ABSTRACT

Measuring Hot Flashes: Examination of an Alternate Criterion for Ambulatory Hot Flash Detection in Post-Menopausal Women

William Fisher, Ph.D.

Mentor: Gary R. Elkins, Ph.D.

Hot flashes are a highly prevalent and severe symptom experienced in menopause and in cancer treatment. Objective measurement of hot flashes, essential for conducting research in this area, is achieved using skin conductance monitors. The current gold standard skin conductance level (SCL) criterion for detecting hot flashes is a ≥ 2 μmho rise in a 30-s period, however this standard criterion has shown limited sensitivity in ambulatory trials. A recent study of breast cancer survivors found an alternate criterion (≥ 1.2 μmho rise in a 30-s period) improved performance. The purpose of this dissertation is the comparison of the alternate and standard SCL criterion in population of healthy post-menopausal women, the determination of racial/ethnic group differences, and the determination of any relationship of SCL-derived hot flashes to hot flash related daily interference. This study is a retrospective analysis of data collected from a sample of 140 healthy post-menopausal women in a clinical trial of hypnosis for hot flashes. Skin conductance data sampled over one day using ambulatory SCL-monitors were analyzed using the standard and alternate criteria for hot flash detection. Indices of SCL-criterion

performance*, sensitivity*, *specificity, positive* and *negative predictive values,* were subsequently calculated and compared. Comparison of these indices were made to examine racial/ethnic differences, however, limited sample sizes only allowed for the statistical examination of white versus non-whites. Correlations of the standard and alternate criteria-derived hot flashes with the Hot Flash Related Daily Interference Scale were performed. The results of this study showed that the alternate criterion has a distinct advantage in terms of sensitivity (77.4%, alternate vs. 56.9%, standard; $p < .001$), with only minor deficit to specificity (80.1% alternate vs. 86.1%, standard). Sensitivity of the alternate criterion differed significantly between Whites vs. Non-Whites (83.58%, vs. 73.81%, respectively; $p < .05$). No significant correlations were revealed. The alternate criterion is recommended for use in the ambulatory study of healthy, post-menopausal women. Additional research is needed to: 1) determine specific racial/ethnic group differences in SCL performance, 2) determine relative performance in perimenopause, and 3) ascertain if alternate physiologic measures (e.g. ECG; heart-rate variability) could improve performance.

Measuring Hot Flashes: Examination of an Alternate Criterion for Ambulatory Hot Flash Detection in Post-Menopausal Women

by

William Fisher, B.A., M.A.

A Dissertation

Approved by the Department of Psychology and Neuroscience

_____

Jaime L. Diaz-Granados, Ph.D., Chairperson

Submitted to the Graduate Faculty of
Baylor University in Partial Fulfillment of the
Requirements for the Degree
of

Doctor of Philosophy

Approved by the Dissertation Committee

_____

Gary R. Elkins, Ph.D., Chairperson

_____

Jaime L. Diaz-Granados, Ph.D.

_____

Wade C. Rowatt, Ph.D.

_____

Matthew S. Stanford, Ph.D.

_____

Thomas A. Fergus, Ph.D.

_____

James W. Ellor, Ph.D.

Accepted by the Graduate School

December, 2013

_____

J. Larry Lyon, Ph.D., Dean

Page bearing signatures is kept on file in the Graduate School.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would first like to acknowledge my chair, Dr. Gary R. Elkins, for his mentorship, support, guidance and example. I am deeply indebted to Dr. Elkins for the boundless opportunities that work and study at Baylor University and the Mind-Body Medicine Research Laboratory has provided. I would also like to acknowledge my committee members, Dr. Jaime L. Diaz-Granados, Dr. Wade C. Rowatt, Dr. Matthew S. Stanford, Dr. Thomas A. Fergus, and Dr. James W. Ellor for agreeing to serve on my committee, and the generous contributions of their time, encouragement, and guidance in this process.

I would also like to acknowledge the assistance of my colleagues at the Mind-Body Medicine Research laboratory for their kind assistance with this project, Vicki Patterson, Dr. Cassie Kendrick and Alisa Johnson. Finally, I would like to acknowledge Dr. Janet S. Carpenter of Indiana University School of Nursing for her advice, encouragement and support.

DEDICATION

To my beloved parents, Dr. Larry and Jere Fisher:

Dad, though I will never fill your shoes, it has been a privilege to walk in your footsteps.

Mom, your endless support, encouragement and love is what has made this possible

CHAPTER ONE

Introduction

Hot Flashes

Hot flashes, (aka 'hot flushes', 'vasomotor events', and 'flushing') are a prevalent and severe symptom that can affect post-menopausal women. Hot flashes have been shown to interfere with sleep, mood and quality of life (Freeman & Sherif, 2007; Kronenberg, 1994; Stein, Jacobsen, Hann, Greenberg, & Lyman, 2000). Etiologic study of hot flashes has revealed hot flashes to be complex physiological events, though the exact pathophysiology remains unknown. Research in women suggests that the experience of hot flashes begins with feeling chilled and an inspiratory sigh (Woodward, Greville, & Freedman, 1995). Data taken from samples of women with hot flashes show they are concurrent with subsequent increases in heart rate, metabolic rate and sweating (Carpenter, Andrykowski, Freedman, & Munn, 1999; Carpenter, Gilchrist, Chen, Gautam, & Freedman, 2004; Freedman, 1998; Kronenberg, 1990; Kronenberg, Cote, Linkie, Dyrenfurth, & Downey, 1984). Hot flashes have been described as sudden and transient episodes of heat and sweating with possible co-occurring palpitations and anxiety.(Carpenter, 2005; Kronenberg, 1994).

Hot Flash Prevalence

Hot flashes are highly prevalent; a systematic review of sixty-six papers containing data on hot flash prevalence from around the world found that estimates of prevalence in postmenopausal women range up to 80%, with a median of 41.5%

(Freeman & Sherif, 2007). As indicated in the systematic review, estimates of global prevalence vary; for example, a pan-European study of 4023 postmenopausal women ranging in age from 50 to 64 years, found a hot flash prevalence of 73% (Scoutellas, O'Neill, Lunt, Reeve, & Silman, 1999). The largest evaluation of menopausal symptoms performed anywhere in the world to-date, the Study of Women's Health Across the Nation (SWAN), examined hot flash prevalence in a sample of 16,065 women across multiple ethnic groups in North America finding that the prevalence of hot flashes was 18% among women of Japanese ethnicity, 21% among women of Chinese ethnicity, 31% among White women, 35% among Hispanic women, and 46% among African American women (Avis et al., 2001; Gold et al., 2000). Given these estimates, it can be safely presumed that millions of women are currently experiencing symptoms.

