Menopause: The Journal of The North American Menopause Society
Vol. 22, No. 11, pp. 1190-1196
DOI: 10.1097/GME.0000000000000459

Confirmatory factor analysis of the Pittsburgh Sleep Quality Index in women with hot flashes

Julie L. Otte, PhD, RN,[1] Kevin L. Rand, PhD,[2] Carol A. Landis, PhD, RN, FAAN,[3] Misti L. Paudel, PhD,[4] Katherine M. Newton, PhD,[5,6] Nancy Woods, PhD, RN, FAAN,[3] and Janet S. Carpenter, PhD, RN, FAAN[1]

Abstract

Objective: Women, especially those with hot flashes, report poor sleep quality during various stages of the menopausal transition and postmenopause. Sleep measurements vary widely because of the copious instruments available. The Pittsburgh Sleep Quality Index (PSQI) is a frequently used questionnaire that produces a single score for sleep quality. This one-factor structure has not received consistent support in the literature. The goal of this analysis was to determine the best factor structure of the PSQI in women with hot flashes.

Methods: A confirmatory factor analysis was conducted on PSQI baseline data from three randomized controlled clinical trials enrolling perimenopausal and postmenopausal women with hot flashes (N = 849) from the Menopause Strategies: Finding Lasting Answers for Symptoms and Health network. Several a priori factor models were compared.

Results: One-factor and two-factor models did not fit the data. A three-factor model comprising sleep efficiency, perceived sleep quality, and daily disturbance showed good fit; however, the sleep medication item was dropped because of poor fit and low rates of sleep medication use. The three-factor model was examined in African-American and white subsamples and was found to be similar in both groups; however, two items showed small group differences in strength as indicators.

Conclusions: Sleep quality in midlife women with hot flashes, as measured by the PSQI, seems to comprise three correlated factors. Minor measurement differences detected between groups are of research interest but do not necessitate different scoring practices. Additional research is needed to further define sleep quality and its associations with health-related outcomes.

Key Words: Sleep – Sleep quality – Measurement – Menopause – Women – Factor analysis.

Sleep complaints are twice as prevalent in women than in men, and disturbed sleep contributes to daytime fatigue and poor quality of life.[1] Women report poor sleep quality during various stages of the menopausal transition and postmenopause. This is especially true for women experiencing hot flashes, which are prevalent in as many as 50% of women during the late menopausal transition and 48% of women during early postmenopause.[2]

Selecting a measure of sleep to use in research studies largely depends on the research questions. Current self-report sleep measures typically address a single sleep complaint or combinations of sleep complaints, such as (a) sleep duration or total sleep time; (b) sleep latency (minutes to fall asleep); (c) number and duration of nighttime awakenings; (d) quality of perceived sleep; (e) sleep efficiency (amount of time in bed spent asleep vs amount of time spent in bed); (f) sleep

Received September 19, 2014; revised and accepted February 12, 2015.
From the [1]School of Nursing, Indiana University, Indianapolis, IN; [2]Department of Psychology, Indiana University–Purdue University, Indianapolis, IN; [3]University of Washington School of Nursing, Seattle, WA; [4]University of Minnesota School of Public Health and Epidemiology, Minneapolis, MN; [5]Harvard Medical School, Boston, MA; and [6]Group Health Research Institute, Seattle, WA.
The network sites that participated in these studies included the following: Boston, MA (Massachusetts General Hospital and Brigham and Women's Hospital; principal investigators: Lee Cohen, MD; 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; Bette Caan, PhD); Philadelphia, PA (University of Pennsylvania; principal investigator: Ellen W. Freeman, PhD); Seattle, WA (Group Health Research Institute; principal investigators: Katherine Newton, PhD; Susan Reed, MD).

The Data Coordinating Center of the network is based at the Fred Hutchinson Cancer Research Center (principal investigators: Andrea LaCroix, PhD; Katherine A. Guthrie, PhD [Garnet L. Anderson, PhD, was previously a principal investigator]; statistician: Joseph C. Larson, MS). The chairperson is Kris E. Ensrud, MD (University of Minnesota). The project coordinator is Judy Hannah, PhD (National Institute on Aging/US National Institutes of Health, Bethesda, MD).

Funding/support: The Menopause Strategies: Finding Lasting Answers for Symptoms and Health studies were supported by a cooperative agreement issued by the National Institute on Aging, in collaboration with the Eunice Kennedy Shriver National Institute of Child Health and Development, the National Center for Complementary and Alternative Medicine, and the Office of Research and Women's Health, and by National Institute on Aging grants U01AG032659, U01AG032669, U01AG032682, U01AG032699, and U01AG032700. In Indiana, the 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.

