



Published in final edited form as:

Menopause. 2015 November ; 22(11): 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, Kevin L. Rand, PhD, Carol A. Landis, PhD, RN, FAAN, Misti L. Paudel, PhD, Katherine M. Newton, PhD, Nancy Woods, PhD, RN, FAAN, and Janet S. Carpenter, PhD, RN, FAAN

School of Nursing Indiana University, Indianapolis, IN (Otte, Carpenter); Department of Psychology, Indiana University-Purdue University, Indianapolis, IN (Rand); University of Washington School of Nursing (Landis and Woods); University of Minnesota School of Public Health and Epidemiology (Paudel); Harvard Medical School, Boston, MA; Group Health Research Institute, Seattle, WA (Newton)

Abstract

Objective—Women report poor sleep quality during various stages of the menopause transition and post-menopause, especially those with hot flashes. Sleep measurements vary widely due to 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 the PSQI baseline data from three randomized controlled clinical trials enrolling peri- and postmenopausal women with hot flashes (N=849) from the Menopause Strategies: Finding Lasting Answers for Symptoms and Health (MsFLASH) network. Several a priori factor models were compared.

Results—One- 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 due to poor fit and low rates of sleep medication use. The three-factor model was examined in African American (AA) and Caucasian subsamples and 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, appears 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.

MESH Key Words

sleep; sleep quality; measurement; menopause; women; factor analysis

Introduction

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

Selecting a measure of sleep to use in research studies largely depends on the research question(s). Current self-report sleep measures typically address single 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 medication use, and (g) daytime dysfunction (daytime sleepiness, inability to function during the day, feeling tired or fatigued).³⁻⁵ Instruments address either a single or combination of the above aspects and are often specific to the overarching sleep problem being addressed. Recommendations have been made for the standardization of sleep measures in order to facilitate meta-analyses in research in patients with cancer and insomnia⁶, but a variety of instruments continue to be used.⁷

The Pittsburgh Sleep Quality Index (PSQI) is one of the most frequently used questionnaires for the assessment of self-reported sleep quality in studies of men and women with and without chronic illnesses.^{4,8} The PSQI was created to capture sleep quality in a sample of psychiatric patients. The original intent of the questionnaire was to: (a) create a reliable and valid standardized measure of sleep quality; (b) provide a tool to distinguish good versus poor sleep quality; (c) create a measure of sleep quality that was user friendly for patients, clinicians, and researchers to interpret; and (d) provide a clinical tool that assessed a list of sleep disturbances that impact sleep quality for psychiatric populations. The PSQI contains 7 individual sub-scales and a *single factor* global score of > 5 representing poor overall sleep quality. The global score of the PSQI has been deemed a simple measure for use in clinical and research to identify good versus poor sleep. The PSQI single-factor based score for sleep quality has consistently reported acceptable reliability and validity in various populations⁹⁻¹² and has been translated into several different languages.⁹

The single factor structure of the PSQI has been analyzed to understand if the global sleep quality score derived from all of the individual 7- subscale scores is the best representation of sleep quality. This research, conducted in various adult samples with (depression, breast cancer, post-renal transplantation, rheumatoid arthritis) and without chronic illnesses (community dwelling English and Spanish men and women, Nigerian students, non-depressed men and women), has yielded inconsistent results.¹³⁻¹⁸ Of seven studies, four found that a model comprising three factors (e.g. Sleep Efficiency, Perceived Sleep Quality, and Daytime Functioning) was a better fit to the data than a 1-factor global score for sleep quality.¹³⁻¹⁶ One study found a 1-factor model (Sleep Quality)¹⁹ and two studies reported 2-factor models (Sleep Efficiency, Perceived Sleep Quality) were most appropriate for the populations being studied.^{17,18} The Otte et al. study also found racial differences in a nested-model comparison of African American and Caucasian breast cancer survivors of Sleep

Quality and Sleep Latency as well as Sleep Efficiency and Sleep Quality.¹⁸ The findings from these studies raise questions about whether it is appropriate to use a single global score to differentiate good from poor sleep quality across clinical populations. There were several limitations to these previous studies that make it difficult to generalize the findings of sleep quality to other populations, such as women with hot flashes. These include sample heterogeneity (mixed gender, wide age ranges and specific focus on chronic illnesses) and varied sample sizes from 107 to 1174.¹⁸ There have been no multi-group analyses using factor modeling to determine if mixed sample groups limit the factor findings.¹⁵ In addition, findings have been limited to research studies and it is unclear how findings can be translated into clinical practice.