Impact of Hot Flashes

The impact of hot flashes is myriad and potentially quite serious. In a study of 70 post-menopausal breast-cancer survivors, researchers surveyed patients about the presence and impact of hot flashes (Stein et al., 2000). The researchers found that when compared to women without hot flashes, women with hot flashes experienced 66% more fatigue, 63% poorer sleep quality, and 20% poorer physical health. This finding was significant ($p < .05$), even after controlling for relevant medical, demographic, and cancer-treatment variables.

Hot flashes may also be linked to cardiovascular disease. In a study of 110 post-menopausal women, cardiovascular function was examined (Lambrinoudaki et al., 2012). The researchers found that Carotid IMT, a surrogate marker of subclinical atherosclerosis and cardiovascular risk was found to be significantly increased in women with vasomotor

symptoms as compared to asymptomatic women. This result was independent of cardiovascular risk factors and endogenous hormone levels. The Study of Women's Health Across the Nation Heart Study examined 492 women, 45-58 years of age who were free of clinical cardiovascular disease (Thurston, Sutton-Tyrell, Everson-Rose, Hess, & Matthews, 2008). Using a brachial artery ultrasound, electron beam tomography, self-reported hot flashes, and a blood sample to measure estradiol concentrations, the researchers found that hot flashes were associated with significantly lower flow-mediated dilation (β=-1.01, p=.01) and greater coronary artery and aortic calcification in age- and race-adjusted models (OR= 1.48; OR=1.55, respectively). Hot flashes showed a significant association with flow-mediated dilation (β=.097, p<.03), and aortic calcification (OR=1.63), in models adjusted for cardiovascular disease risk factors and estradiol.

Further, hot flashes interrupt sleep. In a sleep study with physiological monitoring of finger temperature and sternal skin conductance, nine postmenopausal women with severe hot flashes and five asymptomatic premenopausal women were observed (Erlik et al., 1981). During cumulative sleep, 47 objectively determined hot flashes were observed with 45 of those associated with a waking episode determined by polygraphic techniques. A similar association was not observed in premenopausal women. The results of a polysomnographic study of 24 breast cancer survivors involving hot flash measurement using ambulatory skin conductance monitors found that hot flashes were significantly correlated with increased theta (r=.28) and delta (r=.32) electroencephalographic activities (J. Savard et al., 2004). Furthermore, they found that hot flashes were significantly associated with percentage of time awake (p =.038), and

the number of changes to lighter stages of sleep (p=.02). Savard and colleagues concluded that hot flashes were associated with less efficient, more disrupted sleep. In a study examining the impact of hot flashes on quality of life among postmenopausal women being treating for breast cancer, 70 women completed questionnaires (Stein et al., 2000). The results of this study revealed that women with hot flashes reported significantly (p<.001) more fatigue, poorer sleep quality and poorer physical health compared to women without hot flashes.

Hot flashes have also been negatively associated with sexual function. In a prospective study of 43 women, data on hot flashes and regularity of intercourse was obtained at two points during the menopausal transition (McCoy, Cutler, & Davidson, 1985). The frequency and regularity of sexual intercourse and the rating of hot flashes were collected for both a 4-week period prior to interview and a 10-week period prior to interview. The data yielded significant negative correlations between hot flash ratings and frequency of intercourse in 4-week per-interview (r =-.42, p <.02), and 10-week pre-interview (r= -.46, p <.01). Further correlations were shown in an assessment of sexual function in mid-aged Ecuadorian women (Chedraui et al., 2012). In this study, 904 healthy women aged 40-50 completed questionnaires including a six-item Female Sexual Function Index (FSFI-6; Isidori et al., 2010) and a general socio-demographic questionnaire containing personal/partner data. Total FSFI-6 scores were positively correlated with coital frequency, however hot flash intensity was found in a multivariate regression analysis to be negatively associated (β=-1.29, 95% CI (-1.99 to -.60), p <.001).

Psychological factors have also been attributed to hot flashes. In a psychological analysis of menopausal hot flashes, 61 symptomatic women were asked to complete

measures to ascertain their experience with hot flashes (Hunter & Liao, 1995). In this study, depressed mood, anxiety and low self-esteem could discriminate between those who reported problematic hot flashes versus those who did not. Psychological stress has also been associated with increased hot flashes. A study of 21 post-menopausal women involved participants undergoing hot flash monitoring during stressful and non-stressful laboratory conditions (L. C. Swartzman, R. Edelberg, & E. Kemmann, 1990). In this study, participants were randomly assigned to a non-stress control or to a stress session, where a variety of stress stimuli (paced arithmetic task, loud noise, stressful film, social stressor, personally relevant stressor, vigilance task), were presented. There were significantly more (43%; $F(1,20) = 9.39$, $p = .007$) objectively identified hot flashes (via sternal skin conductance monitoring) and subjective report of hot flashes (57%; $F(1, 20) = 6.53$, $p < .02$) during the stress session than during the non-stress session.

Given the scope and impact of hot flashes, the scientific study of the etiology and treatment of hot flashes is of great value. Like any scientific endeavor, precise measurement is critical, and the methodology of hot flash measurement has received considerable study involving two primary methods of hot flash measurement: self-report & objective measures.

Self-Report Measure of Hot Flashes

Self-reported hot flashes have been collected using hot flash diaries, logs, and event-marker devices (Sievert, 2013). Daily symptom diaries are commonly used in symptom measurement as they provide a rich source of data, they support dynamic as well as static analysis, they do not rely on retrospective reflection, and a series of events can be documented and examined for causal relationships (Richardson, 1994).

Currently, there is only one validated hot flash diary. The Hot Flash Daily Diary is an instrument that tallies the frequency of hot flashes and the severity of the reported hot flashes (mild, moderate, severe, very severe) daily for a week (Sloan et al., 2001). This instrument provides a total weekly hot flash frequency, can yield an average daily frequency and severity rating, and is designed to provide a hot flash composite score (product of number of hot flashes by their severity rating). Concurrent validity was demonstrated through correlation analyses with toxicity, quality of life (QOL), and preference data compiled over a series of the authors' clinical trials. Further evidence of validity was obtained from questions posed to participants in five cross-over trials finding concordance with treatment outcomes and results from the diary. Reliability of the diary was evidenced in a four-arm (n=50/arm) clinical trial of venlafaxine, which showed that diary results were nearly identical at baseline for each trial arm (Loprinzi et al., 2000). Sloan and colleagues reported that confirmatory information was available from identical findings from several subsequent clinical studies of clonidine (Goldberg et al., 1994; Pandya et al., 2000).