Financial disclosure/conflicts of interest: None reported.
Address correspondence to: Julie L. Otte, PhD, RN, 1111 Middle Drive, Indianapolis, IN 46202. E-mail: jlelam@iu.edu

medication use; and (g) daytime dysfunction (daytime sleepiness, inability to function during the day, and feeling tired or fatigued).[3-5] Instruments address either a single aspect or a combination of the abovementioned aspects and are often specific for the overarching sleep problem being addressed. Recommendations have been made for the standardization of sleep measures to facilitate meta-analyses of research on patients with cancer and insomnia,[6] but a variety of instruments continue to be used.[7]

The Pittsburgh Sleep Quality Index (PSQI) is one of the questionnaires most frequently used for the assessment of self-reported sleep quality in studies of men and women with or without long-term illnesses.[4,8] The PSQI was created to capture sleep quality in a sample of patients with psychiatric conditions. The original intent of the questionnaire was to (a) create a reliable and valid standardized measure of sleep quality; (b) provide a tool for distinguishing good sleep quality from poor sleep quality; (c) create a measure of sleep quality that was user-friendly for individuals, clinicians, and researchers to interpret; and (d) provide a clinical tool for assessing a list of sleep disturbances that impact sleep quality in populations with psychiatric conditions. The PSQI contains seven individual subscales and a single-factor global score higher than 5 representing poor overall sleep quality. The global PSQI score has been deemed a simple measure for use in clinical research to identify good versus poor sleep. The PSQI single-factor score for sleep quality has consistently reported acceptable reliability and validity in various populations[9-12] and has been translated into several different languages.[9]

The single-factor structure of the PSQI has been analyzed to understand whether the global sleep quality score derived from all of the seven individual subscale scores is the best representation of sleep quality. This research—conducted in various adult samples with long-term illnesses (depression, breast cancer, post–renal transplantation, and rheumatoid arthritis) and without long-term illnesses (community-dwelling English and Spanish men and women, Nigerian students, and nondepressed men and women)—has yielded inconsistent results.[13-18] Of seven studies, four found that a model comprising three factors (eg, sleep efficiency, perceived sleep quality, and daytime functioning) was a better fit to the data than was a one-factor global score for sleep quality.[13-16] One study found a one-factor model (sleep quality),[19] whereas two studies reported two-factor models (sleep efficiency and perceived sleep quality), as being most appropriate for the populations being studied.[17,18] A study by Otte et al[18] also found racial differences in sleep quality and sleep latency, as well as in sleep efficiency and sleep quality, in a nested-model comparison of African-American and white breast cancer survivors. The findings from these studies raise questions about whether it is appropriate to use a single global score to differentiate good sleep quality from poor sleep quality across clinical populations. These previous studies had several limitations that make it difficult to generalize the findings on sleep quality to other populations, such as women with hot flashes. These include sample heterogeneity (mixed sex, wide age ranges, and specific focus on long-term illnesses) and varied sample sizes (from 107 to 1,174).[18] No multigroup analyses using factor modeling have determined whether mixed sample groups limit factor findings.[15] In addition, findings have been limited to research studies, and it is unclear how findings can be translated into clinical practice.

Because sleep is a significant problem in women during the menopausal transition, it would be informative to evaluate whether the single-factor measure from the PSQI, compared with the previously reported two-factor model (by Otte et al[18]) and three-factor model (by Cole et al[15]), best reflects good versus poor sleep quality in a large sample of relatively healthy midlife women experiencing hot flashes in late perimenopause and early postmenopause.[13-15] Additional nested-model comparisons between African-American and white women can identify potential racial differences that have been reported in a previous study of women with breast cancer.[18] The goal of this analysis is to provide recommendations on the best factor structure of the PSQI for analyzing sleep quality in research trials of women with hot flashes.

METHODS

Procedures

Baseline data from the PSQI were pooled from three samples of midlife women participating in Menopause Strategies: Finding Lasting Answers for Symptoms and Health (MsFLASH) controlled randomized trials. Study 1 was a randomized controlled trial of escitalopram versus placebo for the treatment of hot flashes.[20-22] Study 2 was a three-by-two factorial design study that yielded three comparison groups: (a) yoga versus usual activity; (b) exercise versus usual activity; and (c) ω-3 fatty acids versus placebo pill for hot flashes.[23-26] Study 3 was a randomized controlled trial of low-dose estradiol versus placebo, and extended-release venlafaxine versus placebo for menopausal symptoms.[27] Data from midlife women ($N = 899$) in these three studies were evaluated (study 1, $n = 205$; study 2, $n = 355$; study 3, $n = 339$). Descriptions of the procedures used in the MsFLASH trials have been published elsewhere.[20-27]

Setting

Participants were recruited from five MsFLASH network sites (Seattle, Boston, Oakland, Indianapolis, and Philadelphia). Participants were recruited from July 2009 to October 2012 through targeted mailings to midlife women using purchased mailing lists.