Since sleep is a significant problem in women during the menopause transition, it would be informative to evaluate whether the single factor measure from the PSQI compared to the previously reported Otte et al. 2-factor model and Cole et al. 3-factor models best reflects good versus poor sleep quality in a large sample of relatively healthy midlife women in late perimenopause and early post-menopause with hot flashes.^{13–15} Additional nested model comparisons between African American and Caucasian women can identify potential racial differences which 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 to analyze 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 control trial of escitalopram versus placebo for the treatment of hot flashes.^{20–22} Study 2 was a 3 by 2 factorial design study that yielded 3 comparison groups: (a) yoga versus usual activity; (b) exercise versus usual activity; and (c) omega 3 fatty acids versus placebo pill for hot flashes.^{23–26} Study 3 was a randomized control trial of low-dose estradiol versus placebo and venlafaxine XR versus placebo for menopausal symptoms.²⁷ Data from a total of 899 midlife women from these 3 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 aged-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 and income level, employment status, height, and weight.

Sleep was assessed using the PSQI,⁴ a 19-item scale that provides 7 component scores (ranges 0–3): sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbance, use of sleep medications, and daytime dysfunction. The 7 component scores are combined to produce a global sleep quality index score. The global scores range from 0–21, with scores above 5 reflecting poor global sleep quality⁴ and scores above 8 reflecting poor sleep and high daytime fatigue burden.⁹

Statistical Analysis

In order to determine the best factor model and scoring method for the PSQI among women with hot flashes, a confirmatory factor analysis using LISREL 8.8²⁸ was conducted. Because the single global score of the PSQI is frequently used in the literature, we examined the fit of a 1-factor model. Because recent studies (Cole et al.¹⁵ and Otte et al.¹⁸) found support both for 2-factor and 3-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.

The subscale scores from the 7 PSQI components represent ordinal rather than continuous data and should not be analyzed using methods that assume they are continuous variables with metric properties.²⁹ Instead, PRELIS 2.8²⁸ was used to estimate the polychoric correlations among the subscale scores.²⁹ Polychoric correlations are used to describe the associations between observed ordinal variables (i.e., PSQI subscale scores) that represent underlying phenomena (i.e., sleep characteristics) that are normally distributed and continuous. These polychoric correlations were then used to create an asymptotic covariance matrix (similar to a correlation matrix) that represents all of the inter-relationships among the PSQI subscales. It is this matrix of associations among the PSQI subscales that was used to examine the various factor models (using the weighted least squares approach).

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) chi-square (χ^2); (b) the Root Mean Square Error of Approximation (RMSEA) and its 90% confidence interval; and (c) the Comparative Fit Index (CFI). The χ^2 is an index of absolute fit between the hypothesized and observed covariance matrices. A non-significant χ^2 is evidence of acceptable fit. The 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 $\leq .06$ indicate acceptable fit.³⁰ The CFI is an incremental fit index, with values $\geq .95$ 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

The majority of the women were non-Hispanic (97.2%), Caucasian (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 of American Indian (1.2%), Asian/Pacific Islander (2.2%), and women who self-identified as “Other” (2.8%). The mean age was 54.47 years ($SD = 3.83$). The mean global PSQI score was 7.82 ($SD = 3.47$). The majority of 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 >8 indicating poor sleep quality and high fatigue burden.

Responses to the PSQI were scored into the seven subscales⁴ and 94.4% of the sample had complete baseline data on all component scores. The means, standard deviations, and correlations among the subscales are shown in Table 1. The number of missing values for each subscale ranged from 7 to 17, with a total of 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 the confirmatory factor analysis, resulting in a final sample size of 849. We compared women with complete data to women with missing data and found that there were no differences on age ($M = 54.48$ & 54.16 for complete and missing data respectively). There was a small, but significant relationship between education and missing data points; women with less education were more likely to have missing data, $r_s = -.078$, $p = 0.019$. There also was a relationship between race/ethnicity and missing data, $\chi^2(5) = 19.04$, $p = .002$; specifically, proportionately more African American women had missing data compared to Caucasian women (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% & 2.16%, respectively, $\chi^2(4) = 9.51$, $p = .049$). This could be attributed to women without bed partners not providing a response for the bed partner related item.

