



Published in final edited form as:

Climacteric. 2011 February ; 13(6): 171–180. doi:10.3109/13697131003717070.

Modeling hot flushes and quality of life in breast cancer survivors

K. L. Rand, J. L. Otte^{*}, D. Flockhart[†], D. Hayes[†], A. M. Storniolo[†], V. Stearns^{**}, N. L. Henry[†], A. Nguyen[†], S. Lemler[†], J. Hayden[†], S. Jeter^{**}, and J. S. Carpenter^{*}

Department of Psychology, Indiana University-Purdue University, Indianapolis

Abstract

Objectives—To evaluate the relationships among measures of hot flushes, perceived hot flush interference, sleep disturbance, and measures of quality of life while controlling for potential covariates (patient and treatment variables).

Methods—Breast cancer survivors ($n = 395$) due to receive aromatase inhibitor therapy provided demographic information, physiological hot flush data via sternal skin conductance monitoring, hot flush frequency via written diary and electronic event marker, hot flush severity and bother via written diary, and questionnaire data via the Hot Flash Related Daily Interference Scale, Pittsburgh Sleep Quality Index, the EuroQOL, Hospital Anxiety and Depression Scale and the Center for Epidemiologic Studies Depression Scale.

Results—Confirmatory factor analysis supported a two-factor model for hot flush symptoms (frequency and severity). Although there was strong convergence among self-reported hot flush measures, there was a high degree of unexplained variance associated with physiological measures. This suggests that self-report and physiological measures do not overlap substantially. The structural model showed that greater hot flush frequency and severity were directly related to greater perceived interference with daily life activities. Greater perceived interference, in turn, directly predicted greater sleep disruption, which predicted lower perceived health state and more symptoms of anxiety and depression.

Conclusions—Findings suggest hot flush interference may be the most appropriate single measure to include in clinical trials of vasomotor symptom therapies. Measuring and ameliorating patients' perceptions of hot flush interference with life activities and subjective sleep quality may be the most direct routes to improving quality of life.

Keywords

Hot Flushes; Menopause; Breast Cancer; Quality of Life; Structural Model

Introduction

A combination of factors underscores the growing need for evidence-based recommendations for hot flush measurement in clinical trials. First, the sheer number of

Correspondence: Dr K. L. Rand, Department of Psychology, IUPUI, 402 North Blackford Street, LD 124, Indianapolis, IN 46202, USA.

^{*}School of Nursing, Indiana University

[†]School of Medicine, Indiana University

[‡]Breast Oncology Program, University of Michigan Comprehensive Cancer Center

^{**}Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins School of Medicine, USA

Conflict of interest: Nil.

women reaching menopause and experiencing hot flushes has increased tremendously due to demographic shifts occurring with the aging Baby Boomer population of the United States¹. In addition, interest in establishing the efficacy of non-hormonal treatments has peaked following publication of findings from the Women's Health Initiative indicating a shift in the risk/benefit balance of hormone therapy²⁻⁸. Also, there is growing recognition that breast cancer survivors are a particularly symptomatic group⁹⁻¹². Compared to the average woman, hot flushes of breast cancer survivors are more frequent, severe, distressing, of longer duration^{9,10} and can interfere with compliance to life-saving medications, such as tamoxifen¹³⁻¹⁵. Finally, there is growing interest in establishing evidence for non-hormonal treatments based on both subjective reports and objective sternal skin conductance monitoring to improve upon existing research which has largely focused solely on subjective measurement (see reviews, references¹⁶⁻²⁴).

There is little guidance available to direct hot flush measurement. Currently, researchers use a variety of methods such as retrospective or prospective paper diaries, questionnaires, or sternal skin conductance monitoring. Draft recommendations intended primarily for industry studies of hormonal preparations were issued in January, 2003 by the Food and Drug Administration Center for Drug Evaluation and Research²⁵. These recommendations endorse the use of frequency and severity as co-primary outcomes when studying treatment efficacy/effectiveness. However, they do not specify whether frequency should be measured subjectively or objectively (e.g. via sternal skin conductance monitoring)²⁶⁻²⁸. There were also recommendations issued following a National Institutes of Health Workshop on hot flush measurement²⁹ supporting use of a combination of objective and subjective measures. However, these do not specify which subjective measure to use (i.e. frequency, severity, bother/distress, duration, and/or interference) or how to resolve known discrepancies between subjective and objective measures²⁶⁻²⁸.

Recommendations for hot flush measurement in the clinical trial setting should be developed with the following considerations. First, understanding the extent of redundancy and amount of unexplained variance in hot flush measures could provide guidance for eliminating one or more hot flush measures to reduce participant burden. Second, understanding which hot flush measure more robustly predicts important clinical outcomes may provide some indication of the relative value of including that particular measure.

The purpose of this analysis was to explore relationships among subjective and objective hot flush measures and their relationships to quality of life measures ascertained in breast cancer survivors prior to initiation of adjuvant aromatase inhibitor therapy in the Consortium on Breast Cancer Pharmacogenetics Exemestane-Letrozole Pharmacogenetics trial. Our goal was to determine whether evidence-based recommendations for measuring objective and/or subjective hot flushes in breast cancer survivors could be made. We set out, first, to examine the relationships among objective and subjective hot flush measures in breast cancer survivors using confirmatory factor analysis with structural equation modeling (SEM), and, second, to examine the potentially differential relationships among hot flush measures, sleep disturbance, and quality of life (perceived health state, anxiety symptoms, depressive symptoms) in breast cancer survivors.

Methods

Sample and setting

Patients were recruited for a multicenter, randomized, clinical trial designed to test associations between the phenotypic effects of aromatase inhibitor therapy on surrogate markers of drug response and the presence of specific germline genetic variants in aromatase and candidate genes in estrogen metabolism and response. Patients were recruited from

clinics associated with the Indiana University Melvin and Bren Simon Cancer Center, the University of Michigan Comprehensive Cancer Center, and the Johns Hopkins/Sidney Kimmel Comprehensive Cancer Center. All procedures were approved by three levels of ethical and scientific review involving each site's local institutional review board, cancer center scientific review committee, and General Clinical Research Center advisory committee. Inclusion criteria were: (1) postmenopausal; (2) histologically proven ductal carcinoma *in situ* (stage 0) or stage I–III invasive carcinoma of the breast that was estrogen receptor- and/or progesterone receptor-positive by immunohistochemical staining and considering aromatase inhibitor therapy; (3) completed adjuvant chemotherapy if indicated; (4) due to receive an aromatase inhibitor as initial adjuvant hormonal treatment or following adjuvant tamoxifen; (5) ECOG performance status 0–2; (6) aware of the nature of the diagnosis, understanding of the study regimen, its requirements, risks, and discomforts; and (7) able and willing to provide informed consent.