Though self-report of hot flashes, using diaries such as the Hot Flash Daily Diary, has some advantages as a means of recording hot flashes(such as reliable data, low expense and low participant demand) there have been some criticisms of solely using self-reported measures to record hot flashes. A recent review of objective and subjective measures of hot flashes by Sievert (2013) explains that regarding self-report diaries of hot flashes,

> "The experience of hot flashes can be queried as yes/no, as frequency per day or week, or in terms of severity (e.g., a five point scale). Language choice can change the frequency of hot flashes reported, as illustrated in Japan where, as researchers used different words to describe "sudden heat" (Zeserson, 2001),

> reports of hot flashes increased from 12% (Lock, 1986) to 22% (Melby, Lock, & Kaufert, 2005), 37% (Negata et al., 1999) and 45% (Anderson, Yoshizawa, Gollschewski, Atogami, & Courtney, 2004)." (Sievert, 2013, p. 574)

Additionally, self-reported hot flashes using diaries or event markers are subject to recall bias, an issue that Hanisch and colleagues suggest may be more salient in an older population (Hanisch et al., 2009). Hanisch and colleagues compared three techniques of measuring hot flashes, including skin conductance level (SCL)sampled sternally, paper diaries and event markers in 47 men with prostate cancer(Hanisch et al., 2009). The result of this study revealed that paper diaries produced the lowest hourly hot flash rate ($M = 0.17$), which was substantially less than the rates of the objective profile ($M = 0.28$) and event markers ($M = 0.23$). The authors further concluded that the recalled data on paper diaries may not map reliably onto real time. Thus, data suggests that event markers are the optimal self-report measure of hot flashes. However, as the objective profile identified the highest hourly hot flash rate, the authors' position is that the most reliable method of measuring hot flashes is the combined use of event-marking and sternal SCL.

Objective Measure of Hot Flashes

Numerous physiologic instruments have been studied for their ability to detect hot flashes, including finger and core temperature (Tataryn et al., 1981), skin blood-flow and water evaporation (Frödin, Ålund, & Varenhorst, 1985), heart rate and respiratory exchange ratio (Freedman, 2000) skin conductance (Carpenter et al., 2012; de Bakker & Everaerd, 1996; Freedman, 1989; Hanisch, Palmer, Donahue, & Coyne, 2007; Leora C Swartzman, Robert Edelberg, & Ekkehard Kemmann, 1990; Tataryn et al., 1981), and hygrometry (a measure of humidity; Freedman & Wasson, 2007). SCL, sampled

sternally, has been widely adopted as the gold-standard of objective hot flash measurement.

Fundamentals of Skin Conductance Measurement

Skin conductance measurement of hot flashes operates by monitoring SCL increases with sweating. Two silver/silver chloride electrodes are attached to the upper chest on either side to the sternum. Sternal measurement for hot flashes is used as SCL levels are fairly inactive at that site under a normal variety of conditions (Rickles & Day, 1968), and, when examined, the sternal measurement of hot flashes was found to be more sensitive than palmar measurement (de Bakker & Everaerd, 1996). A constant voltage current (0.5 V) is passed from one electrode to the other and conductance is sampled via polygraph or ambulatory monitor typically recorded at a rate of 1 a sharp and rapid rise following by a sloping return to baseline, or ‘‘swishy tail,’’ that can be distinguished from the ‘‘sawtooth’’pattern characteristic of activity or other sweating-related artifact (Carpenter et al., 1999).

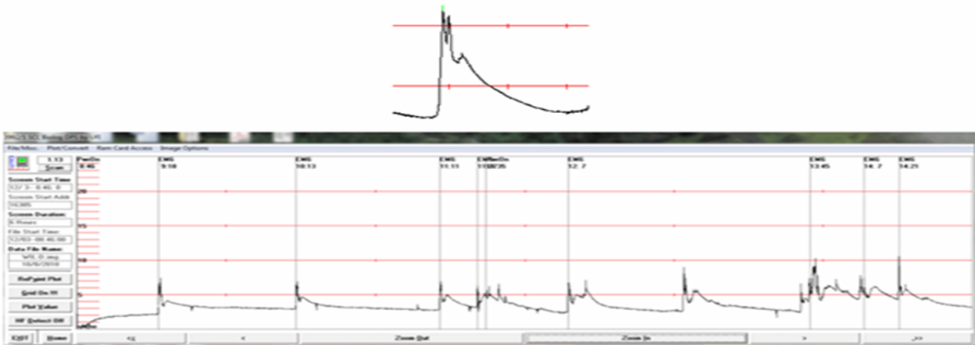


Figure 1. Example of SCL Trace and Hot Flash Shape. Skin conductance level is on the y-axis of the trace, with vertical lines denoting participant self-report of onset of hot flash using an event marker.

Commonly Used Terminology and Analyses

In order to understand how SCL criteria have been investigated, it is necessary to first review the commonly used terminology and analyses associated with subjective-referent methods. A commonly used means of examining SCL criterion performance in detecting hot flashes uses self-report as a referent and involves calculations from the four possible outcomes in the application of the criterion (Table 1):

True Positive: In this outcome, a criterion-derived hot flash corresponds with subjective-report of hot flashes within a defined epoch (e.g., The first incident, the criterion of 2.0 μmh or greater rise sternal skin conductance over 30 seconds, co-occurs with a self-reported hot flash on an event marker, and occurs within 10 minutes of a second incident).

False Positive: In this outcome, a criterion-derived hot flash fails to correspond with subjective-report of hot flashes within a defined epoch (e.g., The first incident, a criterion-selected event involving a ≥2.0 μmh rise in sternal skin conductance over 30 seconds, fails to correspond within 10 minutes to a self-reported hot flash on an event marker).

False Negative: In this outcome, a self-report of hot flashes on an event marker fails to correspond with a criterion-derived hot flash within a defined epoch (e.g., The first incident, a self-reported hot flash on an event marker, does not correspond within 10 minutes to a criterion-matching event involving a ≥2.0 μmh rise in sternal skin conductance over 30 seconds).

True Negative: In this outcome, neither a self-report of hot flashes nor a criterion-derived hot flash occurs within a defined epoch. In this outcome, the negative classification using SCL-criterion is correct. In other words, the monitor using this criterion correctly identified the absence of a hot flash.

Table 1

SCL Hot Flash Detection Confusion Matrix

		Subjective hot flash	
		YES	NO
Criterion-matching skin conductance event	YES	True Positives (TP)	False Positive (FP)
	NO	False Negative (FN)	True Negatives (TN)

Derived from these four possible outcomes, the following performance metrics, common to the study of SCL-derived hot flashes using self-report referents, can be calculated:

Sensitivity (TP/TP+FN): The proportion of detecting an objective hot flash when one has been self-reported. This relates to the power of the metric to detect hot flash.