Confirmatory factor analysis

The sample sizes, means, standard deviations, 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 1-factor model (Model 1), consistent with the global score of the PSQI, did not show acceptable fit on any of the indices. The 2-factor model (Sleep Efficiency, Perceived Sleep Quality) (Model 2)¹⁵ also did not show acceptable fit on any index. A modified 2-factor model (Model 3)¹⁸ showed better fit, with the CFI indicating acceptable fit. The 3-factor model (Sleep Efficiency, Perceived Sleep Quality; Daily Disturbances) (Model 4)¹⁵ showed the best fit, with acceptable values on the RMSEA and CFI.

The 3-factor model (Model4)¹⁵ is presented in Figure 1 with its standardized coefficients. As the model shows, the loading from Perceived Sleep Quality to the sleep medication use subscale is small ($\beta = .15$), which leaves a large amount of unexplained variability in this

item ($\delta = .98$). 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 ($r = .15$). This is likely due to a restriction of range, as 68% of the sample reported using no sleep medication at all.

Given the low rate of sleep medication use and the resultant weak correlations between it and the other PSQI subscales, we examined an alternative 4-factor model where Sleep Medication Use was modeled as a separate factor along with Perceived Sleep Quality, Sleep Efficiency, and Daily Disturbance. This model showed good fit to the data (Model 5 in Table 2). However, the Sleep Medication Use factor did not correlate strongly with any of the other factors ($\psi = .03-.18$), which raises the question of utility of this item in assessing sleep quality among women with hot flashes.

Consequently, we examined the fit of an alternative 3-factor model (Sleep Efficiency, Perceived Sleep Quality, and Daily Disturbance), excluding the sleep medication use item, and this model showed good fit to the data across all indices, including a non-significant χ^2 (Table 2; Model 6). Therefore, we concluded that this 3-factor conceptualization represented the best model of sleep quality among women with hot flashes (see Figure 2).

To examine whether the factor structure was consistent between African American and Caucasian 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 Caucasian ($n = 530$) subsamples had sufficient numbers to allow for statistical comparisons of the factor models. First, a two-group, 3-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 the factor loadings to be equal across both groups, and the resulting 3-factor model showed significantly worse fit to the data, $\chi^2 (df = 6) = 23.81, p = 0.0005$,

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 .72 and .87 for Caucasian and African American women, respectively. This suggests that subjective sleep quality subscale is a slightly stronger indicator of Perceived Sleep Quality factor in African American women compared to Caucasian 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 .73 and .44 for Caucasian and African American women, respectively.. This indicates that among menopausal women, habitual sleep efficiency subscale is a stronger indicator of Sleep Efficiency factor for Caucasian women compared to African American women. Although we were able to detect two differences in the factor loadings of African American compared to Caucasian 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 are 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 Caucasian women.

Next, we constrained the correlations among the three factors to be equal between groups to test if the factor intercorrelations differed in African American and Caucasian women. The resulting model showed no significant difference in fit, $\chi^2_{(df=3)} = 5.34$, $p = 0.1485$, indicating that the correlations among Sleep Efficiency, Perceived Sleep Quality, and Daily Disturbance were the same in African American and Caucasian women. This result suggests that the structure of sleep quality is similar in African American and Caucasian 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.

DISCUSSION

The results of this study showed that sleep quality, as measured by the PSQI, in midlife women in the late menopause transition and early postmenopause with daily hot flashes is multifaceted with three correlated factors: Sleep Efficiency, Perceived Sleep Quality, and Daily Disturbance. This 3-factor model fit well in both African American and Caucasian 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 Caucasian women.