Study procedures

Eligible and interested subjects provided written informed consent and authorization to release medical information. Each subject completed a baseline visit (before the start of aromatase inhibitor) and subsequent visits 1, 3, 6, 12 and 24 months later. Hot flush assessments were completed at each time point except 24 months. Subjects completed quality-of-life questionnaires at each time point. A trained study staff member instructed subjects on the use and care of the Biolog hot flush monitor, connected each subject to the monitor at the time of their clinic visit, and administered questionnaires. For this study, we analyzed the baseline data only.

Subjects wore the monitor and reported hot flushes via diary and event button for up to 36 h. After monitoring was completed, participants followed written instructions for turning off the monitor and returning the monitor via pre-paid postal mail to each study site. Each site downloaded the hot flush data from the monitors and transmitted it to Indiana University for centralized analysis using customized software and established procedures^{9,30,31}.

Measures

Physiological hot flush frequency was assessed using sternal skin conductance monitoring^{9,26,30,31} (UFI, Model 7-day 3991 SCL, Morro Bay, CA, USA). Outcomes from file scoring included hot flush frequency (number of hot flushes per unit of time) and total magnitude or the summed degree of change in skin conductance occurring with each hot flush. Scoring was completed by trained raters with inter-rater reliability exceeding 90%.

Self-reported hot flush frequency, severity, and bother were assessed using written diaries and electronic event markers during the 36-h period. When the subject experienced a hot flush, she was instructed to push the two red buttons on the hot flush monitor, write down the time of the hot flush, and rate severity and bother in a paper diary (0 = not at all and 10 = extremely severe or 10 = extremely bothersome).

We only included monitor and diary data from the 93.2% of women who had at least 20 h of quality monitoring. For women with less than 20 h of quality monitoring, their data were treated as missing. Because the amount of quality monitoring could vary from 20 to 36 h, all hot flush variables were pro-rated for a 24-h period. We used the following formula: $24 \text{ h} \times (\text{frequency or severity or bother during the monitoring period} / \text{number of quality hours of the monitoring period})$. For example, if a woman had 18 hot flushes during the 36 h of monitoring, we calculated a 24-h frequency of 12 hot flushes ($24 \times (18/36) = 12$ hot flushes in 24 h).

The *Hot Flash Related Daily Interference Scale* (HFRDIS) was used to measure perceived hot flush interference. This ten-item scale measures the degree to which hot flushes interfere with nine daily life activities (work, social activities, leisure activities, sleep, mood, concentration, relations with others, sexuality, and enjoyment of life) and overall quality of life. Participants rate the degree to which hot flushes have interfered with each item during the previous week using a 0 (do not interfere) to 10 (completely interfere) scale. A total score is computed by summing items; reliability and validity have been previously supported³⁰. Higher scores indicate greater perceived interference. Cronbach's α in this data set was 0.94.

The *Pittsburgh Sleep Quality Index* (PSQI) measures sleep quality and disturbance retrospectively over the previous month using self-reports. This 19-item scale generates a global sleep quality score and the seven component scores: sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbance, use of sleeping medications, and daytime dysfunction. PSQI items use varying response categories that include recording usual bed time, usual wake time, number of actual hours slept, number of minutes to fall asleep, as well as forced-choice Likert-type responses. Psychometric data suggest the PSQI is a valid and reliable instrument to use in women with breast cancer^{32,33}. Higher scores indicate greater sleep disturbance. Cronbach's α in this data set for the global sleep quality score was 0.72.

An overall rating of perceived health state was given using the *EuroQOL*. This Cantril-like ladder rates best health at the top (100) and worst health (death) at the bottom (0). Participants were asked to provide a single number to describe their own health state that day and a single number to describe the average woman of their own age. The latter was used as a comparison standard and was subtracted from the participant's own rating to yield a single score of relative quality of current health state. Negative values indicate that the woman perceives her health state as being worse than average, and positive values indicate that the woman perceives her health state as being better than average.

Anxiety was assessed with the anxiety subscale of the *Hospital Anxiety and Depression Scale* (HADS-A). This seven-item subscale has been well validated as a measure of anxiety^{34–36}. Higher scores indicate more anxiety symptoms. Cronbach's α in this data set was 0.78.

The *Center for Epidemiologic Studies Depression Scale* (CESD) is a 20-item self-report measure assessing presence and severity of depressive symptoms over the past week. Respondents rate each item on a four-point scale. After four positively worded items are reverse scored, responses are summed to obtain total scores ranging from 0 to 60. Scores of 16 and above are indicative of high depressive symptoms. Psychometric properties of the CESD have been extensively examined and the scale has been widely used in research, including our own breast cancer research^{37–40}. Cronbach's α in this data set was 0.87.

Data analysis

For this study, we examined the pattern of relationships and hypothesized causal connections among hot flush symptoms, perceived interference, sleep disturbance, health state, anxiety and depressive symptoms, and patient/treatment variables using structural equation modeling (SEM). SEM was conducted using LISREL 8.8⁴¹.

First, we conducted a confirmatory factor analysis of hot flushes using the following indicators: (1) monitor-recorded frequency; (2) diary-recorded frequency; (3) button-recorded frequency; (4) average monitor-recorded change in skin conductance; (5) average diary-recorded bother; and (6) average diary-recorded severity. All were pro-rated over 24 h

as described above. This allowed us to create a *measurement model* identifying abstract concepts (i.e. latent variables) based on the patterns of relationships (e.g. correlations) among a set of measured variables (i.e. indicators).

Second, we created a *structural model* to evaluate the hypothesized causal relationships among hot flush symptoms, perceived hot flush interference, sleep disturbance, perceived health state, anxiety and depressive symptoms, and patient/treatment variables.

