Published in final edited form as:

Menopause. 2009 ; 16(4): 653–659. doi:10.1097/gme.0b013e3181a5d0d6.

Comparison of Subjective and Objective Hot Flash Measures Over Time Among Breast Cancer Survivors Initiating Aromatase Inhibitor Therapy

Julie L. Otte, PhD, RN¹, David Flockhart, MD, PhD², Anna Maria Storniolo, MD², Bryan Schneider, MD², Faouzi Azzouz, MS², Suzanne Lemler, BSN, RN², Stacie Jeter, BA, CCRP⁴, Janet S. Carpenter, PhD, RN¹, Daniel Hayes, MD³, Vered Stearns, MD⁴, N. Lynn Henry, MD, PhD³, Anne Nguyen, BS², Jill Hayden, RN³, and Laurie Wright, BS, RN⁴

¹School of Nursing, Indiana University

²School of Medicine, Indiana University

³Breast Oncology Program, University of Michigan Comprehensive Cancer Center

⁴Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins School of Medicine

Abstract

Objective—Hot flashes are valuable indicators of physiologic condition and drug effect; however, subjective and objective measures do not always agree. No study has examined both subjective and objective hot flashes in women prescribed aromatase inhibitors. The study (1) compared subjective and objective hot flash measures, (2) examined changes in subjective and objective hot flashes over time, and (3) evaluated predictors of change in hot flashes in aromatase inhibitor-treated women.

Design—Subjects (n=135) were enrolled in a randomized clinical trial comparing exemestane and letrozole for the treatment of breast cancer. Hot flashes were assessed prior to starting drug and 1, 3, and 6 months later. Subjects wore a sternal skin conductance monitor for ≥ 24 hours at each time point. With each perceived hot flash, women pressed an event button and rated intensity and bother in a paper diary.

Results—Subjects were a mean age of 60 years and mainly Caucasian (92%). Across time points, monitor hot flashes were (1) significantly more frequent than diary and/or event button flashes ($p < .05$) and (2) moderately correlated with subjective measures ($.35 < r < .56$). Monitor hot flashes did not significantly change over time with aromatase inhibitor therapy, whereas both diary and event button frequency significantly varied but in dissimilar patterns (51% non-linear). No consistent predictors of hot flashes across measures or time points were identified.

Conclusions—Findings indicated dissimilarities between subjective and objective measures of hot flashes. Despite statistical significance, there was little clinically meaningful change in hot flashes after initiating aromatase inhibitor therapy.

Address for correspondence and reprints: Julie L. Otte, PhD, RN, Indiana University, 1111 Middle Drive NU236, Indianapolis, IN 46220, (ph) 317-278-6094, (fax) 317-274-4774, (email) jlelam@iupui.edu.

Objective and subjective hot flash measures were analyzed after initiation of two aromatase inhibitors for the treatment of breast cancer. Findings indicated dissimilarities among measures, hot flash frequency and hot flash intensity and bother did not significantly change over time, diary and event button flashes significantly changed but in dissimilar patterns, and few consistent predictors of change in hot flashes were identified.

Keywords

hot flashes; menopause; aromatase inhibitor; breast cancer

INTRODUCTION

Hot flashes are the cardinal symptom of menopause and are important indicators of physiologic condition and of drug effect. Hot flashes can be measured subjectively via self-report and objectively via sternal skin conductance monitoring. Data from American and international populations show that subjective and objective measures do not always agree.^{1–5} During the daytime and in a laboratory setting, concordance between subjective and objective monitoring has been reported as near 100%.⁶ However, concordance between objective and subjective hot flashes is low in nighttime laboratory studies⁵ and in daytime or nighttime ambulatory studies.¹ Women may report hot flashes that do not register objectively or may not report hot flashes that do register objectively.