Measures

Sample characteristics were collected using an adapted multi-item questionnaire that included items from the PSQI and also measured race, age, marital status, educational level, income level, employment status, height, and weight.

Sleep was assessed using the PSQI,[4] a 19-item scale that provides seven component scores (ranges, 0-3): sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbance, sleep medication use, and daytime dysfunction. The seven component scores are combined to produce a global sleep quality index score. The global scores range from 0 to 21, with scores higher than 5 reflecting poor global sleep quality[4] and with scores higher than 8 reflecting poor sleep and high daytime fatigue burden.[9]

Statistical analysis

To determine the best factor model and scoring method for the PSQI among women with hot flashes, we conducted a confirmatory factor analysis using LISREL version 8.8.[28] Because the single global score of the PSQI is frequently used in the literature, we examined the fit of a one-factor model. Because recent studies (Cole et al[15] and Otte et al[18]) found support for both two-factor and three-factor models, we examined the fit of these models in the present sample. Given that parsimonious models are preferred to more complex ones, we examined these models in order of increasing complexity.

Subscale scores from the seven PSQI components represent ordinal rather than continuous data and should not be analyzed using methods that assume that they are continuous variables with metric properties.[29] Instead, PRELIS 2.8[28] was used to estimate polychoric correlations among subscale scores.[29] Polychoric correlations are used to describe associations between observed ordinal variables (ie, PSQI subscale scores) that represent underlying phenomena (ie, sleep characteristics) that are normally distributed and continuous. These polychoric correlations were used to create an asymptotic covariance matrix (similar to a correlation matrix) that represents all of the interrelationships among the PSQI subscales. This matrix of associations among the PSQI subscales was used to examine the various factor models (using weighted least-squares).

The appropriateness of a factor model is evaluated based on how well the theoretical model fits the observed matrix of associations. Several indices have been developed to describe model fit, with each index using different criteria.[30] We chose three of the most commonly used fit indices: (a) χ^2; (b) root-mean-square error of approximation (RMSEA) and its 90% CI; and (c) comparative fit index (CFI). χ^2 is an index of absolute fit between a hypothesized covariance matrix and the observed covariance matrix. A nonsignificant χ^2 is evidence of acceptable fit. RMSEA is a parsimonious index, meaning that the complexity of the hypothesized model is taken into consideration in evaluating its fit with the observed covariance matrix. RMSEA values of 0.06 or lower indicate acceptable fit.[30] CFI is an incremental fit index, with values of 0.95 or higher indicating acceptable fit.[30]

Once the best-fitting model has been determined in the overall sample, its consistency across subsamples can be examined using nested-model comparisons. Such comparisons involve statistically comparing different parameters of the model to find which ones are significantly different in one group versus the other.

RESULTS

Sample characteristics

Most of the women were non-Hispanic (97.2%), white (59.0%), partnered (63.2%), employed (69.7%), and highly educated (79.6% with some college). A substantial portion of the sample was African American (33.8%), with smaller proportions self-identifying as American Indian (1.2%), Asian/Pacific Islander (2.2%), or "other" (2.8%). The mean (SD) age was 54.47 (3.83) years. The mean (SD) global PSQI score was 7.82 (3.47). Most of the women (72.0%) had global scores above the cutoff score of 5, suggesting poor sleep quality, and a sizeable portion of those women (39.8%) had scores higher than 8, indicating poor sleep quality and high fatigue burden.

Responses to the PSQI were scored into the seven subscales,[4] and 94.4% of the sample had complete baseline data on all component scores. Means, SDs, and correlations among the subscales are shown in Table 1. The number of missing values for each subscale ranged from 7 to 17, with 50 participants missing data on at least one subscale. None of the cases with missing values had adequate information to allow for imputation. Therefore, these 50 cases were excluded from confirmatory factor analysis, resulting in a final sample size of 849. There were no differences in age between women with complete data (mean, 54.48 y) and women with missing data (mean, 54.16 y). There was a small but significant relationship between education and missing data points; women with less education were more likely to have missing data ($r_s = -0.078$, $P = 0.019$). There also was a relationship between race/ethnicity and missing data ($\chi^2 = 19.04$, $df = 5$, $P = 0.002$); specifically, proportionately more African-American women than white women had missing data (7.89% vs 2.19%, respectively). Similarly, never-married women had proportionately more missing data than women who were married or divorced/separated (8.47% vs 3.40% and 2.16%, respectively; $\chi^2 = 9.51$, $df = 4$, $P = 0.049$). This could