Specificity (TN/FP+TN): The proportion of epochs correctly identified as a non-flash period. This relates to the test's ability to identify negative results.

Positive Predictive Value (TP/TP+FP): The proportion of criterion-indicated hot flashes with corresponding self-report. This relates to the positive precision of the metric.

Negative Predictive Value (TN/TN+FN): The proportion of negative results that are correctly indicated. This relates to the negative precision of the metric.

Validation of Skin Conductance Measurement of Hot Flashes

In 1981, Tataryn and colleagues , responding to reports that hot flashes correspond to prominent physiologic changes, undertook an investigation to see if measurement of physiologic changes could be used to objectively assess the occurrence of menopausal hot flashes (Tataryn et al., 1981). The study consisted of the laboratory observation of 8 postmenopausal and 4 premenopausal women in continuous recordings of finger and core temperature and sternal skin conductance. Each subject was studied for 8 or 16 hours during which only cool, clear fluids by mouth were allowed, excluding tea and coffee. Participants were lightly clothed and at bed rest in thermally controlled environment and not allowed to smoke. The onset of subjective hot flashes was indicated by the participant via an event marker. The postmenopausal women recorded a total of 120 subjective flashes over 104 hours of study. Results showed that 81% of the subjective hot flashes were associated with changes in finger temperature, skin conductance and core temperature, and 76% were associated with simultaneous changes in all three parameters. The changes in finger temperature and skin conductance were attributed to cutaneous vasodilation and perspiration, respectively. It was further noted that with some hot flashes, rises in skin temperature were not recorded even though flushing and changes in skin conductance were observed. It was concluded that the measurement of sternal skin conductance was the most sensitive and specific single indicator of hot flashes, with a change in resting skin conductance showing a 97% concordance with subjective hot flashes.

Subsequently, in 1989, a study was conducted to evaluate whether sternal skin conductance monitoring for hot flashes was feasible in an ambulatory format (Freedman,

1989). In this study, 8 premenopausal women (age 33-46) who reported regular menstrual cycles and 11 women (age 45-58) who had cease menstruating at least one year earlier and who reported at least four daily hot flashes underwent monitoring over three days. On the first session, the women received a 4-hour baseline laboratory recording session, physiologically monitored using a 12-channel polygraph recording skin temperature and skin conductance while supine in a temperature controlled room, using an event marker to indicate hot flash onset. On a subsequent day, participants wore an ambulatory recorder consisting of a 4-channel cassette recorded monitoring finger temperature, sternal skin conductance level and ECG from 10:00-22:00 hours while engaging in their normal daily activities. They also had an event marker to indicate hot flash onset. On a third day, participants received a peripheral heating test involving two heating pads placed on the torso (to illicit hot flashes) while supine in a temperature controlled room for 30 minutes. All subjects wore hospital gowns and were to activate an event marker to indicate hot flash onset.

Results revealed a total of twenty subjective hot flashes indicated by event marker occurring in 5 of the women, with a range of 1-9 hot flashes (median =3). Examination of the laboratory polygraph tracings showed that all event marks were accompanied by an increase in SCL of at least 2 μS in 30-s, when scored in 15-s epochs for 4 minutes before and 2 minutes after the peak of the SCL response. Only one SCL increase of 2 μS in 30-s occurred without an event mark, for a concordance rate of 95%. ANOVAS on all skin conductance, finger pad temperature, finger dorsum temperature and heart rate measures showed significant epoch effects ($p<.001; p<.01; p<.01; p<.001$, respectively). Criterion SCL changes were not found in premenopausal women during heating or ambulatory

monitoring. The author concluded from this study that a sternal SCL increase of 2 μS was shown to have good concordance with event markers, finding 95% concordance for hot flashes spontaneously recorded in the laboratory, 86% for those recorded during ambulatory monitoring, and 88.9% for heat-induced hot flashes.

In 1990, a combined physiological method for measuring hot flashes to include skin conductance was proposed: the Physiological Flush Profile (PFP; 1990). The primary goal of this study was to develop a criterion for the physiological identification of hot flashes that can be applied regardless of subjective report. In this study, 21 women (age 37-71, *m*=51) who reported at least six hot flashes daily underwent psychophysiological monitoring in three laboratory sessions. Skin conductance, skin temperature and finger pulse volume data were collected. The results of this study revealed that skin conductance corresponds highly with (*r*=.592) subjective report. The authors concluded the PFP revealed a sensitivity of 87% and a specificity of 88.4%, which the authors suggested was comparable to previous developed criteria. Though their results were not as sensitive as Freedman' s results using the now standard 2.0 μmh rise in sternal skin conductance over 30 seconds, they suggested the small sample size of the original investigation pared with the relatively few number of subjective reports (20) may have led to inaccurate calculations of sensitivity and specificity.

In order to determine which method, the PFP or Freedman's standard criterion was optimal, a study in 1995 examined and compared both methods in a sample of 20 menopausal women 15 with frequency hot flashes, 5 without hot flashes) and 5 women who were not menopausal (de Bakker & Everaerd, 1996). In a laboratory, participants were continuously monitored for 2.5 hours using an event marker for subjective report on

hot flash onset as well as sternal and palmar skin conductance, dorsal and palmar finger temperature and pulse blood volume. The results of this study indicated that the PFP was slightly more sensitive, but the standard criterion was much more specific. In the same article, the authors describe a second study where they conducted a cross-validation on an entirely separate sample of 26 menopausal women (de Bakker & Everaerd, 1996). The results of this study revealed a similar profile with the sensitivity of the PFP showing an advantage over the standard criterion in laboratory study (95% vs. 52% in detecting severe hot flashes), however the standard criterion was much more specific (Positive Predictive Value 44.62% for PFP vs. 100% for the standard SCL criterion). The authors concluded that Swartzmann and colleague's (Leora C Swartzman et al., 1990) position that a profile could be created independent of SCL and subjective report was not optimal and that the Freedman standard criterion was a more specific indicator of hot flashes.

The feasibility and psychometrics of a lightweight automated ambulatory sternal skin conductance monitor to measure the frequency of hot flashes was examined in 1999 (Carpenter et al., 1999). In this study of breast cancer survivors, 19 postmenopausal breast cancer survivors and 5 premenopausal healthy comparison women participated by wearing an ambulatory monitor for 24 hours during normal activities, including sleep. Hot flashes were assessed using subjective diaries (both hot flash daily diaries and event markers) and objective (SCL level) methods. Results of this study revealed a concordance of 72% to event marked hot flashes (after controlling for data from a participant who erroneously activated the event marker). False positive rate was 28% and false-negative rates were 41%. Though the authors suggested that their concordance rates was less than they expected given previous reports of high concordance rates (95%

for laboratory study and 86% in Freedman's ambulatory study), they still concluded that ambulatory hot flash detection using lightweight ambulatory sternally recorded skin conductance monitors was feasible.