While it is possible that women may over-report hot flashes in order to qualify for a treatment study, our previous data from breast cancer survivors indicated that this was not true. Breast cancer survivors seeking treatment for hot flashes tended to under-report hot flashes meeting objective criteria.^{1, 7} Compared to monitor-detected hot flashes, subjective hot flash reports such as paper diaries and electronic event buttons significantly underestimated hot flashes in breast cancer survivors seeking treatment for the symptom. Using at-home objective monitoring, breast cancer survivors under-reported 50%–64% of daytime objective hot flashes and 58%–78% of nighttime objective hot flashes.^{1, 7} In another study of breast cancer survivors, women reported difficulty completing hot flash diaries because of sleep, being too busy, or other reasons.^{1, 7} No published studies using both objective and subjective methods have examined changes in hot flashes related to the profound decrease in serum estrogen concentration after initiation of aromatase inhibitors. Existing studies include only subjective hot flash data and describe the percentage of women reporting hot flashes,⁸ levels of hot flash bother,⁹ or hot flash severity/hot flash diary scores after treatment,^{10–13} with little description of change over time. Therefore, we used longitudinal data from a randomized controlled trial comparing the effects of two aromatase inhibitors for the treatment of breast cancer to compare hot flash measures by (1) determining relationships among subjective and objective hot flash measures, (2) examining changes in subjective and objective hot flash measures over time, and (3) evaluating predictors of change in hot flashes.

MATERIALS AND METHODS

Sample and setting

Patients were recruited for a multi-center randomized clinical trial designed to study the effects of two aromatase inhibitors, exemestane or letrozole, on surrogate markers of response to antiestrogen therapy, and to correlate these effects with germline genetic variants in candidate genes involved in estrogen metabolism and signaling. Patients were recruited from clinics at the Indiana University Melvin and Bren Simon Cancer Center including Wishard Memorial Hospital; the University of Michigan Comprehensive Cancer Center; and the Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins. All procedures were approved by three levels of review: each site's local institutional review board, cancer center scientific review committees, and General Clinical Research Center advisory committees. Women were eligible if they: (a) were post-menopausal, (b) had histologically proven ductal carcinoma in situ (DCIS/stage 0) or stage I–III invasive carcinoma of the breast that was estrogen receptor and/or progesterone receptor positive, (c) had completed any recommended adjuvant chemotherapy (d) were considering treatment with an aromatase inhibitor as initial adjuvant hormonal

treatment or following adjuvant tamoxifen, and (e) had an ECOG performance status of 0–2. Eligible and interested subjects provided a written informed consent and authorization for release of health information.

Study procedures

Study visits, which included hot flash assessment, were completed at baseline (before the start of aromatase inhibitor therapy) and 1, 3, 6, and 12 months after starting aromatase inhibitor therapy. Because this trial is still accruing subjects, not all subjects enrolled had data at all available time points. Subjects were instructed to wear the monitor for a total of 36 hours. However, not all files had a total of 36 hours of data. To maximize the use of the available data, subjects were included in the analyses if they had at least 24 hours of completed diary and monitor hot flash data from baseline, 1, 3 and 6 months (n=135).

Trained study personnel instructed subjects on the use and care of the Biolog® hot flash monitor and connected each subject to the monitor at the time of her clinic visit. Once the monitoring session was completed, participants followed written instructions for turning off and returning the monitor via pre-paid postal mail to the study site. Each site downloaded the hot flash data from the monitors and cleaned and prepared the monitors for the subsequent patient. Downloaded data were sent via email to a central lab at Indiana University. All hot flash monitor data were analyzed at Indiana University using customized software and established procedures with trained analysts, as described previously.^{7, 14–16}

Measures

Physiological hot flash frequency was assessed using sternal skin conductance monitoring (UFI, Model 3991 SCL, Morro Bay, CA).^{7, 15–17} Skin conductance was monitored using single use sensors (UFI, Morro Bay, CA) and a 0.5 constant voltage circuit built into the front end of a single channel Biolog® ambulatory recorder (UFI, Model 7-day 3991 SCL, Morro Bay, CA). The electrodes were 1.5 cm in diameter and filled with 0.05 M potassium chloride Velvachol/glycol gel. Electrodes were attached 1.5 inches below the collarbones and 2 inches on either side of the sternal mid-line which measure sweat gland activity, an indicator of sympathetic activation.⁶ Using established scoring procedures,¹ it is possible to differentiate a hot flash from exercise and other sweating episodes. Because the electrodes are placed on the sternum, it is not a good indicator of general sympathetic activity due to stress, fear, or anxiety as is typically measured using palmar skin conductance. The monitor is a solid-state device containing a microprocessor and four megabytes of memory powered by a 9 volt battery. It measures 1.3 × 2.8 × 5 inches and weighs 8 ounces. This machine is programmed to sample 12-bit skin conductance data once per second. It is placed in a bag and worn around the waist or across the shoulders. An event button on the monitor can be pressed by participants to signal the subjective perception of a hot flash. This event button timestamps the sternal skin conductance data and is used to help interpret the data during data analysis. At the end of the monitoring session, the monitor is connected to a personal computer to download data and determine hot flash frequency. An objective hot flash was defined as each discrete $\geq 2 \mu\text{mho}$ rise in sternal skin conductance occurring within a 30 second period. The time between subsequent hot flashes measured by the monitor is 15 minutes. Once a hot flash is counted, another hot flash is not counted until 15 minutes had elapsed. This has been the standard in prior published studies.¹ The total number of objective hot flashes was summed for each monitoring session. Scoring was completed by trained raters with inter-rater reliability exceeding 95%.