Because the HFRDIS, PSQI, CESD and HADS-A had large numbers of items (ranging from 7 to 20), individual items were aggregated into three parceled indicators per latent variable. This was done to reduce the complexity of the structural model and increase the reliability of the indicators⁴². Each indicator reflected the mean of the constituent scale items. Details on the parceling are available from the authors. Perceived health state was modeled as a latent variable with a single indicator, the relative quality of current health state derived from the EuroQOL scale, as described above.

Several covariates previously shown to be related to hot flushes in breast cancer survivors were included in the structural model. These included: (1) minority racial or ethnic status (coded yes or no); (2) body mass index; (3) age at first breast cancer surgery; (4) previous chemotherapy (yes, no); (5) previous use of tamoxifen (yes, no); (6) alcohol use (yes, no); (7) history of ever smoking tobacco (yes, no); and (8) use of selective serotonin reuptake inhibitors (SSRI) or other hot flush treatments (yes, no). Because this latter variable included mainly SSRI users ($n = 99$ women, 25% of sample) rather than other treatment users ($n = 18$, 4.5% of sample), we evaluated SSRI use and other hot flush treatments as a combined covariate. This did not substantially change the results of the model. Therefore, we present the structural model with SSRI use modeled because it is conceptually cleaner than combining it with other hot flush treatments. All covariates were modeled as latent variables with single indicators.

Fit statistics included χ^2 and the root mean square error of approximation (RMSEA)⁴³. The former measures the absolute fit between the hypothesized model and the observed pattern of relationships among measured variables. A non-significant χ^2 suggests there is no difference between the hypothesized and observed patterns of relationships and, hence, that the hypothesized model is acceptable. The RMSEA statistic adjusts the measure of absolute fit based on the complexity of the hypothesized model, with more complex models receiving a 'penalty'. Smaller values of RMSEA indicate better model fit, with values less than 0.10 indicating acceptable fit and values less than 0.05 indicating close fit^{44,45}.

Model parameters were estimated using full information maximum likelihood estimation, which uses all the observed data from every participant and is superior to traditional methods of handling missing data (e.g. listwise deletion)⁴⁶.

Results

Sample description

The sample included 395 participants who were mainly white (84%), with smaller numbers who were African-American (8%) or Asian-American (2%). Participants were a mean age of 59.29 years (standard deviation, SD = 8.70 years) and a mean of 16.36 months past their first surgery for breast cancer (SD = 20.38 months). Participants were categorized as flushers ($n = 284$) or non-flushers ($n = 96$). Flushers had at least one hot flush on any recording method (monitor or diary or event button). Fifteen participants were missing hot flush data. Flushers were significantly younger, had significantly lower body mass index,

and reported significantly greater hot flush-related interference than their non-flusher counterparts (see Table 1).

Description of variables

Descriptive statistics for continuous-level covariates (i.e. age, body mass index), hot flushes and other variables are shown in Table 2. Consistent with previous research⁴⁷, there was a high degree of variability across women in their experiences of hot flushes, with standard deviations exceeding the means for monitor frequency, diary frequency, button frequency and perceived interference.

The bivariate correlations among study measures are shown in Table 2. The strongest correlation was between diary hot flush frequency and button hot flush frequency ($r = 0.94$), indicating that using both measures is redundant and unnecessary. Similarly, there was a strong correlation between mean bother and severity ($r = 0.85$), suggesting that these may be redundant measures.

Hot flush measurement model

Based on previous research⁴⁷, we hypothesized that hot flush symptoms were characterized by two latent variables: HF Frequency and HF Severity. We modeled monitor HF frequency, diary HF frequency, and button HF frequency as indicators of HF Frequency. We modeled mean magnitude of change in skin conductance, mean HF bother, and mean HF severity (all averaged per hot flush) as indicators of HF Severity. This initial model did not show acceptable fit to the data, $\chi^2_{(d.f. = 8)} = 40.60$, $p < 0.001$, RMSEA = 0.10. Examination of the modification indices revealed that mean magnitude of change in skin conductance was an indicator of both HF Severity and HF Frequency, and we revised the model accordingly. This new model showed close fit to the data, $\chi^2_{(d.f. = 7)} = 0.77$, $p = 0.998$, RMSEA = 0.00. The error terms associated with button HF frequency and mean HF severity, however, were not significantly different from zero. In addition, the error term for mean HF severity was negative, which is an unacceptable value. Therefore, we fixed the error terms for button HF frequency and mean HF severity to zero, and the resulting model showed close fit to the data, $\chi^2_{(d.f. = 9)} = 2.93$, $p = 0.967$, RMSEA = 0.00 (see Figure 1). As the figure shows, the latent variables of HF Frequency and HF Severity are slightly correlated ($r = 0.23$), which is consistent with previous research⁴⁷. Surprisingly, monitor HF frequency was not the strongest indicator of HF Frequency. Although monitor-registered HF frequency previously has been considered an ‘objective’ measure, this result shows it has considerable measurement error associated with it ($\theta = 0.86$). In contrast, this model suggests that button HF frequency is the strongest indicator of HF Frequency, followed closely by diary HF frequency. In fact, button HF frequency and diary HF frequency appear to be redundant indicators.

As with HF Frequency, the ‘objective’, physiological measure of HF Severity (i.e. mean magnitude change in skin conductance) had considerable measurement error ($\theta = 0.77$). Moreover, mean magnitude change in skin conductance was a stronger indicator of HF Frequency than HF Severity. Together, these results suggest that magnitude of change in skin conductance is a ‘noisy’ and therefore less useful, indicator of HF symptoms. The strongest indicator of HF Severity was mean severity, followed closely by mean bother. In fact, these indicators appear to be redundant, indicating that future studies could simply measure one.

Hot flush and quality-of-life outcomes structural model

Next we examined the overall model, including HF symptoms, perceived interference, sleep disturbance, perceived health state, and anxiety and depressive symptoms with covariates. Correlations among covariates are presented in Table 3.