Challenges with the Ambulatory SCL Measure of Hot Flashes

Sternal skin conductance monitoring for hot flashes has become the gold standard in the physiological measurement of hot flashes. However, there have been concerns over the sensitivity of the instrument in ambulatory settings. In 2004 the National Institutes of Health (NIH) conducted a special workshop, "NIH Workshop on Objective Measure of Hot Flashes" (Miller & Li, 2004), whose summary report identified improvements in sternal skin conductance systems as a critical need. Several studies have shown ambulatory versus in-laboratory SCL monitoring of hot flashes provides much lower levels of concordance between objective and subjective hot flashes (Carpenter et al., 1999; Carpenter & Rand, 2008; Hanisch et al., 2007; E. Mann & Hunter, 2011; Otte et al., 2009; Sievert et al., 2008; Thurston, Blumenthal, Babyak, & Sherwood, 2005; Thurston, Hernandez, Del Rio, & De La Torre, 2011). A systematic review of the concordance of self-reported and sternal skin conductance measures of hot flushes in symptomatic perimenopausal and postmenopausal women found that SCL measurement of hot flashes using standard criterion and self-reported hot flashes as a referent appear to have adequate *specificity,* however, only moderate *sensitivity* (Figure 2; E. Mann & Hunter, 2011).

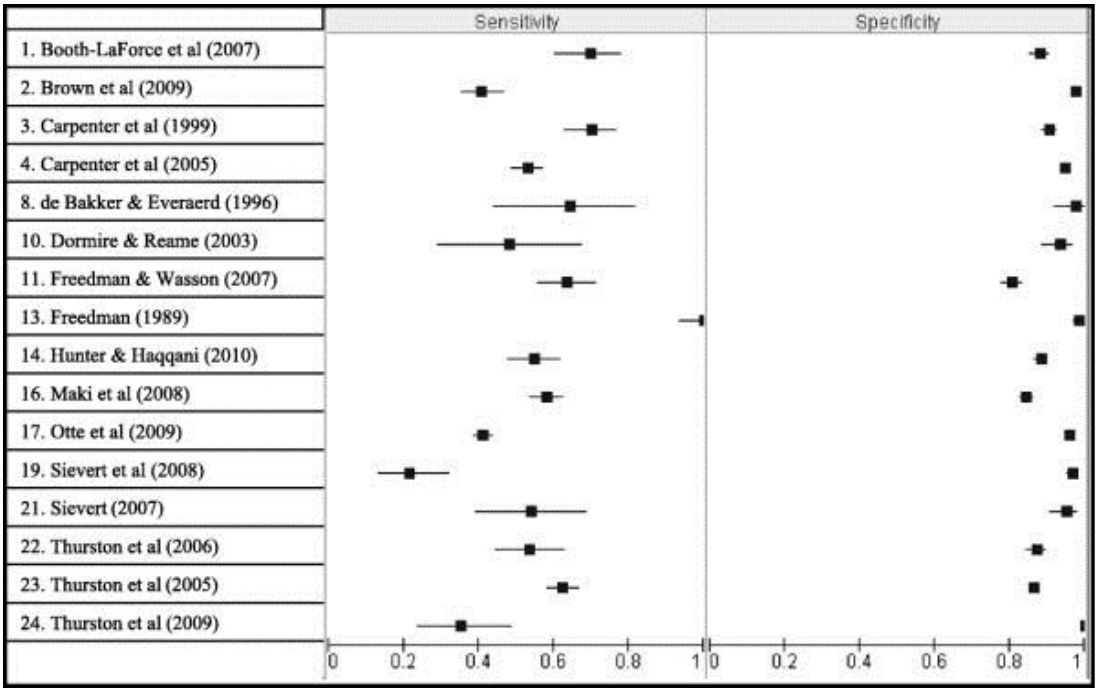


Figure 2. Forest plot of the sensitivity and specificity of self-reported measures to identify sternal skin conductance-defined hot flashes. Adapted with permission from Lippincott Williams and Wilkins/Wolters Kluwer Health: Mann, E., & Hunter, M. S. (2011). Concordance between self-reported and sternal skin conductance measures of hot flushes in symptomatic perimenopausal and postmenopausal women: a systematic review. *Menopause*, *18*(6), 709-722. doi: 10.1097/gme.0b013e318204a1fb

Alternate Models of SCL Hot Flash Detection

Several studies have been conducted in order to attempt to improve the sensitivity of SCL measurement in ambulatory settings. In a study that examined SCL criterion for detecting hot flashes in prostate cancer survivors, 8 participants were recruited to undergo physiological monitoring (Hanisch et al., 2007). Participants were fitted with ambulatory SCL monitors and tested in two phases, a 4-hr laboratory evaluation, and a 24-hour evaluation outside of a lab. Subjective hot flashes were assessed using a hot flash questionnaire and via event marking using the SCL recording device. Results

revealed that 7 men experienced 21 self-reported hot flashes over the course of the study. The standard criterion matched twelve of the self-reported hot flashes, with sensitivity and positive predictive value calculated at 55% and 92%, respectively. The authors performed a Receiver Operation Characteristic curve statistic (Green & Swets, 1966), determining the a 1.78 μmh rise in sternal skin conductance over 45 seconds yielded the optimal criterion for their sample of prostate cancer survivors, providing a sensitivity of 68.2%, and a specificity and positive predictive values at 100%. The sample size in this study suffers the same issues as previous studies of ambulatory hot flash monitoring, with very small sample size, yielding relatively few subjectively recorded hot flashes. Additionally as this sample was entirely male, the generalizability of this criterion to women with hot flashes is limited. This study is important; however, as it suggests that a small change in the criterion used for the gold-standard SCL monitors may yield important improvements in the ability of the measure to detect hot flashes.