Self-reported hot flash frequency was assessed using written diaries and the electronic event button during each monitoring session. Self-reported intensity and bother were also assessed with diaries. When a subject experienced a hot flash, she pushed two red buttons on the hot

flash monitor, wrote down the time of the hot flash, and rated intensity and bother in a paper diary (0=not at all and 10=extremely severe or 10=extremely bothersome).^{7, 18–21} Intensity and bother of subjective hot flash frequency were individually summed for a total score at each time point.

Statistical analysis

Data analyses were performed using the Statistical Package for the Social Sciences (SPSS) (v. 15.0, SPSS Inc., Chicago, IL) and Statistical Analysis Software (SAS) (SAS 8.2 System for Windows; SAS Institute, Cary, NC). Demographic variables and hot flash outcomes were evaluated using descriptive and frequency statistics. Data were evaluated for completeness for diary and monitor sessions for all four time-points (baseline, 1 month, 3 month, 6 month). Subjects were included in the analyses if they provided both objective and self-reported hot flash data at the four time points. Hot flash monitor files were considered complete if recording sessions contained at least 24-hours of data. Missing data were found to be randomly distributed by subject and time point. Objective and self-reported hot flash frequency outcomes were standardized to 24-hour time periods to reduce missingness.

To determine the relationships among subjective and objective hot flash measures, we compared hot flash measures in four separate analyses. We hypothesized that there would be no significant differences among the hot flash measures. First, we examined concordance between monitor and event button frequency at each time point through frequency statistics. We evaluated the total number of hot flashes recorded by the monitor alone, by the monitor and event button, and by the event button only. Second, we compared hot flash frequency as assessed using the monitor, diary and event button at each time point using paired *t* tests. Third, using previously published methods,¹ the accuracy of each subjective measure (diary, event marker) was calculated using 4 indices; sensitivity, specificity, positive predictive value, and negative predictive value for the diary and event button compared to the reference standard of monitor hot flashes.^{6, 7} These 4 indices were calculated using all available hot flashes for each week during waking hours. The binary data (hot flash/no hot flash) were clustered because each patient was measured continuously for at least 24 hours, with hot flashes occurring at multiple times during each weekly monitoring session. The confidence intervals for specificity and sensitivity were computed as described by Zhou et al.²² by using the ratio estimator for the variance.^{23, 24} To identify true negatives (no objective or subjective hot flash), each weekly monitoring period was divided into 96 consecutive 15-minute periods of data. The total number of true negatives for each woman at baseline was calculated as the total number of the 96 data segments that contained neither an objective flash nor a subjective flash (no diary, no event marker). For this analysis, we used only a subset of women who had baseline data that was scored using a more detailed method as previously described.¹ Because this scoring method is more time consuming, it was not used for all subjects' data. Fourth, we examined Pearson correlations among hot flash measures (monitor, diary, event mark) at each time point.

To determine changes in subjective and objective hot flash measures over time, hot flash measures were analyzed as follows. We hypothesized that hot flashes would significantly increase over time on all measures. For the sample as a whole, repeated-measure analyses of covariance were used to determine significant differences over time in each hot flash measure controlling for baseline hot flash frequency. In addition, for each individual woman, the pattern of change in monitor hot flashes was evaluated over time and women were classified into one of four groups: (1) no change in hot flashes, (2) increasing hot flashes over time, (3) decreasing hot flashes over time, and (4) non-linear change (e.g., increase and decrease) over time. Women were similarly grouped according to pattern of change in diary-reported hot flash frequency.