TABLE 1. *Complete data sample sizes, means, SDs, and correlations for seven Pittsburgh Sleep Quality Index component scores*

	n	Mean (SD)	1	2	3	4	5	6
(1) Subjective sleep quality	890	1.62 (0.74)	–					
(2) Sleep latency	885	1.33 (1.01)	0.39[a]	–				
(3) Sleep duration	890	0.83 (0.93)	0.46[a]	0.32[a]	–			
(4) Habitual sleep efficiency	888	0.85 (1.06)	0.30[a]	0.27[a]	0.55[a]	–		
(5) Sleep disturbances	892	1.64 (0.57)	0.28[a]	0.17[a]	0.15[a]	0.16[a]	–	
(6) Sleep medication use	884	0.61 (1.03)	0.07[b]	0.15[a]	0.01	0.03	0.10[a]	–
(7) Daytime dysfunction	882	0.97 (0.75)	0.31[a]	0.16[a]	0.19[a]	0.08[a]	0.29[a]	0.06

Score range, 0 to 3.
Ns for correlations vary between 875 and 890 because of missing data.
[a] $P < 0.01$.
[b] $P < 0.05$.

TABLE 2. *Fit indices for different factor models of the Pittsburgh Sleep Quality Index*

Model	Factor	χ^2	*df*	*P*	RMSEA (90% CI)	CFI
1	One factor (overall sleep quality)	107.84	14	<0.00001	0.089 (0.074-0.100)	0.90
2	Two factors (sleep efficiency and perceived sleep quality)[15]	62.71	13	<0.00001	0.067 (0.051-0.084)	0.94
3	Two factors (sleep efficiency and perceived sleep quality)[18]	41.97	8	<0.00001	0.071 (0.051-0.093)	0.96
4	Three factors (sleep efficiency, perceived sleep quality, and daily disturbance)[15]	31.92	11	0.00079	0.047 (0.029-0.067)	0.98
5	Four factors (sleep efficiency, perceived sleep quality, daily disturbance, and sleep medication use)	25.31	9	0.00265	0.046 (0.025-0.068)	0.98
6	Three factors (sleep efficiency and perceived sleep quality; daily disturbance; without sleep medication use indicator)[15]	12.27	6	0.05600	0.035 (0.000-0.063)	0.99
6a	Three factors (sleep efficiency and perceived sleep quality; daily disturbance; without sleep medication use indicator [two groups])[15]	23.08	12	0.02709	0.048 (0.016-0.078)	0.99

RMSEA, root-mean-square error of approximation; CFI, comparative fit index.

be attributed to women without bed partners not providing a response for the bed partner–related item.

Confirmatory factor analysis

Sample sizes, means, SDs, and correlations among the seven PSQI subscales are shown in Table 1. Univariate skew and kurtosis statistics were examined for each component score, and none of the indicators showed excessive skew or kurtosis.[31] Hence, none of the data were transformed. The fit indices for the tested models are shown in Table 2. The one-factor model (model 1), consistent with the global PSQI score, did not show acceptable fit on any of the indices. The two-factor model (sleep efficiency and perceived sleep quality; model 2)[15] also did not show acceptable fit on any index. A modified two-factor model (model 3)[18] showed better fit, with the CFI indicating acceptable fit. The three-factor model (sleep efficiency, perceived sleep quality, and daily disturbance; model 4)[15] showed the best fit, with acceptable values for RMSEA and CFI.

The three-factor model (model 4),[15] with its standardized coefficients, is presented in Figure 1. As the model shows, the loading from perceived sleep quality to the sleep medication use subscale is small ($\beta = 0.15$), leaving a large amount of unexplained variability in this item ($\delta = 0.98$). The small loading of the sleep medication use item is consistent with its weak correlation with the other subscales of the PSQI (Table 1). As the correlation table shows, even the strongest association of sleep medication use with sleep latency was small ($r = 0.15$). This is probably attributable to restriction of range, as 68% of the sample reported using no sleep medication at all.