Another approach has been examined for the SCL detection of hot flashes using support vector machines as a method of pattern-matching hot flashes. In a study of 31 women with hot flashes who underwent 24 hours of ambulatory sternal skin conductance monitoring, analyses were completed using support vector machines (SVM) in order to classify hot flashes based on their pattern and this was compared to the standard SCL criterion (Thurston et al., 2011). A support vector machine is a sophisticated algorithm that involves a training phase in which the model is developed and a testing phase were the model's performance is evaluated. The process involved preprocessing the SCL data to normalize the SCL signal and remove noise. Momentary voltage drops and baseline drift were also corrected and then exponential smoothing was applied to reduce signal

noise. The SVM model was then trained, qualifying the characteristics associated with hot-flash and non-hot flash segments. Testing was then performed using the trained model. Quantification of standard criterion of a 2.00 μmh rise in sternal skin conductance over 45 seconds to indicate hot flashes showed good specificity (98%), positive predictive value (91%), and negative predictive value (90%), however, it provided a relatively low sensitivity (57%). The SVM technique improved this performance returning good sensitivity (87%), specificity (90%), positive predictive value (90%) and negative predictive value (96%). Though these results are extremely encouraging, specialized expertise is required in order to build an SVM, and although a publically available software library exists, the complexity of this method, specialized software and the time required to perform the analyses will put this technique beyond the reach of many researchers until packaged software is made available. Until such a product is created and achieves wide-spread dissemination, the quest for alternate methods of detecting hot flashes using the most commonly employed monitors and software continues.

Development of an Alternate SCL Criterion for Monitoring Hot Flashes in Breast Cancer Survivors.

In a study of 56 women with breast cancer, researchers collected ambulatory sternal skin conductance data from participants over one home-based daytime recording of hot flashes (M.-H. Savard, Savard, & Ivers, 2013). The study was limited to French-Canadian white women (ages 30-67 years old, $M = 52$), who were undergoing treatment for breast cancer. Six of the participants (11%) were taking medications to manage hot flashes (i.e., venlafaxine, paroxetine or gabapentin), 17 participants (30%) were using another psychotropic medications (e.g., benzodiazepine) and 7 (13%) were using both

types of medication. The researchers used ambulatory skin conductance recorders and two silver/silver chloride electrodes, measuring 1.5 cm in diameter, filled with 0.05 M potassium chloride Valvachol/glycol gel. The electrodes were attached on the chest, 4 cm apart on the sternum. SCL was recorded at a 200-Hz sampling rate, and averaged to 1 hz. The data was collected on PC and traces were visually scanned to detect artifact.

SCL raw data were analyzed with in-house hot flash detection software(UFI, 2006). The researchers examined a range of SCL criterion compared to the standard SCL criterion. Using participant report of onset of hot flash on an event marker, they found that the sensitivity of the standard criterion (≥2 μmho rise in a 30-s period) in identifying hot flashes was 32.7% (95% CI, 26.5-39.5), whereas they found that a more sensitive criteria was a 1.2 μmho rise in a 30-s period, yielding a sensitivity of 60.8%, specificity of 90.4%, and a positive predictive value of 86.4%. The authors concluded that an alternate criterion may be preferable in detecting hot flashes in women with breast cancer. It is notable that though the alternate criteria appeared to be more sensitive, the standard criteria may be more specific (97% vs. 90.4%, statistical significance of difference not reported), with greater positive predictive value (91.5% vs. 86.4%, statistical significance of difference not reported); however, it is the conclusion of the authors that, for ambulatory measurement, the alternate criterion was an improvement over the standard criterion as they suggest that it is more important that the selected criterion be able to detect a maximal number of events identified by women as a hot flash (M.-H. Savard et al., 2013).

There were a number of limitations to this study. First, the average number of hot flashes reported by participants was not reported; rather, only a total number of hot

flashes manually detected by participants were reported (199 hot flashes). Second, the sample of study was taken from the entire menopausal spectrum. While this might be advantageous in clinical study, changes of physiological markers, including skin conductance, have shown significant differences between pre- and post-menopausal women possibly confounding the results of a validation study (Tataryn et al., 1981). Third, the sample used in the study was ethnically homogenous, which does not answer to how the criterion performs across different ethnic groups. Fourth, this study did not compare the results of using the alternate SCL criteria to participants' perceived impact of hot flashes. One might predict that an improved measure of hot flashes should correspond to the perceived impact of hot flashes; however, this remains an unanswered empirical question. Lastly, the participants in this study were undergoing hormone therapy and taking other pharmacy, including medications specifically to reduce hot flashes. The effect of these agents on the measure of hot flashes via sternal skin conductance is unknown. Those limitations aside, the results of this study are important because if, as Savard and colleagues (2013) suggest, the alternate skin conductance criteria (1.2 μmho rise in a 30-s period) is more sensitive in ambulatory monitors, this small change may nearly double the sensitivity of the gold-standard instrument for measuring hot flashes. Additionally, this study is important for its potential in broad and immediate dissemination as this criterion change can be easily achieved using the current gold-standard SCL monitors and hot flash detection software. Further study, however, is critically needed to replicate their results with this criterion and ascertain how this criterion performs in healthy post-menopausal women and among different ethnicities.

Purpose

The purpose of this dissertation is to examine an alternate criterion for hot flash detection in ambulatory skin conductance monitors, explore ethnic variability, and determine the correlation of the alternate criterion-derived hot flashes to hot-flash related daily interference in healthy post-menopausal women. Current SCL criterion, although very effective in laboratory trials, have demonstrated only moderate sensitivity in ambulatory trials. An alternate criterion has been proposed for ambulatory use in breast cancer patients with potentially substantial gains in sensitivity. It is unknown, however, how this criterion performs in a post-menopausal sample and among diverse ethnic groups. The relationship of this measure to hot flash related daily interference is also unknown. This study addresses these empirical questions.

CHAPTER TWO

Materials and Methods

Objectives

The objectives of this dissertation are threefold. The *first objective* is to compare an alternate skin conductance criterion (≥ 1.2 μmho rise in a 30-s period) to the standard criterion (≥ 2 μmho rise in a 30-s period) in detecting hot flashes using ambulatory sternal skin conductance hot flash monitors in symptomatic post-menopausal women using participant self-report on an event marker as the referent. This objective is achieved by the analyses of four indices of monitor performance, *sensitivity* (the proportion of detecting an objective hot flash when one has been self-reported), *specificity* (the proportion of recording the absence of a hot flash when none has been self-reported), *positive predictive value* (the proportion of criterion-indicated hot flashes with corresponding self-report) and *negative predictive value* (the proportion of correctly identified negative epochs). The *second objective* is to determine group differences between the alternate and standard skin conductance criterion among self-identified Black, Hispanic and White ethnic groups in terms of the indices of monitor performance: *sensitivity, specificity, positive predictive value* and *negative predictive value*. The *third objective* is to determine and compare the correlations of the alternate SCL criteria versus the standard criteria to each of the ten components of the Hot Flash Related Daily Interference Scale (HFRDIS) including hot-flash related interference with *Work, Social Activities, Leisure Activities, Sleep, Mood, Concentration, Relationships, Sexuality, Enjoyment of Life* and *Overall Quality of Life* in symptomatic post-menopausal women.