To determine predictors of change, hot flashes were evaluated in two steps. It was hypothesized that body mass index (BMI), use of selective serotonin reuptake inhibitors, prior number of

years on tamoxifen, or prior number of years of hormone replacement therapy might be predictors of changes in hot flashes over-time. First, linear regression was performed using covariates of BMI, use of selective serotonin reuptake inhibitors, prior number of years on tamoxifen, and prior number of years of hormone replacement therapy. Second, we compared characteristics of the four patterns of change in hot flash groups (described in the previous paragraph) using chi-square tests and analysis of variance. Groups were compared on BMI, number of years on tamoxifen, number of years on hormone therapy, and use of hot flash treatments.

Differences in hot flash parameters between the two aromatase inhibitors were not measured since the sample was not powered to evaluate drug effect. This will be evaluated once the full sample is obtained and the parent study is completed.

RESULTS

The 135 subjects were a mean age of 59.76 years (SD=8.32). They were all non-Hispanic/Latina and mainly Caucasian (92%). Randomization to aromatase inhibitor therapy was 57% letrozole and 43% exemestane. For the 48% of women who had previously taken hormone replacement therapy, the mean time from discontinuation was 5.7 years (SD=5.6). For the 34% of women who had previously taken tamoxifen, the mean time since stopping was 2 years (SD=1.4) and 39% of them had stopped within one month or less of starting the aromatase inhibitor (13% of the total sample). Prior breast cancer treatments included surgery (100%), chemotherapy (49%), and radiation therapy (78%).

Comparisons among subjective and objective hot flash measures

The agreement between monitor and event button at each time point is shown in Figure 1. Overall, this figure shows that more hot flashes were recorded on the monitor than with subjective methods. At all time points, approximately 50% of hot flash events were recorded by the monitor only, another 20% were recorded by both monitor and event button, and 30% were recorded by the event button only (no monitor hot flash). Thus, at all time points, approximately 70% of detected hot flash events were recorded by the hot flash monitor whereas only 50% were subjectively reported.

Mean hot flash frequency with standard deviation bars for monitor, diary, and event button measures at each time point are shown in Figure 2. At baseline, there were no significant differences among measures. At 1 month, monitor hot flash frequency was significantly higher than diary ($p=0.03$) and event button frequency ($p=0.02$). At 3 months, monitor hot flash frequency was significantly higher than event button frequency ($p=0.04$). At six months, both monitor and diary frequency were significantly higher than event button ($p=0.03$ and 0.04 , respectively). Thus, at three of the four time points, significantly more hot flashes were recorded on the monitor compared to one or both subjective methods.

Results of the detailed scoring of the 46 subjects' baseline data are shown in Table 1 and again document that the monitor captured more hot flashes compared to subjective methods. Sensitivity was low for diary and event button hot flash frequency. The upper limit of 95% confidence intervals indicated that sensitivity rarely exceeded 29% for diary and event button-reported hot flash frequency. When a woman had a monitor hot flash, she was likely to record it in her diary 22% of the time and with the event button 25% of the time. Specificity was high for both diary and event button hot flash frequency. Using the lower limit of the 95% confidence interval, specificity rarely fell below 97.7%. The positive predictive value was moderate (41–44%) and the negative predictive value was high (95%). Subsequent t tests indicated that hot flashes that were reported with the event button but not confirmed on the monitor (false positives) were significantly more intense (mean=3.61 vs. 2.98) and more bothersome

(mean=2.90 vs. 1.99) than hot flashes that were reported and confirmed on the monitor (true positives) ($p=0.03$).

Pearson correlations among all hot flash measures at each time point are shown in Table 2 and Table 3. Consistent with the previous analyses, monitor frequency was moderately correlated with subjective diary and event button frequency ($.35 < r < .56$). The strength of the relationship appeared to decrease over time. Subjective diary and event button frequency were highly correlated with each other ($.90 < r < .97$) and with hot flash intensity and bother ($.69 < r < .92$). Monitor frequency was weakly to moderately correlated with subjective intensity and bother ($.28 < r < .42$). In addition, subjective intensity and bother were highly correlated with each other ($.94 < r < .97$).