Overall, results of these analyses in midlife women with hot flashes are consistent with previous reports analyzing a three-, two- and one factor models of sleep quality. The nested model comparisons by race show no major differences as with the previous Otte et al. study.¹⁸ The model results are similar to previous factor analyses that also found that a 3-factor model (Sleep Efficiency, Perceived Sleep Quality, Daily Disturbances) best fit the data in depressed and non-depressed adults,¹⁵ 520 Nigerian university students,¹³ 135 post-renal transplant patients,¹⁴ and 3,667 community-dwelling English and Spanish speaking Hispanic and non-Hispanic adults.¹⁶ Conversely, the traditional 1-factor (Sleep Quality) and 2-factor models (Sleep Efficiency; Perceived Sleep Quality) consisted of samples of 197 Chinese women with breast cancer,¹⁹ 417 depressed and non-depressed older adults,¹⁵ 107 patients with rheumatoid arthritis,¹⁷ and 1174 non-depressed breast cancer survivors.¹⁸ Two of the prior studies also produced a better model fit by removing the sleep medication item during the analysis.^{16,17} The findings from these various studies show that the PSQI factor structure differ among studies that have different mixed genders, race, and chronic 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 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 structure can delineate improvements in more specific areas such as sleep efficiency or perceived sleep quality.

Although we found that the factor structure and factor inter-correlations were consistent between African American and Caucasian women, nested model comparisons did reveal 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 to Caucasian women. This difference is important from a measurement and research standpoint. However, it is unlikely to be of any practical or clinical significance.³² For both African American and Caucasian women, subjective sleep quality is the best indicator of Perceived Sleep Quality. Second, we found that assessing the Sleep Efficiency factor differed significantly in African American versus Caucasian women. Specifically, we found that Sleep Efficiency among African American women was primarily captured by the sleep duration subscore. In contrast, the habitual sleep efficiency subscale was a stronger indicator of Sleep Efficiency in Caucasian 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 item may not be a good indicator of sleep quality among women participating in clinical trials for hot flashes, likely due to low use of sleep medication in this population. This is a consistent finding across the 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 (e.g., 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 medications to help with sleep among women. Although it is clinically important to query patients regarding sleep medication use, it is recommended that the item be removed from the PSQI sub-scale scoring until further validity and reliability testing is performed.

The results from this study and the previous factor analyses highlight the fact that the definition of sleep quality varies among researchers, clinicians and patients 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 and without insomnia found that perceived sleep quality is a multifaceted concept that includes: (a) being tired upon waking that lasts throughout the day; (b) the feeling of being rested upon waking; and (c) number of awakenings during the night of sleep.³² It could be suggested that assessing patient'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 in late menopause transition and early postmenopausal experiencing hot flashes limiting generalizability to all midlife women of this age range and menopausal status.

CONCLUSION

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, chronic illness, and gender, making cross-study comparisons of sleep a continued challenge.

Based on the results of this study, alternative PSQI scoring could be considered and further evaluated using the 3-factor model for self-reported sleep in studies of midlife women with hot flashes. Revisions would consist of recalculations of subscale loadings into Sleep Efficiency (sleep disturbance and habitual sleep efficiency subscale), Perceived Sleep Quality (subjective sleep quality and sleep latency subscales, and Daily Disturbance (sleep disturbance and daytime dysfunction subscale). 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 for measuring sleep quality. Improving understanding of the larger concept of sleep quality can further research efforts to facilitate better subjective measurement that can be used for intervention testing that can improve sleep for women during the menopause transition.

Acknowledgments

Conflicts of interest and Source of Funding:

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, 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.

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 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 (Garnet L. Anderson, PhD, was previously a PI) and 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.

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: [Mar 21-2 2005]
3. Berger AM. Patterns of fatigue and activity and rest during adjuvant breast cancer chemotherapy. *Oncol. Nurs. Forum.* 1998 Jan-Feb;25(1):51-62. [PubMed: 9460773]