Specific Aims

This dissertation has three specific aims:

Aim 1: Compare the results of the alternate SCL criterion for identifying hot flashes (≥ 1.2 μmho rise in a 30-s period) to the standard criterion (≥ 2 μmho rise in a 30-s period) using participant self-report of onset of hot flashes on the SCL monitor event marker as a referent to ascertain their relative *sensitivity* (TP/TP+FN)*, specificity* (TN/FP+TN), *positive predictive value* (TP/TP+FP) and *negative predictive value* (TN/TN+FN) in post-menopausal women.

> *H 1.1* The *sensitivity* (TP/TP+FN)*,* of the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show a significantly greater proportion in detecting an objective hot flash when one has been self-reported than the standard criterion (≥ 2 μmho rise in a 30-s period).
>
> *H 1.2* The *specificity* (TN/FP+TN) of the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show, as was shown in previous study of breast cancer survivors, a significantly lesser proportion in recording the absence of a hot flash when none has been self-reported than the standard criterion (≥ 2 μmho rise in a 30-s period).
>
> *H 1.3* The *positive predictive value* (TP/TP+FP) of the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show a significantly greater proportion of criterion-indicated hot flashes with corresponding self-report than the standard criterion (≥ 2 μmho rise in a 30-s period).
>
> *H 1.4* The *negative predictive value* (TN/TN+FN) of the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show a greater proportion of correctly identified negative results than the standard criterion (≥ 2 μmho rise in a 30-s period).

Aim 2: Determine group differences between the alternate and standard skin conductance criterion among self-identified post-menopausal American Indian, Black, Hispanic and White participants, grouped into White and non-White categories, in terms of four indices of monitor performance: *sensitivity, specificity, positive predictive value, and negative predictive value*.

H 2.1 The *sensitivity* of the standard SCL criterion for identifying hot flashes (≥ 2.0 μmho rise in a 30-s period) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show *sensitivity* differs significantly among White and non-White racial/ethnic groups.

H 2.2 The *specificity* of the standard SCL criterion for identifying hot flashes (≥ 2.0 μmho rise in a 30-s period) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show *specificity* differs significantly among White and non-White racial/ethnic groups.

H 2.3 The *positive predictive value* of the standard SCL criterion for identifying hot flashes (≥ 2.0 μmho rise in a 30-s period) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show *positive predictive value* differs significantly among White and non-White racial/ethnic groups.

*H 2.*4 The *negative predictive value* of the standard SCL criterion for identifying hot flashes (≥ 2.0 μmho rise in a 30-s period) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show the *negative predictive value* differs significantly among White and non-White racial/ethnic groups.

H 2.5 The *sensitivity* of the alternate SCL criterion for identifying hot flashes (≥ 1.2 μmho rise in a 30-s period) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show *sensitivity* differs significantly among White and non-White racial/ethnic groups.

H 2.6 The *specificity* of the alternate SCL criterion for identifying hot flashes (≥ 1.2 μmho rise in a 30-s period) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show *specificity* differs significantly among White and non-White racial/ethnic groups.

H 2.7 The *positive predictive value* of the alternate SCL criterion for identifying hot flashes (≥ 1.2 μmho rise in a 30-s period) in ambulatory hot flash detection

over one day's waking hours in post-menopausal women will show *positive predictive value* differs significantly among White and non-White racial/ethnic groups.

H 2.8 The *negative predictive value* of the alternate SCL criterion for identifying hot flashes (≥ 1.2 μmho rise in a 30-s period) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show the *negative predictive value* differs significantly among White and non-White racial/ethnic groups.

Aim 3: Determine and compare the relative correlations of the alternate SCL criteria versus the standard criteria to the 10 components of the HFRDIS including the hot-flash related interference with *Work, Social Activities, Leisure Activities, Sleep, Mood, Concentration, Relationships, Sexuality, Enjoyment of Life* and *Overall Quality of Life* in post-menopausal women.

H 3.1 Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the standard SCL criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with work as indicated on scores on the *Work* construct of the HFRDIS in post-menopausal women.

H 3.2 Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the standard SCL criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with social activities as indicated on scores on the *Social Activities* construct of the HFRDIS in post-menopausal women.

H 3.3 Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the standard SCL criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with leisure activity as indicated on scores on the *Leisure* construct of the HFRDIS in post-menopausal women.

H3.4 Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the standard SCL criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with sleep as indicated on scores on the *Sleep* construct of the HFRDIS in post-menopausal women.

H3.5 Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the standard SCL criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related effect on mood as indicated on scores on the *Mood* construct of the HFRDIS in post-menopausal women.

H3.6 Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the standard SCL criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with concentration as indicated on scores on the *Concentration* construct of the HFRDIS in post-menopausal women.

H3.7 Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the standard SCL criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with relationships as indicated on scores on the *Relationships* construct of the HFRDIS in post-menopausal women.

H3.8 Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the standard SCL criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related effect on sexuality as indicated on scores on the *Sexuality* construct of the HFRDIS in post-menopausal women.

H3.9 Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the standard SCL criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with enjoyment of life as indicated on scores on the *Enjoyment of Life* construct of the HFRDIS in post-menopausal women.

H3.10 Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the standard SCL criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with overall quality of life as indicated on scores on the *Overall Quality of Life* construct of the HFRDIS in post-menopausal women.

H 3.11 Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with work as indicated on scores on the *Work* construct of the HFRDIS in post-menopausal women.

H 3.12 Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot

flash-related interference with social activities as indicated on scores on the *Social Activities* construct of the HFRDIS in post-menopausal women.

H 3.13 Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with leisure activity as indicated on scores on the *Leisure* construct of the HFRDIS in post-menopausal women.

H3.14 Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with sleep as indicated on scores on the *Sleep* construct of HFRDIS in post-menopausal women.

H3.15 Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related effect on mood as indicated on scores on the *Mood* construct of the HFRDIS in post-menopausal women.

H3.16 Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with concentration as indicated on scores on the *Concentration* construct of HFRDIS in post-menopausal women.

H3.17 Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with relationships as indicated on scores on the *Relationships* construct of the HFRDIS in post-menopausal women.

H3.18 Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related effect on sexuality as indicated on scores on the *Sexuality* construct of the HFRDIS in post-menopausal women.

H3.19 Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with enjoyment of life as indicated on scores on the *Enjoyment of Life* construct of the HFRDIS in post-menopausal women.

H3.20 Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with overall quality of life as indicated on scores on the *Overall Quality of Life* construct of the HFRDIS in post-menopausal women.

H 3.21 The correlations between the monitor-verified hot flashes using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) and the standard criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will differ significantly in their relationship to hot flash-related interference with work as indicated on scores on the *Work* construct of the HFRDIS.