Changes in subjective and objective hot flash measures over time

Changes over time for each hot flash frequency measure when controlling for baseline are shown in Table 4. Monitor hot flashes did not significantly change over time. Both diary and event button frequencies changed over time though in a dissimilar pattern. Diary hot flashes decreased significantly from 3 to 6 months by about 1 hot flash. In contrast, event button frequency significantly increased between 1 to 3 months (by approximately 1 hot flash) and then significantly decreased by nearly 2 hot flashes at 6 months. Hot flash intensity and bother over time are shown in Figure 3. There were no significant differences in intensity or bother over time.

When data were graphed for each individual woman (not shown), there was no consistent pattern of change in objective hot flashes over time. Using monitor hot flash frequency, 27% of subjects had no change in hot flashes over time, 10% had increasing hot flashes over time, 12% had decreasing hot flashes over time, and 51% had a non-linear change in hot flashes (e.g., increase then decrease). Using diary hot flash frequency, 22% had no change in hot flashes, 7% had increasing hot flashes over time, 13% had decreasing hot flashes over time, and 58% had a nonlinear change.

Predictors of change in hot flashes

Using linear regression after controlling for baseline hot flash frequency, the following predictors were found: (1) more monitor hot flashes at 6 months were related to longer duration of prior hormone therapy compared to shorter duration of hormone therapy ($p<0.00$). No significant predictors were identified for diary hot flash frequency. Since subjective diary and event button were highly correlated, predictors of event button frequency were not evaluated. Predictors of greater hot flash intensity were: (1) at one month, higher BMI ($p=0.04$) compared to lower BMI and (2) at 6 months, hot flash treatment use (antidepressant or herbal) compared to no treatment ($p=0.04$). Predictors of hot flash bother were not evaluated because intensity and bother were highly correlated.

At the individual level, women with a non-linear change in monitor hot flashes (51%) had taken tamoxifen for significantly more years compared to women showing other patterns of change in hot flashes over time (e.g., compared to those whose monitor hot flashes increased over time, decreased or stayed stable) ($p<0.00$). In addition, women whose diary hot flashes increased over time had significantly higher BMI compared to other women ($p<0.00$).

DISCUSSION

The main purpose of this study was to evaluate hot flashes over time in relation to subjective and objective measures in women initiating aromatase inhibitors. This study provides the first comprehensive assessment of objective and subjective hot flashes over time after initiation of

aromatase inhibitors for the treatment of breast cancer. Findings indicated dissimilarities between subjective and objective measures. More hot flashes were recorded on the monitor than with the diary or event button. Although monitor hot flash frequency and hot flash intensity and bother did not significantly change over time, diary and event button flashes significantly changed but in dissimilar patterns. Given these differences among measures, there were few consistent predictors of change in hot flashes identified.

The detailed sensitivity and specificity analyses revealed similar data to previous publications. Compared to a previous study,¹ sensitivity was slightly worse for both diary and event button, whereas specificity, positive predictive value, and negative predictive value were comparable. The agreement between objective and subjective frequency was worse compared to other studies that found 44–51% true positives and 56% false negatives^{3, 4} and compared to another study reporting lower true positives (19%) and more false negative scores (81%) in a group of Bangladeshi women.²⁵ In summary, these and other data suggest that women tend to report fewer hot flashes than are recorded on the hot flash monitor. However, whether these monitor-only-detected hot flashes are clinically relevant is unclear. Different information may be being picked up by the subjective and objective methods but the clinical difference may not be significant. It is puzzling that subjective hot flashes reported by the event button and not confirmed by the monitor (false positive) were reported as more intense and bothersome than those reported by the event button and confirmed on the monitor (true positives). This may indicate that the monitor and self-reports are capturing different aspects of the hot flash experience and provides evidence that the perception of a hot flash is just as important to investigate in the absence of an objective marker of that hot flash.