The covariances among all latent variables were estimated and this model showed close fit to the data, $\chi^2_{(d.f. = 221)} = 247.62$, $p = 0.106$, RMSEA = 0.02. The factor loading of mean magnitude change in skin conductance as an indicator of HF Severity, however, was non-significant in the overall model. Therefore, we revised the hot flush symptom model such that mean magnitude change in skin conductance was an indicator of HF Frequency only, and this model showed close fit to the data, $\chi^2_{(d.f. = 222)} = 247.63$, $p = 0.114$, RMSEA = 0.02. We then added the hypothesized causal paths. Non-significant causal paths were removed from the model in a stepwise fashion until all remaining paths were significant. The final structural model is shown in Figure 2. As the figure shows, both latent variables of HF Frequency and HF Severity predicted greater perceived interference. This finding is counter to previous research that showed that only HF Severity predicted greater perceived interference⁴⁷. Hot flush symptoms had no direct relationships with measures of patient quality of life. Perceived interference predicted greater sleep disturbance ($\beta = 0.26$). In turn, sleep disturbance had strong, deleterious relationships with all measures of quality of life, predicting lower perceived health state ($\beta = -0.29$), more anxiety ($\beta = 0.41$) and more depressive symptoms ($\beta = 0.49$). Together, these results are consistent with a model in which hot flush symptoms negatively influence quality of life indirectly, through perceived interference and sleep disturbance.

Initially, African-American and Asian-American women were combined into the same minority status, due to the low number of Asian-American women ($n = 8$). As a double-check on our results, we conducted two additional analyses. First, we re-ran the structural model with African-American (yes, no) and Asian-American (yes, no) included as separate variables. This model showed significant misfit with the data, $\chi^2_{(d.f. = 235)} = 276.52$, $p = 0.033$, RMSEA = 0.02, probably due to there being so few Asian-American women in the sample. Being African-American was significantly correlated with depression ($r = -0.10$), body mass index ($r = 0.16$), and receiving chemotherapy ($r = -0.12$). Being Asian-American was significantly correlated with body mass index ($r = -0.14$) and ever-smoked ($r = -0.10$). Neither racial variable correlated with any of the hot flush variables. Second, we ran a model with just African-American (yes, no) as a variable (grouping the Asian-American women with the White women). This model fit the data, $\chi^2_{(d.f. = 232)} = 254.58$, $p = 0.066$, RMSEA = 0.02. Being African-American showed the same correlations with depression, body mass index, and chemotherapy as in the previous model. In both of the new models, all of the path coefficients among the main study variables (i.e. hot flush frequency, severity, interference, sleep, depression, anxiety, health state) were the same as the original structural model.

Discussion

The measurement model identified redundancy in hot flush measures and suggests that measurement could be simplified to include (1) either diary or button frequency (but not both) and (2) either diary severity or diary bother (but not both). However, the redundancy in these variables could be an artifact of the measurement methods. Both diary and button frequency were recorded at the same time using prospective real-time recording and participants probably complied with both measures to the same extent³¹. Similarly, both severity and bother were assessed using ten point numeric rating scales in a prospective, real-time manner. Our findings using these methods do not reveal whether similar redundancies would be seen across retrospective and prospective methods or with different rating scales.

The measurement model confirms the importance of studying the separate effects of hot flush frequency and severity. This finding is consistent with previous research⁴⁷ where these separate dimensions independently predicted perceived interference, but is in contrast to strategies used by Loprinzi and colleagues²⁴, who combine frequency and severity to create a single variable (i.e. hot flush score). Our results suggest that frequency and severity may have separate predictors and sequelae that merit further investigation.

Based on the structural model, perceived hot flush interference may be the most appropriate single measure to use in treatment studies aimed at decreasing hot flushes and improving quality of life. Hot flush-related daily interference was the most proximal measure to sleep, health state, anxiety, and depressive symptoms. Both hot flush severity and frequency predicted greater perceived interference, which is counter to our previous finding that perceived interference was predicted only by hot flush severity⁴⁷. In turn, greater hot flush interference mediated through sleep disturbance predicted lower perceived health state and more symptoms of anxiety and depression, which is consistent with previous research on hot flushes, sleep disturbance and psychological symptoms⁴⁸. Other hot flush measures were not directly linked to these other outcomes, which is consistent with previous research showing that sleep disturbance mediates the relationship between physical symptoms and quality of life. For example, Beck and colleagues found that sleep disturbance partially mediated the relationship between pain and fatigue among cancer patients⁴⁹.

The structural model also highlights the central importance that sleep disturbance plays in these women's lives. Sleep disturbance was the strongest predictor of quality of life, having associations ranging from 1.5 to 3.5 times stronger than the next strongest predictor of any outcome. In addition, the strongest predictor of sleep disturbance was hot flush interference, which is consistent with the notion that stressors negatively impact well-being primarily to the extent that they interfere with efforts to achieve life goals^{50,51}. It is also worth noting that hot flush symptoms did not directly predict sleep disturbance, but did so indirectly through perceived interference. Although this finding contrasts with studies showing that women report that hot flushes disturb sleep⁵², it is consistent with research showing that, when measured objectively, hot flushes may not be directly associated with quality of sleep^{48,53,54}. Instead, it appears that women experiencing hot flushes are more likely to report *subjective* dissatisfaction with their sleep^{53,54}, suggesting that cognitive appraisals about the impact of hot flushes are a central mechanism of influence on sleep quality. Consistent with this possibility, research on pharmacological interventions for hot flushes frequently observes a substantial placebo effect²⁴. Therefore, interventions aimed at changing the way hot flush symptoms are appraised (e.g. cognitive-behavioral therapy) may prove beneficial in improving subjective sleep quality and quality of life.

Because SSRIs are commonly used to treat a variety of issues among breast cancer survivors (e.g. hot flushes, anxiety and depressive symptoms), it is worth noting our findings related to SSRI use in the present sample. Surprisingly, SSRI use was not directly related to hot flushes or depressive symptoms. Moreover, SSRI use paradoxically predicted greater perceived hot flush interference, greater sleep disturbance, and more anxiety symptoms, possibly as a reflection of their treatment indications (SSRI use in more symptomatic women) or side-effect profile (SSRI use causing sleep disturbances)⁵⁵. However, this study was not designed to test the efficacy of SSRIs in treating hot flush symptoms, and these results should be interpreted with caution. Further research is needed to examine the relationships among hot flushes and quality of life indicators in populations receiving treatment for hot flush symptoms.