Given the low rate of sleep medication use and its resultant weak correlations with the other PSQI subscales, we examined an alternative four-factor model where sleep medication use was modeled as a separate factor along with

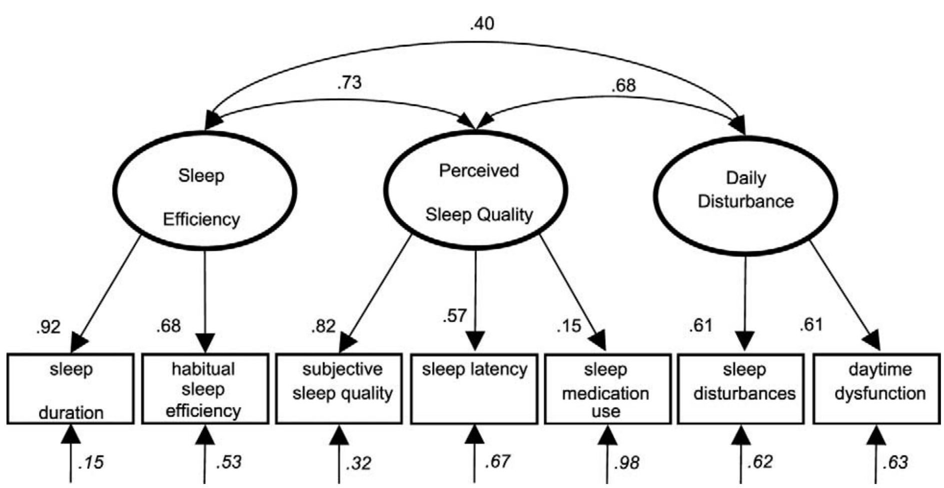

FIG. 1. Three-factor model of sleep quality measured with the Pittsburgh Sleep Quality Index in midlife women.[15] All coefficients are standardized and significant at $P < 0.05$, unless otherwise noted. Ovals, latent variables; rectangles, measured variables; single-headed arrows, factor loadings; double-headed arrows, correlations. Italicized coefficients are error terms representing the proportion of variation in the variable not accounted for by the model.

perceived sleep quality, sleep efficiency, and daily disturbance. This model showed good fit to the data (Table 2, mode 5). However, the sleep medication use factor did not correlate strongly with any of the other factors ($\psi = 0.03$-0.18), questioning the utility of this item in assessing sleep quality among women with hot flashes.

Consequently, we examined the fit of an alternative three-factor model (sleep efficiency, perceived sleep quality, and daily disturbance), excluding the sleep medication use item; this model showed good fit to the data across all indices, including a nonsignificant χ^2 (Table 2, model 6). Therefore, we concluded that this three-factor conceptualization represented the best model of sleep quality among women with hot flashes (Fig. 2).

To examine whether the factor structure was consistent between African-American and white subsamples, we conducted several two-group comparisons. Although the sample consisted of women from several racial/ethnic groups, only the African-American (n = 304) and white (n = 530) subsamples had sufficient numbers to allow for statistical comparisons of factor models. First, a two-group three-factor model was examined, with all parameters freed to vary between groups. This model showed good fit to the data across all indices (Table 2, model 6a). We constrained factor loadings to be equal across both groups, and the resulting three-factor model showed significantly worse fit to the data ($\Delta\chi^2 = 23.81$, $df = 6$, $P = 0.0005$), indicating that the factor loadings differed between African-American and white women.

Subsequently, we conducted a series of nested-model comparisons to determine which factor loadings onto the PSQI subscales differed between the two groups (Table 3). The loading values for the perceived sleep quality factor onto the subjective sleep quality subscale differed significantly between groups, with standardized loadings of 0.72 and 0.87 for white and African-American women, respectively. This suggests that the subjective sleep quality subscale is a slightly stronger indicator of perceived sleep quality factor in African-American women compared with white women. In addition, the loading values for the sleep efficiency factor onto the habitual sleep efficiency subscale differed significantly between groups, with standardized loadings of 0.73 and 0.44 for white and African-American women, respectively. This indicates that, among postmenopausal women, the habitual sleep efficiency subscale is a stronger indicator of sleep efficiency factor for white women compared with African-American women. Although we were able to detect two differences in the factor loadings of African-American versus white women, these results do not require different PSQI scoring procedures for the two groups. For both groups of women, subjective sleep quality and habitual sleep efficiency were reliable positive indicators of their respective constructs. These loading differences may indicate subtle cultural differences in expectations or habitual practices with regard to sleep and sleep-related behaviors.

However, these differences are mainly of interest for research purposes and do not necessitate different scoring procedures for African-American versus white women.

Next, we constrained the correlations among the three factors to be equal between groups to test whether factor intercorrelations differed in African-American and white women. The resulting model showed no significant difference in fit ($\Delta\chi^2 = 5.34$, $df = 3$, $P = 0.1485$), indicating that correlations among sleep efficiency, perceived sleep quality, and daily disturbance were the same in African-American and white women. This result suggests that the structure of sleep quality is similar in African-American and white women with hot flashes and that the pattern of relationships among the factors is the same for both groups. Taken together, these results suggest that the three-factor model is the best conceptualization of the PSQI among women with hot flashes and that there are no major racial differences.