4. Buysse DJ, Reynolds CF 3rd, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Res.* 1989 May; 28(2): 193–213. [PubMed: 2748771]
5. Miaskowski C, Lee KA. Pain, fatigue, and sleep disturbances in oncology outpatients receiving radiation therapy for bone metastasis: A pilot study. *Journal of Pain & Symptom Management.* 1999 May; 17(5):320–332. [PubMed: 10355211]
6. Buysse DJ, Ancoli-Israel S, Edinger JD, Lichstein KL, Morin CM. Recommendations for a standard research assessment of insomnia. *Sleep.* 2006 Sep; 29(9):1155–1173. [PubMed: 17040003]
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 Nov; 32(6):E98–E126. [PubMed: 16270104]
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 Dec 15; 4(6):563–571. [PubMed: 19110886]
9. Carpenter JS, Andrykowski MA. Psychometric evaluation of the Pittsburgh Sleep Quality Index. *J. Psychosom. Res.* 1998 Jul; 45(1):5–13. [PubMed: 9720850]
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 Nov; 43(11):1317–1318. [PubMed: 7594173]
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 Sep; 53(3): 737–740. [PubMed: 12217446]
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 Feb; 27(2):140–148. [PubMed: 15157038]
13. Aloba OO, Adewuya AO, Ola BA, Mapayi BM. Validity of the Pittsburgh Sleep Quality Index (PSQI) among Nigerian university students. *Sleep Med.* 2007 Apr; 8(3):266–270. [PubMed: 17368977]
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 and Biological Rhythms.* 2010 October; 8(4):274–281. 2010.
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 Jan 1; 29(1):112–116. [PubMed: 16453989]
16. Tomfohr LM, Schweizer CA, Dimsdale JE, Loreda 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 Jan 15; 9(1):61–66. [PubMed: 23319906]
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):1–12. [PubMed: 23390921]
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.* 2012 Aug 24.
19. Ho RT, Fong TC. Factor structure of the Chinese version of the Pittsburgh Sleep Quality Index in breast cancer patients. *Sleep Med.* 2014 May; 15(5):565–569. [PubMed: 24759325]
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 Jan 19; 305(3):267–274. [PubMed: 21245182]
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 Jan; 21(1):45–58. [PubMed: 23760428]
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 Jun; 19(6):664–671. [PubMed: 22228321]

23. Cohen LS, Joffe H, Guthrie KA, et al. Efficacy of omega-3 for vasomotor symptoms treatment: a randomized controlled trial. *Menopause*. 2014 Apr; 21(4):347–354. [PubMed: 23982113]
24. Newton KM, Reed SD, Guthrie KA, et al. Efficacy of yoga for vasomotor symptoms: a randomized controlled trial. *Menopause*. 2014 Apr; 21(4):339–346. [PubMed: 24045673]
25. Sternfeld B, Guthrie KA, Ensrud KE, et al. Efficacy of exercise for menopausal symptoms: a randomized controlled trial. *Menopause*. 2014 Apr; 21(4):330–338. [PubMed: 23899828]
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. *Contemporary clinical trials*. 2013 May; 35(1):25–34. [PubMed: 23462342]
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 internal medicine*. 2014 Jul 1; 174(7):1058–1066. [PubMed: 24861828]
28. Joreskog, KG.; Sorbom, D. LISREL (Version 8.8) [Computer software]. Chicago: Scientific Software International; 2006.
29. Joreskog, K. Structural equation with ordinal variables using LISREL. Chicago: Scientific Software International; 2005. p. 1-77.
30. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*. 1999; 6:1–55.
31. Kline, RB. Principles and practice of structural equation modeling. 3rd. ed.. New York: 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 Mar; 31(3):383–393. [PubMed: 18363315]