Unfortunately, neither the measurement nor the structural model reveals which measure serves as the best phenotype for genotype–phenotype association studies. In the first model,

the physiological measures were characterized by considerable measurement error. This finding is inconsistent with our previous work that assumed physiological measures were ‘objective’ or error-free. In the present study, however, we had more hot flush indicators available, allowing us to make fewer assumptions in constructing the model. As a result, we discovered that self-reported hot flush measures show strong convergence, but do not overlap substantially with physiological hot flush measures. On the one hand, findings suggest physiological measures may not be an appropriate phenotype as they are ‘noisy’ and less reliable indicators of hot flush symptoms than self-report. For example, skin conductance magnitude has been shown to be a poor indicator of hot flush severity or bother⁵⁶. On the other hand, physiological measures may be a more appropriate phenotype because they are less influenced by factors other than the physiological experience of hot flushes. For example, women may sleep through hot flushes that occur at night and fail to report them. In addition, personality traits, such as optimism, may cause some women to discount physiological phenomena⁵⁷, resulting in underreporting of hot flush symptoms. Indeed, in the present study, the mean level of monitor-recorded hot flush frequency was significantly greater than either diary-recorded ($t(334) = 4.05, p < 0.001$) or button-recorded frequency ($t(340) = 4.20, p < 0.001$, paired-sample t tests). It is important to note, however, that the substantial measurement error associated with physiological indicators in the present model does not confirm that self-report measures are more accurate than physiological ones. It merely reflects the lack of overlap between the two. Our future planned analyses of the original study aims as funded will carefully evaluate associations between genotypes and both subjective and objective hot flush phenotypes.

There are limitations to this study worth noting. First, this study focused on patient-reported quality-of-life measures. Outcomes in other domains (e.g. interpersonal relationships) or assessed by different means (e.g. clinician ratings) may show different relationships with hot flush symptoms and merit further investigation. Second, this study was limited to American breast cancer survivors, so the findings may not generalize to other populations. For example, it may be that hot flush symptoms and their relationships with quality of life are different for healthy women compared to breast cancer survivors. Indeed, prior research indicates hot flushes in breast cancer survivors are more frequent, severe and bothersome than in healthy women without cancer⁹. However, the physiological occurrence of the hot flush seems to be the same in both groups of women^{58,59}. Finally, the data analyzed were cross-sectional, and longitudinal research is needed to better understand the nature of these relationships over time.

Conclusions

Study findings provide an empirical basis for hot flush measurement recommendations by suggesting that (1) diary *or* button frequency be used, (2) diary bother *or* severity be used, (3) perceived interference may be the most appropriate single measure of therapies designed to decrease hot flushes and improve quality of life, and (4) further studies should explore the best hot flush phenotype to use for evaluating genotype–phenotype associations.

Acknowledgments

Source of funding: Supported in part by a Pharmacogenetics Research Network Grant # U-01 GM61373 that supports the Consortium on Breast Cancer Pharmacogenomics (COBRA), a Clinical Pharmacology training grant 5T32-GM08425 from the National Institute of General Medical Sciences, National Institutes of Health, Bethesda, MD, Indiana University GCRC grant M01RR00750 from the National Institutes of Health, Bethesda, MD, University of Michigan GCRC grant M01-RR00042 from the National Institutes of Health, Bethesda, MD, Johns Hopkins University School of Medicine GCRC grant M01-RR00052 from the National Institutes of Health, Bethesda, MD, and by the Fashion Footwear Charitable Foundation of New York/QVC Presents Shoes on Sale™.

References

1. Hobbs, F.; Stoops, N. Census 2000 Special Reports. Washington, DC: US Census Bureau; Demographic Trends in the 21st Century. CENSR-4
2. Cauley JA, Robbins J, Chen Z, et al. Effects of estrogen plus progestin on risk of fracture and bone mineral density: the Women's Health Initiative randomized trial. *JAMA* 2003;290:1729–38. [PubMed: 14519707]
3. Chlebowski RT, Hendrix SL, Langer RD, et al. Influence of estrogen plus progestin on breast cancer and mammography in healthy postmenopausal women: The Women's Health Initiative randomized trial. *JAMA* 2003;289:3243–53. [PubMed: 12824205]
4. Chlebowski RT, Wactawski-Wende J, Ritenbaugh C, et al. Estrogen plus progestin and colorectal cancer in postmenopausal women. *N Engl J Med* 2004;350:991–1004. [PubMed: 14999111]
5. Cushman M, Kuller LH, Prentice R, et al. Estrogen plus progestin and risk of venous thrombosis. *JAMA* 2004;292:1573–80. [PubMed: 15467059]
6. Manson JE, Hsia J, Johnson KC, et al. Estrogen plus progestin and the risk of coronary heart disease. *N Engl J Med* 2003;349:523–34. [PubMed: 12904517]
7. Rossouw JE, Anderson GL, Prentice RL, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. *JAMA* 2002;288:321–33. [PubMed: 12117397]
8. Shumaker SA, Legault C, Rapp SR, et al. Estrogen plus progestin and the incidence of dementia and mild cognitive impairment in postmenopausal women. The Women's Health Initiative Memory Study: a randomized controlled trial. *JAMA* 2003;289:2651–62. [PubMed: 12771112]
9. Carpenter JS, Johnson D, Wagner L, Andrykowski M. Hot flashes and related outcomes in breast cancer survivors and matched comparison women. *Oncol Nurs Forum* 2002;29:E16–25. [PubMed: 11979290]
10. Harris PF, Remington PL, Trentham-Dietz A, Allen CI, Newcomb PA. Prevalence and treatment of menopausal symptoms among breast cancer survivors. *J Pain Symptom Manage* 2002;23:501–9. [PubMed: 12067774]
11. Lo Presti A, Ruvolo G, Gancitano RA, Cittadini E. Ovarian function following radiation and chemotherapy for cancer. *Eur J Obstet Gynecol Reprod Biol* 2004;113(Suppl 1):S33–40. [PubMed: 15041128]
12. North American Menopause Society. *Menopause Practice: A Clinician's Guide*. Cleveland, OH: North American Menopause Society; 2004.
13. Lash TL, Fox MP, Westrup JL, Fink AK, Silliman RA. Adherence to tamoxifen over the five-year course. *Breast Cancer Res Treat* 2006;99:215–20. [PubMed: 16541307]
14. Kahn KL, Schneider EC, Malin JL, Adams JL, Epstein AM. Patient centered experiences in breast cancer: predicting long-term adherence to tamoxifen use. *Med Care* 2007;45:431–9. [PubMed: 17446829]
15. Demissie S, Silliman RA, Lash TL. Adjuvant tamoxifen: predictors of use, side effects, and discontinuation in older women. *J Clin Oncol* 2001;19:322–8. [PubMed: 11208822]
16. Nelson HD, Vesco KK, Haney E, et al. Nonhormonal therapies for menopausal hot flashes: systematic review and meta-analysis. *JAMA* 2006;295:2057–71. [PubMed: 16670414]
17. Nedrow A, Miller J, Walker M, Nygren P, Huffman LH, Nelson HD. Complementary and alternative therapies for the management of menopause-related symptoms: a systematic evidence review. *Arch Intern Med* 2006;166:1453–65. [PubMed: 16864755]
18. Carroll DG. Nonhormonal therapies for hot flashes in menopause. *Am Fam Physician* 2006;73:457–64. [PubMed: 16477892]
19. Bruno D, Feeney KJ. Management of postmenopausal symptoms in breast cancer survivors. *Semin Oncol* 2006;33:696–707. [PubMed: 17145350]
20. Low Dog T. Menopause: a review of botanical dietary supplements. *Am J Med* 2005;118 2(12): 98–108S. [PubMed: 16414334]
21. Carpenter JS. State of the science: Hot flashes and cancer. Part 2: Management and future directions. *Oncol Nurs Forum* 2005;32:969–78. [PubMed: 16136195]