Regardless of the measurement method used, our findings indicate that the clinical impact of aromatase inhibitors on hot flashes is minimal despite the significant drop in serum estrogen concentration that presumably occurs. Although some statistically significant changes in hot flashes were seen when evaluating group data, these amounted to only 1–2 hot flashes per 24-hour monitoring period. Whether these differences represent measurement error with the diary and event button, normal variation in hot flashes, or true change in hot flashes as a result of aromatase inhibitors remains unclear. In addition, when evaluating data from individual subjects, the pharmacological impact of aromatase inhibitors varied. Taking an aromatase inhibitor did not produce a progressive increase in hot flash frequency. Most subjects demonstrated a non-linear pattern over time of increasing/decreasing or decreasing/increasing hot flashes. The authors' hypothesize three explanations for this finding: 1) functional estrogen is already very low when post-menopausal women start an aromatase inhibitor and the estrogen sensitive concentration range is above that brought about by an aromatase inhibitor; 2) the size of the change in estrogen levels is a main factor in regulation of central thermoregulatory apparatus and that the real concentration change seen with menopause is greater; or 3) women who have already been through menopause are less sensitive to further estrogen depletion, through some compensatory biologic mechanism, involving pathways that remain unclear but could be local within the central nervous system.

Given the differences among hot flash measures, it is not surprising that predictors of hot flashes were not consistent across measures. One predictor, higher BMI, has previously been shown to significantly predict increased subjective hot flash frequency in post-menopausal women.²⁶ A randomized, controlled study of raloxifene indicated that the highest number of subjective hot flashes at baseline was seen in women who were fewer years postmenopausal, who had had a surgical menopause, and who had received previous estrogen or estrogen/progestin therapy.²⁷ However, these variables did not consistently predict hot flashes over time after initiating raloxifene.

Limitations of this study included a lack of racial diversity and heterogeneity on other factors such as prior use of tamoxifen, prior use of hormone therapy, time post-menopausal, and varied prior cancer treatments. Prior tamoxifen use within 4 weeks of starting an aromatase inhibitor (13% of the sample) could explain why some women did not experience a significant increase in hot flashes from baseline. In addition, concurrent prescription and over-the-counter medications (other than hot flash treatments) were not evaluated to determine their role in hot flashes. The burden of asking women to write in a paper diary and push an event button could have reduced concordance for either one of these measures with the objective hot flash measure. Subjects were not asked to document whether hot flashes occurred during sleep or wake. A large number of nighttime (sleeping) hot flashes could have contributed to the low sensitivity findings. In addition, analyzing a 24-hour time frame for objective hot flash monitoring could be considered limited as hot flash frequency can be variable from day to day. Lastly, since we only evaluated post-menopausal women, these results cannot be generalized to women experiencing hot flashes in the perimenopause as a result of processes other than a medication.

Future studies should consider using only one subjective measure of hot flash frequency, documentation of sleep and wake times, and either a measure of intensity or bother. Although this analysis did not examine genetic polymorphisms in drug-metabolizing enzymes as a determinant of hot flashes, future analyses are planned and will be performed after all participants have been enrolled and assessed over time and laboratory data are processed.

Acknowledgments

Sources of Support. Supported in part by a Pharmacogenetics Research Network Grant # U-01 GM61373 (DAF) that supports the Consortium on Breast Cancer Pharmacogenomics (COBRA), a Clinical Pharmacology training grants 5T32-GM08425 (NN, MR) 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™ (DFH). The description of the study design can be found on www.ClinicalTrials.gov (NCT00228956).

REFERENCES

1. Carpenter JS, Monahan PO, Azzouz F. Accuracy of subjective hot flush reports compared with continuous sternal skin conductance monitoring. *Obstet Gynecol* 2004;104:1322–1326. [PubMed: 15572497]
2. Sievert LL. Variation in sweating patterns: Implications for studies of hot flashes through skin conductance. *Menopause* 2007;14:742–751. [PubMed: 17519802]
3. 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;9:367–376. [PubMed: 12218726]
4. Thurston RC, Blumenthal JA, Babyak MA, Sherwood A. Emotional antecedents of hot flashes during daily life. *Psychosom Med* 2005;67:137–146. [PubMed: 15673636]
5. 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;27:513–522. [PubMed: 15165649]
6. Freedman RR. Laboratory and ambulatory monitoring of menopausal hot flashes. *Psychophysiology* 1989;26:573–579. [PubMed: 2616704]
7. Carpenter JS, Andrykowski MA, Freedman RR, Munn R. Feasibility and psychometrics of an ambulatory hot flash monitoring device. *Menopause* 1999;6:209–215. [PubMed: 10486790]
8. Goss PE, Ingle JN, Martino S, et al. A randomized trial of letrozole in postmenopausal women after five years of tamoxifen therapy for early-stage breast cancer. *N Engl J Med* 2003;349:1793–1802. [PubMed: 14551341]
9. Whelan TJ, Goss PE, Ingle JN, et al. Assessment of quality of life in MA.17: A randomized, placebo-controlled trial of letrozole after 5 years of tamoxifen in postmenopausal women. *J Clin Oncol* 2005;23:6931–6940. [PubMed: 16157934]