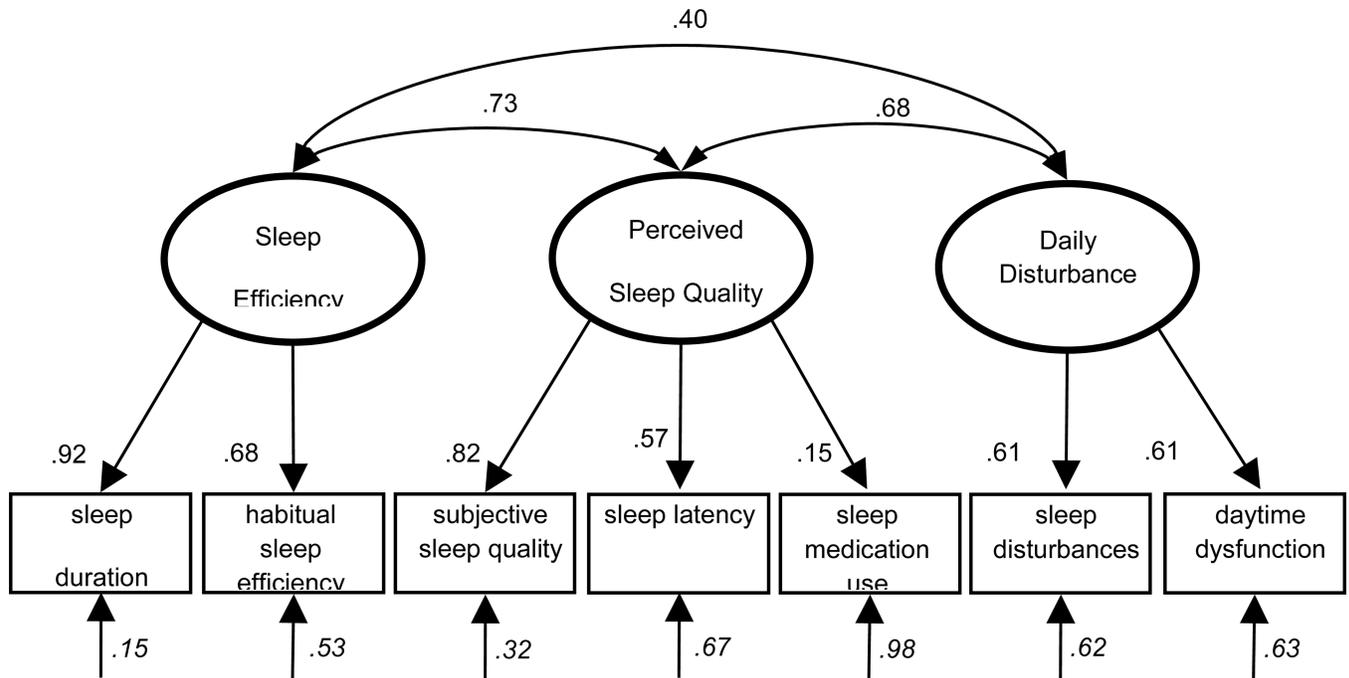


Figure 1. 3-Factor Model of Sleep Quality Measured with the Pittsburgh Sleep Quality Index in Midlife Women¹⁵

Note: All coefficients are standardized and significant at $p < .05$, 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.