22. Stearns V, Loprinzi CL. New therapeutic approaches for hot flashes in women. *J Support Oncol* 2003;1:11–21. [PubMed: 15352639]
23. Wymenga AN, Sleijfer DT. Management of hot flushes in breast cancer patients. *Acta Oncol* 2002;41:269–75. [PubMed: 12195746]
24. Sloan JA, Loprinzi CL, Novotny PJ, Barton DL, Lvasseur BI, Windschitl H. Methodologic lessons learned from hot flash studies. *J Clin Oncol* 2001;19:4280–90. [PubMed: 11731510]
25. US Department of Health and Human Services (DHHS). Estrogen and estrogen/progestin drug products to treat vasomotor symptoms and vulvar and vaginal atrophy symptoms – recommendations for clinical evaluation: Draft guidance. Rockville: Food and Drug Administration Center for Drug Evaluation and Research (CDER); Jan. 2003
26. Carpenter JS, Monahan PO, Azzouz F. Accuracy of subjective hot flush reports compared with continuous sternal skin conductance monitoring. *Obstet Gynecol* 2004;104:1322–6. [PubMed: 15572497]
27. Sievert LL, Freedman RR, Garcia JZ, et al. Measurement of hot flashes by sternal skin conductance and subjective hot flash report in Puebla, Mexico. *Menopause* 2002;9:367–76. [PubMed: 12218726]
28. Savard J, Davidson JR, Ivers H, et al. The association between nocturnal hot flashes and sleep in breast cancer survivors. *J Pain Symptom Manage* 2004;27:513–22. [PubMed: 15165649]
29. Miller HG, Li RM. Measuring hot flashes: Summary of a National Institutes of Health workshop. *Mayo Clin Proc* 2004;79:777–81. [PubMed: 15182093]
30. Carpenter JS. The Hot Flash Related Daily Interference Scale: a tool for assessing the impact of hot flashes on quality of life following breast cancer. *J Pain Symptom Manage* 2001;22:979–89. [PubMed: 11738160]
31. Carpenter JS, Andrykowski MA, Freedman RR, Munn R. Feasibility and psychometrics of an ambulatory hot flash monitoring device. *Menopause* 1999;6:209–15. [PubMed: 10486790]
32. Carpenter JS, Andrykowski MA. Psychometric evaluation of the Pittsburgh Sleep Quality Index. *J Psychosom Res* 1998;45(1 Spec No):5–13. [PubMed: 9720850]
33. 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;28:193–213. [PubMed: 2748771]
34. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983;67:361–70. [PubMed: 6880820]
35. Snaith RP. The Hospital Anxiety and Depression Scale. *Health Qual Life Outcomes* 2003;1:29. [PubMed: 12914662]
36. Jacobsen PB, Donovan KA, Trask PC, et al. Screening for psychologic distress in ambulatory cancer patients. *Cancer* 2005;103:1494–502. [PubMed: 15726544]
37. Carpenter JS, Andrykowski MA, Wilson J, et al. Psychometrics for two short forms of the Center for Epidemiologic Studies-Depression Scale. *Issues Ment Health Nurs* 1998;19:481–94. [PubMed: 9782864]
38. Radloff IS. The CES-D Scale: A self-report depression scale for research in the general population. *Appl Psych Meas* 1977;1:385–401.
39. McCallum J, Mackinnon A, Simons L, Simons J. Measurement properties of the Center for Epidemiological Studies Depression Scale: an Australian community study of aged persons. *J Gerontol B Psychol Sci Soc Sci* 1995;50:S182–9. [PubMed: 7767702]
40. Coyle CP, Roberge J. The psychometric properties of the Center for Epidemiological Studies-Depression Scale (CES-D) when used with physical disabilities. *Psychol Health* 1992;7:69–81.
41. LISREL [computer program] Version 8.8. Chicago: Scientific Software International; 2008.
42. Little TD, Cunningham WA, Shahar G, Widaman KF. To parcel or not to parcel: exploring the question, weighing the merits. *Structural Equation Modeling* 2002;9:151–73.
43. Steiger JH. Structural model evaluation and modification: an interval estimation approach. *Multivariate Behavioral Res* 1990;25:173–80.
44. Browne, MW.; Cudeck, R. Alternative ways of assessing model fit. In: Bollen, KA.; Long, JS., editors. *Testing Structural Equation Models*. Newbury Park, CA: Sage; 1993. p. 136-62.

45. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling* 1999;6:1–55.
46. Enders CK. The performance of the full information maximum likelihood estimator in multiple regression models with missing data. *Educ Psychol Measurement* 2001;61:713–40.
47. Carpenter JS, Rand KL. Modeling the hot flash experience in breast cancer survivors. *Menopause* 2008;15:469–75. [PubMed: 18467951]
48. Thurston RC, Blumenthal JA, Babyak MA, Sherwood A. Association between hot flashes, sleep complaints, and psychological functioning among healthy menopausal women. *Int J Behav Med* 2006;13:163–72. [PubMed: 16712434]
49. Beck SL, Dudley WN, Barsevick A. Pain, sleep disturbance, and fatigue in patients with cancer: using a mediation model to test a symptom cluster. *Oncol Nursing Forum* 2005;32:542.
50. Affleck G, Tennen H, Zautra A, Urrows S, Abeles M, Karoly P. Women's pursuit of personal goals in daily life with fibromyalgia: a value-expectancy analysis. *J Consult Clin Psychol* 2001;69:587–96. [PubMed: 11550725]
51. Brunstein JC. Personal goals and subjective well-being: a longitudinal study. *J Personality Soc Psychol* 1993;65:1061–70.
52. Kronenberg F. Hot flashes: epidemiology and physiology. *Ann N Y Acad Sci* 1990;592:52. [PubMed: 2197954]
53. Young T, Rabago D, Zgierska A, Austin D, Finn L. Objective and subjective sleep quality in premenopausal, perimenopausal, and postmenopausal women in the Wisconsin sleep cohort study. *Sleep* 2003;26:673–7. [PubMed: 14572119]
54. Polo-Kantola P, Erkkola R, Irjala K, Helenius H, Pullinen S, Polo O. Climacteric symptoms and sleep quality. *Obstet Gynecol* 1999;94:219–24. [PubMed: 10432131]
55. Wilson S, Argyropoulos S. Antidepressants and sleep: a qualitative review of the literature. *Drugs* 2005;65:927–47. [PubMed: 15892588]
56. Carpenter JS, Azzouz F, Monahan PO, Storniolo AM, Ridner SH. Is sternal skin conductance monitoring a valid measure of hot flash intensity or distress? *Menopause* 2005;12:512–9. [PubMed: 16145304]
57. Caltabiano ML, Holzheimer M. Dispositional factors, coping and adaptation during menopause. *Climacteric* 1999;2:21–8. [PubMed: 11910676]
58. Carpenter JS, Gilchrist JM, Chen K, Gautam S, Freedman RR. Hot flashes, core body temperature, and metabolic parameters in breast cancer survivors. *Menopause* 2004;11:375–81. [PubMed: 15243274]
59. Freedman RR. Biochemical, metabolic, and vascular mechanisms in menopausal hot flashes. *Fertil Steril* 1998;70:332–7. [PubMed: 9696230]

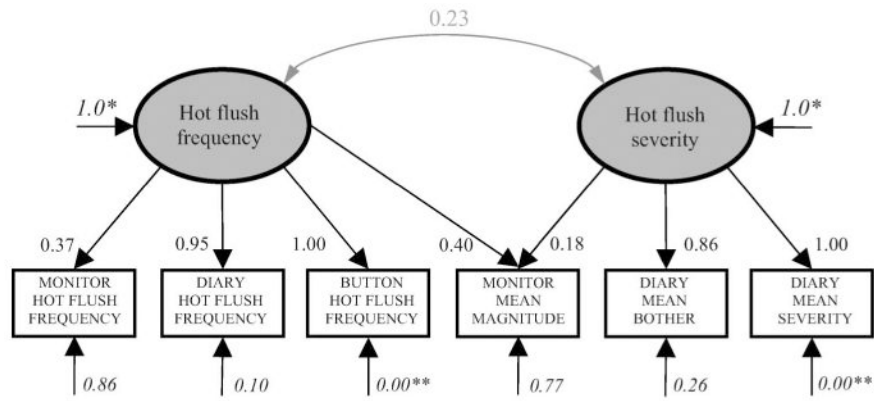


Figure 1. Hot flush model. Prior to standardization, coefficient fixed to 1.0 (*) or 0 (**). Ovals represent latent variables and rectangles the measured variables. Bold black arrows indicate factor loadings (causal relationships). Grey curved arrows indicate correlations (non-causal relationships). Italicized coefficients are error terms, representing the proportion of variation in the variable not accounted for by the model. All coefficients significant at $p < 0.05$, unless otherwise noted