10. Fallowfield L, Cella D, Cuzick J, Francis S, Locker G, Howell A. Quality of life of postmenopausal women in the Arimidex, Tamoxifen, Alone or in Combination (ATAC) adjuvant breast cancer trial. *J Clin Oncol* 2004;22:4261–4271. [PubMed: 15514369]
11. Fallowfield LJ, Bliss JM, Porter LS, et al. Quality of life in the intergroup exemestane study: A randomized trial of exemestane versus continued tamoxifen after 2 to 3 years of tamoxifen in postmenopausal women with primary breast cancer. *J Clin Oncol* 2006;24:910–917. [PubMed: 16484701]
12. Thurlimann B, Keshaviah A, Coates AS, et al. A comparison of letrozole and tamoxifen in postmenopausal women with early breast cancer. *N Engl J Med* 2005;353:2747–2757. [PubMed: 16382061]
13. Jones SE, Cantrell J, Vukelja S, et al. Comparison of menopausal symptoms during the first year of adjuvant therapy with either exemestane or tamoxifen in early breast cancer: Report of a Tamoxifen Exemestane Adjuvant Multicenter trial substudy. *J Clin Oncol* 2007;25:4765–4771. [PubMed: 17947724]
14. Carpenter JS, Storniolo AM, Johns S, et al. Randomized, double-blind, placebocontrolled crossover trials of venlafaxine for hot flashes after breast cancer. *Oncologist* 2007;12:124–135. [PubMed: 17227907]
15. Carpenter JS, Gautam S, Freedman RR, Andrykowski M. Circadian rhythm of objectively recorded hot flashes in postmenopausal breast cancer survivors. *Menopause* 2001;8:181–188. [PubMed: 11355040]
16. Carpenter JS, Wells N, Lambert B, et al. A pilot study of magnetic therapy for hot flashes after breast cancer. *Cancer Nurs* 2002;25:104–109. [PubMed: 11984098]
17. Carpenter JS, Elam J, Ridner S, Carney P, Cherry G, Cucullu H. Sleep, fatigue, and depressive symptoms in breast cancer survivors and matched healthy women experiencing hot flashes. *Oncol Nurs Forum* 2004;31:591–598. [PubMed: 15146224]
18. Carpenter JS. Physiological monitor for assessing hot flashes. *Clin Nurse Spec* 2005;19:8–10. [PubMed: 15684886]
19. Carpenter JS. Objective and subjective hot flash reports agree during laboratory monitoring in breast cancer survivors. unpublished.
20. 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–E25. [PubMed: 11979290]
21. Loprinzi CL, Flynn PJ, Carpenter LA, et al. Pilot evaluation of citalopram for the treatment of hot flashes in women with inadequate benefit from venlafaxine. *J Palliat Med* 2005;8:924–930. [PubMed: 16238505]
22. Zhou, X-H.; Obuchowski, NA.; McClish, DK. *Statistical Methods in Diagnostic Medicine*. New York: Wiley; 2002.
23. Cochran, W. *Sampling Techniques*. Third ed.. New York: Wiley; 1977.
24. Rao NK, Scott AJ. A simple method for the analysis of clustered binary data. *Biometrics* 1992;48:577–585. [PubMed: 1637980]
25. Sievert LL, Begum K, Sharmeen T, Chowdhury O, Muttukrishna S, Bentley G. Patterns of occurrence and concordance between subjective and objective hot flashes among Muslim and Hindu women in Sylhet, Bangladesh. *Am J Hum Biol*. 2008
26. Huang AJ, Grady D, Jacoby VL, Blackwell TL, Bauer DC, Sawaya GF. Persistent hot flushes in older postmenopausal women. *Arch Intern Med* 2008;168:840–846. [PubMed: 18443259]
27. Aldrich JM, Quail DC, Levy-Frebault J, et al. Predictors of hot flushes in postmenopausal women who receive raloxifene therapy. *Am J Obstet Gynecol* 2004;191:1979–1988. [PubMed: 15592280]

