); Indianapolis, IN (Indiana University; Principal Investigator: Janet S Carpenter, PhD, RN, FAAN); Oakland, CA (Kaiser Permanente Division of Research; Principal Investigators: Barbara Sternfeld, PhD and Bette Caan, PhD); Philadelphia, PA (University of Pennsylvania; Principal Investigator: Ellen W. Freeman, PhD); Seattle, WA (Group Health Research Institute; Principal Investigators: Katherine M. Newton, PhD and Susan Reed, MD). The Data Coordinating Center of the network is based at the Fred Hutchinson Cancer Research Center; Principal Investigators: Andrea LaCroix, PhD and Katherine A. Guthrie, PhD. The chairperson is Kristine E. Ensrud, MD, University of Minnesota. The project coordinator is Sergei Romashkan, MD, PhD: National Institute on Aging/US National Institutes of Health, Bethesda, MD.

REFERENCES

1. Shaver J, Giblin E, Lentz M, Lee K. Sleep patterns and stability in perimenopausal women. *Sleep* 1988;11:556-561.
2. Shaver JL, Giblin E, Paulsen V. Sleep quality subtypes in midlife women. *Sleep* 1991;14:18-23.
3. Shaver JL, Woods NF. Sleep and menopause: a narrative review. *Menopause* 2015;22:899-915.
4. Shaver JL, Zenk SN. Sleep disturbance in menopause. *J Womens Health Gend Based Med* 2000;9:109-118.
5. Ohayon MM. Severe hot flashes are associated with chronic insomnia. *Arch Intern Med* 2006;166:1262-1268.
6. Ohayon MM, Carskadon MA, Guilleminault C, Vitiello MV. Meta-analysis of quantitative sleep parameters from childhood to old age in healthy individuals: developing normative sleep values across the human lifespan. *Sleep* 2004;27:1255-1273.
7. Fisher TE, Chervenak JL. Lifestyle alterations for the amelioration of hot flashes. *Maturitas* 2012;71:217-220.
8. Attarian HP, Schuman C, SpringerLink (Online service). Clinical handbook of insomnia. In: *Current Clinical Neurology.* 2nd ed. Totowa, NJ: Humana; 2010. Available at: http://login.ezproxy.lib.purdue.edu/login?url= http://dx.doi.org/10.1007/978-1-60327-042-7.
9. Attarian HP, Viola-Saltzman M. *Sleep Disorders in Women: A Guide to Practical Management.* 2nd ed. Totowa, NJ: Humana Press; 2006.
10. Kravitz HM, Ganz PA, Bromberger J, Powell LH, Sutton-Tyrrell K, Meyer PM. Sleep difficulty in women at midlife: a community survey of sleep and the menopausal transition. *Menopause* 2003;10:19-28.
11. Kravitz HM, Joffe H. Sleep during the perimenopause: a SWAN story. *Obstet Gynecol Clin North Am* 2011;38:567-586.
12. National Institutes of Health. *NIH State-of-the-Science Conference Statement onManagement of Menopause-Related Symptoms*. Bethesda, MD: National Institutes of Health; 2005.
13. Kryger MH, Roth T, Dement WC. *Principles and Practice of Sleep Medicine*. 4th ed. Philadelphia, PA: Elsevier/Saunders; 2005.
14. Perlis ML, Smith MT, Wilfred RP. Etiology and pathophysiology of insomnia. In: Kryger MH, Roth T, Dement WC, eds. *Principles and Practice of Sleep Medicine.* 4th ed. Philadelphia, PA: Saunders; 2005: 714–725.
15. Bastien CH, Vallieres A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Med* 2001;2:297-307.

16. Morin CM, Belleville G, Belanger L, Ivers H. The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep* 2011;34:601-608.
17. Thorndike FP, Ritterband LM, Saylor DK, Magee JC, Gonder-Frederick LA, Morin CM. Validation of the insomnia severity index as a web-based measure. *Behav Sleep Med* 2011;9:216-223.
18. Chung KF, Kan KK, Yeung WF. Assessing insomnia in adolescents: comparison of Insomnia Severity Index, Athens Insomnia Scale and Sleep Quality Index. *Sleep Med* 2011;12:463-470.
19. Yu DS. Insomnia Severity Index: psychometric properties with Chinese community-dwelling older people. *J Adv Nurs* 2010;66:2350-2359.
20. Fernandez-Mendoza J, Rodriguez-Munoz A, Vela-Bueno A, et al. The Spanish version of the Insomnia Severity Index: a confirmatory factor analysis. *Sleep Med* 2012;13:207-210.
21. Suleiman KH, Yates BC. Translating the insomnia severity index into Arabic. *J Nurs Scholarsh* 2011;43:49-53.
22. Moscou-Jackson G, Allen J, Smith MT, Haywood C Jr. Psychometric validation of the Insomnia Severity Index in adults with sickle cell disease. *J Health Care Poor Underserved* 2016;27:209-218.
23. Savard MH, Savard J, Simard S, Ivers H. Empirical validation of the Insomnia Severity Index in cancer patients. *Psychooncology* 2004;14: 429-441.
24. Chen PY, Yang CM, Morin CM. Validating the cross-cultural factor structure and invariance property of the Insomnia Severity Index: evidence based on ordinal EFA and CFA. *Sleep Med* 2015;16:598-603.
25. Freeman EW, Guthrie KA, Caan B, et al. Efficacy of escitalopram for hot flashes in healthy menopausal women: a randomized controlled trial. *JAMA* 2011;305:267-274.
26. Newton KM, Carpenter JS, Guthrie KA, et al. Methods for the design of vasomotor symptom trials: the menopausal strategies: finding lasting answers to symptoms and health network. *Menopause* 2014;21:45-58.
27. Carpenter JS, Newton KM, Sternfeld B, et al. Laboratory and ambulatory evaluation of vasomotor symptom monitors from the Menopause Strategies Finding Lasting Answers for Symptoms and Health network. *Menopause* 2012;19:664-671.
28. Cohen LS, Joffe H, Guthrie KA, et al. Efficacy of omega-3 for vasomotor symptoms treatment: a randomized controlled trial. *Menopause* 2014;21: 347-354.
29. Joffe H, Guthrie KA, LaCroix AZ, et al. Low-dose estradiol and the serotonin-norepinephrine reuptake inhibitor venlafaxine for vasomotor symptoms: a randomized clinical trial. *JAMA Intern Med* 2014;174:1058-1066.
30. Newton KM, Reed SD, Guthrie KA, et al. Efficacy of yoga for vasomotor symptoms: a randomized controlled trial. *Menopause* 2014;21:339-346.
31. Sternfeld B, Guthrie KA, Ensrud KE, et al. Efficacy of exercise for menopausal symptoms: a randomized controlled trial. *Menopause* 2014;21:330-338.
32. Sternfeld B, LaCroix A, Caan BJ, et al. Design and methods of a multisite, multi-behavioral treatment trial for menopausal symptoms: the MsFLASH experience. *Contemp Clin Trials* 2013;35:25-34.
33. Forero CG, Maydeu-Olivares A, Gallardo-Pujol D. Factor analysis with ordinal indicators: a Monte Carlo Study comparing DWLS and ULS estimation. *Struct Equ Modeling* 2009;16:625-641.
34. Hu Lt, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling* 1999;6:1-55.
35. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Equ Modeling* 2002;9:233-255.
36. Hall P, Wilson SR. Two guidelines for bootstrap hypothesis testing. *Biometrics* 1991;47:757–762.
37. Rosseel Y. lavaan: an R package for structural equation modeling. *J Stat Softw* 2011;48: 1–36.