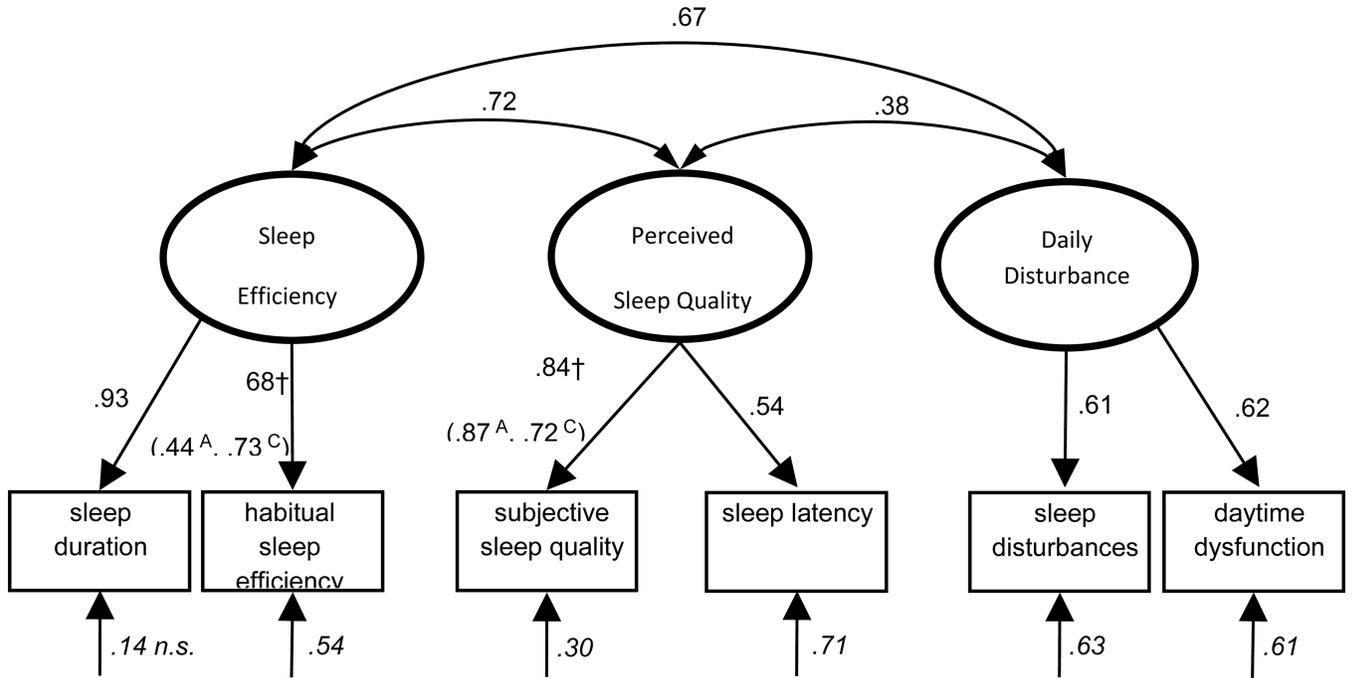


Figure 2. Alternative 3-Factor Model of Sleep Quality Measured with the Pittsburgh Sleep Quality Index in Midlife Women without Sleep Medication Use Item
 Note: All coefficients are standardized and significant at $p < .05$, unless otherwise noted. Ovals = latent variables. Rectangles = measured variables. 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 Caucasian subsamples. ^A = parameter estimate for African American subsample; ^C = parameter estimate for Caucasian subsample.

Complete Data Sample Sizes, Means, Standard Deviations, and Correlations of 7 Pittsburgh Sleep Quality Index Component Scores

Table 1

	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	.39**	--				
3. sleep duration	890	0.83	0.93	.46**	.32**	--			
4. habitual sleep efficiency	888	0.85	1.06	.30**	.27**	.55**	--		
5. sleep disturbances	892	1.64	0.57	.28**	.17**	.15**	.16**	--	
6. sleep medication use	884	0.61	1.03	.07*	.15**	.01	.03	.10**	--
7. daytime dysfunction	882	0.97	0.75	.31**	.16**	.19**	.08**	.29**	.06

Score ranges 0–3.

Note: Ns for correlations vary between 875 and 890 due to missing data.

* p < .05;

** p < .01

Table 2
Fit Indices for Different Factor Models of the Pittsburgh Sleep Quality Index

Model	Factors	χ^2	df	p	RMSEA	90% CI: RMSEA	CFI
1	1-factor (Overall Sleep Quality)	107.84	14	<0.00001	0.089	0.074–0.100	0.90
2	2-factor (Sleep Efficiency, Perceived Sleep Quality) ¹⁵	62.71	13	<0.00001	0.067	0.051–0.084	0.94
3	2-factor (Sleep Efficiency, Perceived Sleep Quality) ¹⁸	41.97	8	<0.00001	0.071	0.051–0.093	0.96
4	3-factor (Sleep Efficiency, Perceived Sleep Quality; Daily Disturbances) ¹⁵	31.92	11	0.00079	0.047	0.029–0.067	0.98
5	4-factor (Sleep Efficiency, Perceived Sleep Quality; Daily Disturbances, Sleep Medication Use)	25.31	9	0.00265	0.046	0.025–0.068	0.98
6	3-factor (Sleep Efficiency, Perceived Sleep Quality; Daily Disturbances; without sleep medication use indicator) ¹⁵	12.27	6	0.05600	0.035	0.000–0.063	0.99
6a	3-factor (Sleep Efficiency, Perceived Sleep Quality; Daily Disturbances; without sleep medication use indicator (2 group)) ¹⁵	23.08	12	0.02709	0.048	0.016–0.078	0.99

Chi Square Tests Comparing Factor Loadings of the Pittsburgh Sleep Quality Index in African America and Caucasian Women

Table 3

Factor	Indicator	λ^A	λ^C	χ^2	df	p
Sleep Efficiency	sleep duration	0.95	0.86	0.81	1	0.368
	habitual sleep efficiency	0.44	0.73	14.66	1	<0.001
Perceived Sleep Quality	subjective sleep quality	0.87	0.72	4.21	1	0.040
	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

Note: λ^A = factor loading in African American subsample; λ^C = factor loading in Caucasian subsample