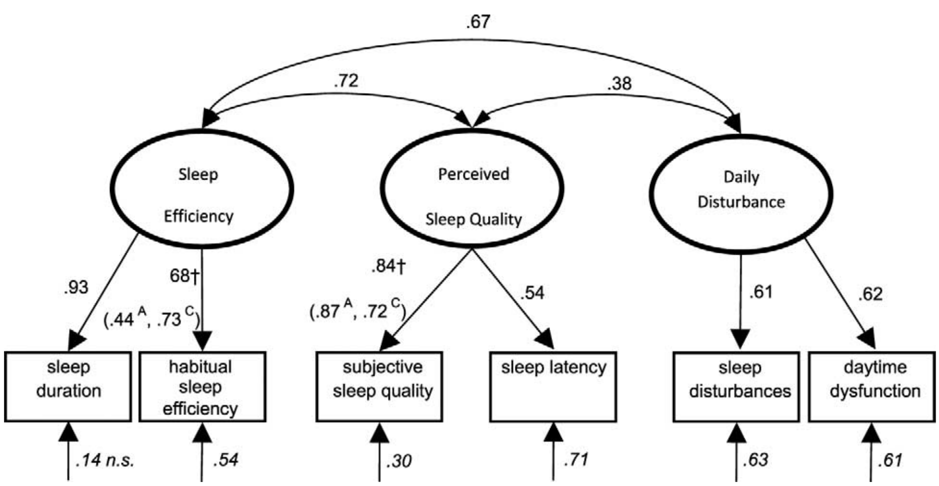

FIG. 2. Alternative three-factor model of sleep quality measured with the Pittsburgh Sleep Quality Index in midlife women without sleep medication use item. All coefficients are standardized and significant at $P < 0.05$, unless otherwise noted. Ovals, latent variables; rectangles, measured variables; single-headed arrows, factor loadings; double-headed arrows, correlations. Italicized coefficients are error terms representing the proportion of variation in the variable not accounted for by the model. [†]Parameters were found to differ significantly between African-American and white subsamples; [A]parameter estimate for African-American subsample; [C]parameter estimate for white subsample. n.s., not significant.

TABLE 3. χ^2 *tests comparing factor loadings of the Pittsburgh Sleep Quality Index in African-American and white women*

Factor	Indicator	λ^A	λ^C	$\Delta\chi^2$	*df*	*P*
Sleep efficiency	Sleep duration	0.95	0.86	0.81	1	0.368
	Habitual sleep efficiency[a]	0.44[a]	0.73[a]	14.66[a]	1[a]	<0.001[a]
Perceived sleep quality	Subjective sleep quality[a]	0.87[a]	0.72[a]	4.21[a]	1[a]	0.040[a]
	Sleep latency	0.51	0.51	0.00	1	1.000
Daily disturbance	Sleep disturbances	0.47	0.54	0.44	1	0.507
	Daytime disturbance	0.54	0.65	1.11	1	0.292

λ^A, factor loading for African-American subsample; λ^C, factor loading for white subsample.
[a]$P < 0.05$.

DISCUSSION

The results of this study showed that sleep quality (as measured by the PSQI) in midlife women experiencing daily hot flashes in the late menopausal transition and early postmenopause is multifaceted, with three correlated factors: sleep efficiency, perceived sleep quality, and daily disturbance. This three-factor model fit well in both African-American and white subsamples, and the correlations among the factors were equivalent across the two groups. This finding suggests that, from a practical perspective, the structure of sleep quality, at least as measured by the PSQI, is consistent in African-American and white women.

Overall, the results of these analyses in midlife women with hot flashes are consistent with previous reports analyzing three-factor, two-factor, and one-factor models of sleep quality. The nested-model comparisons by race show no major differences from the previous study by Otte et al.[18] The model results are similar to previous factor analyses that also found that a three-factor model (sleep efficiency, perceived sleep quality, and daily disturbance) best fit the data in depressed and nondepressed adults,[15] 520 Nigerian university students,[13] 135 patients with renal transplantation,[14] and 3,667 community-dwelling English- and Spanish-speaking Hispanic and non-Hispanic adults.[16] Conversely, the traditional one-factor model (sleep quality) and two-factor model (sleep efficiency and perceived sleep quality) consisted of samples of 197 Chinese women with breast cancer,[19] 417 depressed and nondepressed older adults,[15] 107 patients with rheumatoid arthritis,[17] and 1,174 nondepressed breast cancer survivors.[18] Two of the prior studies also produced a better model fit by removing the sleep medication use item during the analysis.[16,17] The findings from these various studies show that the PSQI factor structure differs among studies that have different mixed sexes, races, and long-term illnesses. Differences in factor structure can also be attributed to the wide range of sample sizes among the studies, which can impact factor structure. The result of this variability in samples and factor structures suggests that future research studies should consider how the PSQI should best be analyzed in light of these findings. The results also suggest that when evaluating intervention efficacy using the PSQI, using different factor structures can delineate improvements in more specific areas such as sleep efficiency or perceived sleep quality.