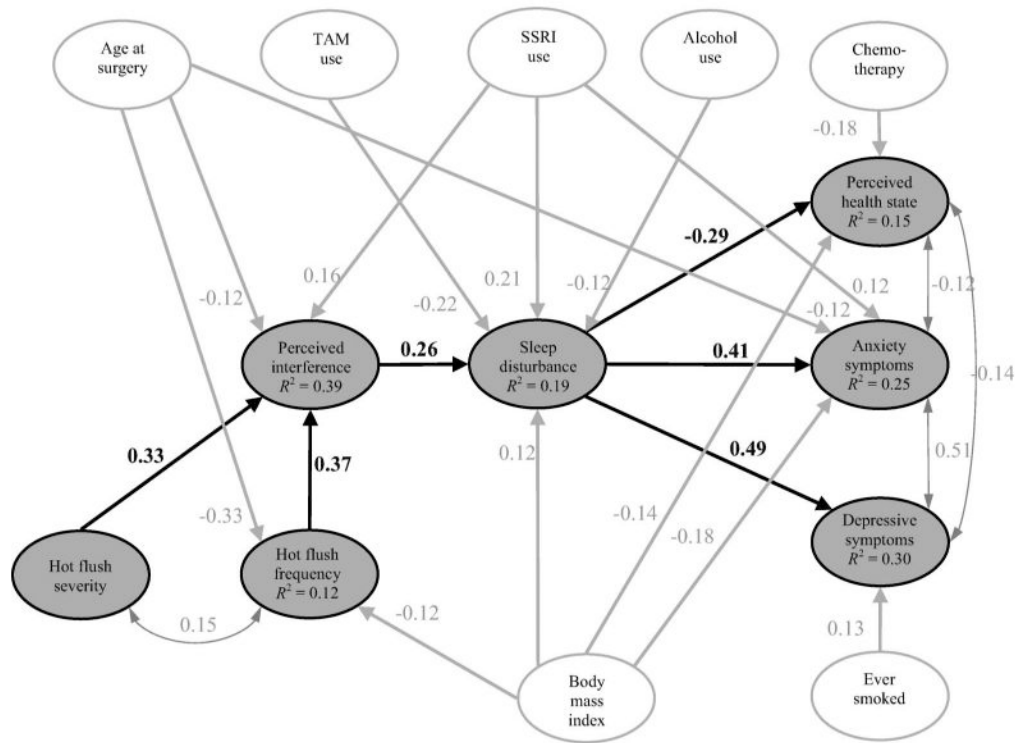


Figure 2. Hot flush symptoms and quality-of-life model. Gray ovals represent hot flush symptoms and related outcomes; white ovals represent patient/treatment covariates. Single-headed arrows indicate regression paths (hypothesized causal relationships). Double-headed arrows indicate correlations (non-causal relationships). R^2 is the proportion of variation in the variable accounted for by the model. All coefficients significant at $p < 0.05$. TAM, tamoxifen; SSRI, selective serotonin reuptake inhibitor

Table 1
Descriptive statistics for flushers and non-flushers

	<i>Flushers (n = 284)</i>		<i>Non-flushers (n = 96)</i>	
	<i>Mean (SD)</i>	<i>Range</i>	<i>Mean (SD)</i>	<i>Range</i>
Age (years)**	57.64 (7.70)	35–84	63.17 (9.28)	41–85
Body mass index*	29.39 (5.97)	18–56	31.37 (7.12)	18–54
Monitor HF frequency (24 h)	5.26 (6.91)	0–42	0	0
Diary HF frequency (24 h)	3.64 (4.21)	0–23	0	0
Button HF frequency (24 h)	3.55 (4.06)	0–22	0	0
Mean change in conductance	3.12 (1.39)	0–13	0	0
Mean bother	2.45 (1.89)	0–10	0	0
Mean severity	2.99 (1.78)	0–9	0	0
Perceived interference**	1.00 (1.53)	0–10	0.14 [†] (0.52)	0–4
Sleep disturbance	6.28 (3.43)	0–17	5.70 (3.62)	0–17
Perceived health state	3.82 (14.81)	–40–45	5.32 (14.66)	–40–35
Anxiety symptoms	4.30 (3.29)	0–19	3.60 (3.00)	0–15
Depressive symptoms	8.28 (7.93)	0–51	7.29 (7.07)	0–32

* Significant difference between flushers and non-flushers at $p < 0.05$;

** significant difference between flushers and non-flushers at $p < 0.01$;

[†] although non-flushers did not experience hot flushes during the monitoring period of this study, they may have experienced them at other times. Given that the HFRDIS asks respondents to report about the previous week, it is reasonable that some non-flushers reported hot flush-related interference

HF, hot flush; SD, standard deviation; HFRDIS, Hot Flash Related Daily Interference Scale

Table 2
Descriptive statistics and uncontrolled correlations for continuous variables in the models (n = 395)

	Descriptive statistics			Pearson's correlations												
	Mean (SD)	Range		1	2	3	4	5	6	7	8	9	10	11	12	
1. Age (years)	59.29 (8.70)	35–89	–	–												
2. Body mass index	29.87 (6.29)	18–56	0.02	–												
3. Monitor HF frequency (24 h)	3.98 (6.42)	0–42	–0.17**	–0.20**	–											
4. Diary HF frequency (24 h)	2.72 (3.97)	0–23	–0.33**	–0.14*	0.46**	–										
5. Button HF frequency (24 h)	2.68 (3.84)	0–22	–0.32**	–0.15**	0.48**	0.94**	–									
6. Mean change in conductance	3.14 (1.43)	0–13	–0.09	–0.06	0.00	0.25**	0.29**	–								
7. Mean bother	2.44 (1.89)	0–10	0.01	0.07	0.00	0.11	0.06	–0.02	–							
8. Mean severity	2.99 (1.79)	0–10	–0.05	0.04	0.04	0.16*	0.13	0.00	0.85**	–						
9. Perceived interference	0.76 (1.38)	0–10	–0.29**	–0.04	0.14*	0.45**	0.40**	0.24**	0.44**	0.44**	–					
10. Sleep disturbance	6.12 (3.47)	0–17	–0.09	0.14*	–0.02	0.11*	0.09	0.05	0.04	0.09	0.27**	–				
11. Perceived health state	4.14 (14.90)	–40–45	0.21**	–0.17**	0.04	–0.10	–0.08	–0.11	0.03	0.01	–0.14**	–0.28**	–			
12. Anxiety symptoms	4.12 (3.21)	0–21	–0.13*	–0.11*	–0.03	0.04	–0.01	0.03	0.14*	0.14*	0.21**	0.28**	–0.23**	–		
13. Depressive symptoms	8.07 (7.72)	0–51	–0.06	0.03	–0.11*	0.01	–0.03	–0.03	–0.03	0.16*	0.14*	0.23**	0.49**	–0.28**	0.65**	

* p < 0.05;

** p < 0.01

HF, hot flush; SD, standard deviation

Table 3
Correlations among covariates in structural model (*n* = 395)

	1	2	3	4	5	6	7
1. Minority status (yes, no)	–						
2. Body mass index	0.08	–					
3. Age at surgery	0.01	0.05	–				
4. Chemotherapy (yes, no)	–0.09	–0.08	–0.40**	–			
5. Tamoxifen (yes, no)	–0.04	–0.02	–0.44**	0.25**	–		
6. Alcohol use (yes, no)	–0.05	–0.21**	–0.12*	–0.01	0.11*	–	
7. Ever smoked (yes, no)	–0.07	–0.07	–0.02	–0.01	0.01	0.11*	–
8. SSRI use (yes, no)	–0.07	0.08	–0.16**	0.06	0.13**	–0.04	0.17**

* *p* < 0.05;

** *p* < 0.01

SSRI, selective serotonin reuptake inhibitor