Figure 1. Percentage and Total Number of Hot Flash Events Recorded By Monitor Only, Monitor with Concurrent Event Button, and Event Button Only

This figure shows the percentage of hot flash events that were recorded with the objective hot flash monitor only, the monitor with event button, and the event button only. Data for all four time points are shown. The figure above indicates that approximately 70% of hot flashes were recorded with the monitor (black plus grey), whereas only 50% were recorded with the event button (white plus grey).

Figure 2. Mean Hot Flash Frequency at Each Time Point

This figure shows the mean number of hot flashes with standard deviation bars for all three hot flash frequency measures (monitor, diary, event button) and at each time point. At baseline, there were no significant differences among measures. At one month, monitor frequency was significantly higher than diary and event button frequency. At three months, monitor frequency was significantly higher than event button frequency. At six months, monitor and diary frequency were significantly higher than event button frequency. Note the large standard deviation bars for each measure at each time point.

Figure 3. Hot Flash Intensity and Bother Over Time

This figure shows mean hot flash intensity and bother with standard deviation bars at each time point. Intensity and bother did not significantly change over time. Women rated intensity and bother of each hot flash using separate 0 to 10 point numeric rating scales (not at all to extremely intense or extremely bothersome). Mean intensity and bother during the 24-hour monitoring sessions at each time point were calculated for each woman. The figure shows the mean of the means.

Table 1
Baseline Diary and Event Button Compared to Sternal Skin Conductance

	Sensitivity (%)		Specificity (%)		PPV [*] (%)		NPV [†] (%)	
	Est ± SD [‡]	95% CI [§]	Est ± SD	95% CI	Est ± SD	95% CI	Est ± SD	95% CI
Diary	22 ± 2.1	(18 26)	98 ± 0.2	(97.7 98.4)	41 ± 3.4	(35 48)	95 ± 0.3	(95 96)
Event								
button	25 ± 2.2	(20 29)	98 ± 0.2	(97.7 98.4)	44 ± 3.4	(38 51)	95 ± 0.3	(95 96)

^{*} PPV = positive predictive value.
[†] NPV = negative predictive value.
[‡] Est ± SD: point estimate ± standard deviation.
[§] CI = confidence interval.

Table 2
Pearson Correlations Among All Hot Flash Measures at Baseline^a and 1 Month^b After Starting Aromatase Inhibitor

	1	2	3	4	5
1. Monitor frequency	---	.54	.56	.42	.36
2. Diary frequency	.37	---	.94	.87	.76
3. Event button frequency	.68	.93	---	.81	.69
4. Total intensity	.33	.89	.80	---	.95
5. Total bother	.27	.83	.75	.95	---

^aCorrelations above diagonal (baseline).

^bCorrelations below diagonal (1 month).

All correlations significant at p<.001.

Table 3
Pearson Correlations Among All Hot Flash Measures at 3 Months^a and 6 Months^b After Starting Aromatase Inhibitor

	1	2	3	4	5
1. Monitor frequency	---	.36	.35	.30	.29
2. Diary frequency	.37	---	.97	.89	.85
3. Event button frequency	.41	.90	---	.85	.79
4. Total intensity	.32	.92	.80	---	.98
5. Total bother	.37	.87	.72	.94	---

^aCorrelations above diagonal (3 months).

^bCorrelations below diagonal (6 months).

All correlations significant at p<.001.

Table 4

Comparison of Hot Flash Frequency Over Time Controlling For Baseline Hot Flash Frequency (n=135).

	1m EM (SE)*	3m EM (SE)	6m EM (SE)
Monitor	4.34(0.43)	4.29(0.50)	3.55(0.35)
Diary	3.19(0.24)	3.31(0.26) ¹	2.87(0.27) ²
Event button	3.11(0.24) ¹	4.29(0.53) ²	2.52(0.22) ³

* Estimated marginal means (EM) and standard errors (SE) are represented.

Differing superscripts within a row denote significant differences, $p < 0.05$.