Although we found that the factor structure and factor intercorrelations were consistent between African-American and white women, nested-model comparisons revealed two measurement differences. First, we found that the subjective sleep quality subscale was a stronger indicator of perceived sleep quality factor in African-American women compared with white women. This difference is important from the standpoint of measurement and research. However, it is unlikely to be of any practical or clinical significance.[32] For both African-American and white women, subjective sleep quality is the best indicator of perceived sleep quality. Second, we found that the sleep efficiency factor differed significantly in African-American versus white women. Specifically, we found that sleep efficiency among African-American women was primarily captured by sleep duration subscore. In contrast, the habitual sleep efficiency subscale was a stronger indicator of sleep efficiency in white versus African-American women. This could mean that African Americans with poor sleep efficiency have shorter sleep duration. Again, this difference is important from a research and measurement perspective but is unlikely to matter in the clinical assessment of sleep quality.

Our results suggest that the sleep medication use item may not be a good indicator of sleep quality among women participating in clinical trials for hot flashes, probably owing to low sleep medication use in this population. This is a consistent finding across multifactor model analyses. In the present sample, 68% of women reported not using any sleep medications at all. Sleep medication use may be an indicator of sleep quality among older women or those being treated for illnesses (eg, breast cancer), as they may be more likely to take medications to improve different aspects of quality of life. Further research is needed to determine what factors drive the use of sleep medications among women. Although it is clinically important to query individuals regarding sleep medication use, it is recommended that the item be removed from PSQI subscale scoring until further validity and reliability testing has been performed.

The results from this study and previous factor analyses highlight the fact that the definition of "sleep quality" varies among researchers, clinicians, and individuals, resulting in a poorly defined concept.[32] The concept of perceived sleep quality tends to have a single meaning based on subjective criteria as a way to interpret the overall perception of an individual's sleep. A qualitative study of adults with or without insomnia found that perceived sleep quality is a multifaceted concept that includes (a) being tired upon waking that lasts throughout the day; (b) feeling of being rested upon waking; and (c) number of awakenings during the night.[32] It could be suggested that assessing an individual's sleep quality requires appraisal of sleep as a continuous assessment of multiple factors and cannot be distilled into a single-factor concept as in the PSQI. However, the data findings in this analysis are limited to women experiencing hot flashes in the late menopausal transition and early postmenopause, limiting generalizability to all midlife women of this age range and menopause status.

CONCLUSIONS

Because of the wide range of research instruments used to measure the different aspects of self-reported sleep, it is challenging to aggregate findings across studies to compare the incidence and prevalence of sleep complaints. This is especially true in women, who continue to remain underrepresented in sleep research.[1] The PSQI, as a measure of sleep quality, is sensitive to race, ethnicity, long-term illness, and sex, making cross-study comparisons of sleep a continuing challenge.

Based on the results of this study, alternative PSQI scoring can be considered and further evaluated using the three-factor model for self-reported sleep in studies of midlife women with hot flashes. Revisions would consist of recalculations of subscale loadings for sleep efficiency (sleep disturbance and habitual sleep efficiency subscales), perceived sleep quality (subjective sleep quality and sleep latency subscales), and daily disturbance (sleep disturbance and daytime dysfunction subscales). There is also the issue of how sleep quality is defined and the problem with a single score reflecting a complex concept. Further work is needed to address this issue on measuring sleep quality. Improving our understanding of the larger concept of sleep quality can further research efforts to facilitate better subjective measurement, which can be used to test interventions that can improve sleep in women during the menopausal transition.

REFERENCES

1. Attarian HP, Viola-Saltzman M. *Sleep Disorders in Women: A Guide to Practical Management*. 2nd ed. Totowa, NJ: Humana Press; 2006.
2. National Institutes of Health. *NIH State-of-the-Science Conference Statement on Management of Menopause-Related Symptoms*. Bethesda, MD: National Institutes of Health; 2005.
3. Berger AM. Patterns of fatigue and activity and rest during adjuvant breast cancer chemotherapy. *Oncol Nurs Forum* 1998;25:51-62.
4. Buysse DJ, Reynolds CF III, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res* 1989;28:193-213.
5. Miaskowski C, Lee KA. Pain, fatigue, and sleep disturbances in oncology outpatients receiving radiation therapy for bone metastasis: a pilot study. *J Pain Symptom Manage* 1999;17:320-332.
6. Buysse DJ, Ancoli-Israel S, Edinger JD, Lichstein KL, Morin CM. Recommendations for a standard research assessment of insomnia. *Sleep* 2006;29:1155-1173.
7. Berger AM, Parker KP, Young-McCaughan S, et al. Sleep wake disturbances in people with cancer and their caregivers: state of the science. *Oncol Nurs Forum* 2005;32:E98-E126.
8. Buysse DJ, Hall ML, Strollo PJ, et al. Relationships between the Pittsburgh Sleep Quality Index (PSQI), Epworth Sleepiness Scale (ESS), and clinical/polysomnographic measures in a community sample. *J Clin Sleep Med* 2008;4:563-571.
9. Carpenter JS, Andrykowski MA. Psychometric evaluation of the Pittsburgh Sleep Quality Index. *J Psychosom Res* 1998;45:5-13.
10. Gentili A, Weiner DK, Kuchibhatla M, Edinger JD. Test-retest reliability of the Pittsburgh Sleep Quality Index in nursing home residents. *J Am Geriatr Soc* 1995;43:1317-1318.
11. Backhaus J, Junghanns K, Broocks A, Riemann D, Hohagen F. Test-retest reliability and validity of the Pittsburgh Sleep Quality Index in primary insomnia. *J Psychosom Res* 2002;53:737-740.
12. Beck SL, Schwartz AL, Towsley G, Dudley W, Barsevick A. Psychometric evaluation of the Pittsburgh Sleep Quality Index in cancer patients. *J Pain Symptom Manage* 2004;27:140-148.
13. Aloba OO, Adewuya AO, Ola BA, Mapayi BM. Validity of the Pittsburgh Sleep Quality Index (PSQI) among Nigerian university students. *Sleep Med* 2007;8:266-270.
14. Burkhalter H, Sereika SM, Engberg S, Wirz-Justice A, Steiger J, De Geest S. Structure validity of the Pittsburgh Sleep Quality Index in renal transplant recipients: a confirmatory factor analysis. *Sleep Biol Rhythms* 2010;8:274-281.
15. Cole JC, Motivala SJ, Buysse DJ, Oxman MN, Levin MJ, Irwin MR. Validation of a 3-factor scoring model for the Pittsburgh Sleep Quality Index in older adults. *Sleep* 2006;29:112-116.
16. Tomfohr LM, Schweizer CA, Dimsdale JE, Loredo JS. Psychometric characteristics of the Pittsburgh Sleep Quality Index in English speaking non-Hispanic whites and English and Spanish speaking Hispanics of Mexican descent. *J Clin Sleep Med* 2013;9:61-66.
17. Nicassio PM, Ormseth SR, Custodio MK, Olmstead R, Weisman MH, Irwin MR. Confirmatory factor analysis of the Pittsburgh Sleep Quality Index in rheumatoid arthritis patients. *Behav Sleep Med* 2014;12:1-12.
18. Otte JL, Rand KL, Carpenter JS, Russell KM, Champion VL. Factor analysis of the Pittsburgh Sleep Quality Index in breast cancer survivors. *J Pain Symptom Manage* 2013;45:620-627.
19. Ho RT, Fong TC. Factor structure of the Chinese version of the Pittsburgh Sleep Quality Index in breast cancer patients. *Sleep Med* 2014;15:565-569.
20. 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.
21. 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.
22. 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.
23. 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.
24. Newton KM, Reed SD, Guthrie KA, et al. Efficacy of yoga for vasomotor symptoms: a randomized controlled trial. *Menopause* 2014;21:339-346.
25. Sternfeld B, Guthrie KA, Ensrud KE, et al. Efficacy of exercise for menopausal symptoms: a randomized controlled trial. *Menopause* 2014;21:330-338.
26. Sternfeld B, LaCroix A, Caan BJ, et al. Design and methods of a multi-site, multi-behavioral treatment trial for menopausal symptoms: the MsFLASH experience. *Contemp Clin Trials* 2013;35:25-34.
27. 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.
28. Joreskog KG, Sorbom D. *LISREL (Version 8.8) [Computer Software]*. Chicago, IL: Scientific Software International; 2006.
29. Joreskog K. *Structural Equation With Ordinal Variables Using LISREL*. Chicago, IL: Scientific Software International; 2005:1-77.
30. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equation Model* 1999;6:1-55.
31. Kline RB. *Principles and Practice of Structural Equation Modeling*. 3rd ed. New York, NY: Guilford Press; 2010.
32. Harvey AG, Stinson K, Whitaker KL, Moskovitz D, Virk H. The subjective meaning of sleep quality: a comparison of individuals with and without insomnia. *Sleep* 2008;31:383-393.