H 3.22 The correlations between the monitor-verified hot flashes using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) and the standard criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will differ significantly in their relationship to hot flash-related interference with social activities as indicated on scores on the *Social Activities* construct of the HFRDIS.

H 3.23 The correlations between the monitor-verified hot flashes using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) and the standard criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will differ significantly in their relationship to hot flash-related interference with leisure activities as indicated on scores on the *Leisure* construct of the HFRDIS.

H 3.24 The correlations between the monitor-verified hot flashes using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) and the standard criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will differ significantly in their relationship to hot flash-related interference with sleep as indicated on scores on the *Sleep* construct of the HFRDIS.

H 3.25 The correlations between the monitor-verified hot flashes using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) and the standard criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will differ significantly in their relationship to hot flash-related impact on mood as indicated on scores on the *Mood* construct of the HFRDIS.

H 3.26 The correlations between the monitor-verified hot flashes using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) and the standard criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will differ significantly in their relationship to hot flash-related interference with concentration as indicated on scores on the *Concentration* construct of the HFRDIS.

H 3.27 The correlations between the monitor-verified hot flashes using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash)

and the standard criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will differ significantly in their relationship to hot flash-related interference with relationships as indicated on scores on the *Relationships* construct of the HFRDIS.

H 3.28 The correlations between the monitor-verified hot flashes using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) and the standard criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will differ significantly in their relationship to hot flash-related interference with sexuality as indicated on scores on the *Sexuality* construct of the HFRDIS.

H 3.29 The correlations between the monitor-verified hot flashes using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) and the standard criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will differ significantly in their relationship to hot flash-related impact on the enjoyment of life as indicated on scores on the *Enjoyment of Life* construct of the HFRDIS.

H 3.30 The correlations between the monitor-verified hot flashes using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) and the standard criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will differ significantly in their relationship to hot flash-related impact on quality of life as indicated on scores on the *Overall Quality of Life* construct of the HFRDIS

Participants

Data used in this study was collected during a NIH-funded clinical trial of hypnosis for the treatment of hot flashes (Elkins, Fisher, Johnson, Carpenter, & Keith, 2012). Participants were recruited for that study from December, 2008 to April, 2012 via advertisement in newspapers and billboards, television, and professional referrals, and television and billboard advertisement. Eligibility requirements involved participants being postmenopausal, no menstruation in the past 12 months, and either (a) a medically documented history of a follicle-stimulating hormone level greater than 40 or (b) a bilateral oophorectomy. Participants were also required to have a self-report of seven hot flashes per day at minimum, or 50 days per week, at baseline. Additionally, the participants in this study were required to have discontinued estrogen/progestin

containing products: one week or longer for prior vaginal hormonal products (rings, creams, gels), four weeks or longer for prior transdermal estrogen alone or estrogen/progestin products, eight weeks or longer for prior oral estrogen and/or progestin therapy, eight weeks or longer for prior intrauterine progestin therapy, three months or longer for prior progestin implants and estrogen alone injectable drug therapy, and six months or longer for prior estrogen pellet therapy or progestin injectable drug therapy.

Exclusion criteria included: a) receiving any other treatment for hot flashes, b) using any complementary or alternate medical treatments for vasomotor symptoms (including soy, black cohosh, phytoestrogens, and any other mind-body techniques), c) History of psychosis, borderline personality disorder of serious psychopathology that were considered contraindications to treatment with clinical hypnosis. Potential participants were telephone screened for eligibility and those who met eligibility criteria completed baseline measures including the objective measure of hot flashes using SCL-monitors (Biolog Model 3991; UFI, 2012), the Hot Flash Daily Interference Scale (Carpenter, 2001), demographic questionnaires and other secondary outcome measures. Of those participants who were screened for eligibility (531), 140 participants were deemed eligible, completed baseline measures and underwent physiological monitoring (Elkins et al., 2012).

Measures

Biolog® Hot Flash Monitor

The Biolog® Hot Flash Monitor (UFI, 2012), is the most commonly used instrument for the sternal skin conductance measure of hot flashes. The monitor is housed in a compact plastic case measureing 5" x 2.5" x 1.25", and weighs approximately 8 ounces with a battery. The monitor is a solid state device, containing a microprocessor programmed to sample 12-bit skin conductance data at 1 Hz, at a range of 0.01 to 39.9 μmho and 4 megabytes of memory. The monitor is powered with non-proprietary, commercially available 9V alkaline batteries, which the manufacturer suggests that the monitor can maintain in constant operation for approximately 14 days. The monitor was used with Biopac EL-507 silver/silver chloride electrodes for electrodermal activity and a 0.5 constant voltage circuit (BIOPAC Systems, 2012). Electrodes are 1.0 cm in diameter and filled with 0.5% chloride solid gel. Electrodes are attached 1.5" below the collarbones and 2" on either side of the sternal mid-line. Participants are instructed to remove the electrodes prior to bathing and trained on how to replace them immediately after bathing. The monitor has dual buttons and when pressed simultaneously will record an event marker at the time of button press. Participants are instructed to press both buttons simultaneously at the onset of hot flashes. Customized software (FlashTrax, version 1.2, UFI, Morro Bay, CA) is used to evaluate hot flashes providing the ability to analyze traces based on variable SCL-criterion, defaulting to the standard (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash), with a 15-minute lock-out period post-flash. Software output provides a visual output of the trace (Figure 2), and can also calculate

true positives, false positives and false negatives for every event (self-report on marker or SCL-event) detected.

Hot Flash Related Daily Interference Scale

The Hot Flash Related Daily Interference Scale (HFRDIS; Carpenter, 2001), was developed to measure the impact of hot flashes on overall quality of life as well as on nine other specific activities and experiences (i.e. work, social activities, leisure activities, sleep, mood, concentration, relations with others, sexuality, enjoyment of life). The HFRDIS consists of a series of 0–10 point numeric rating scales modeled on the Brief Pain Inventory (Daut, Cleeland, & Flannery, 1993) and the Fatigue Symptom Inventory (Hann et al., 1998). Like the Brief Pain Inventory and the Fatigue Symptom Inventory which invite respondents to rate the degree that the symptom (pain or fatigue) interferes with various daily activities as well as overall enjoyment or quality of life, the HFRDIS was designed in the same format. Scale reliability was examined using inter-item correlations at two time points. Test re-test was not employed as it was expected to be low given the unreliable nature of the phenomena. Inter-item correlations were examined for internal consistency with Time 1 correlations ranging from 0.55 to 0.90 and 0.59 to 0.95 at Time 2. Item-total correlations ranged from 0.77 to 0.93 at Time 1 and 0.82 to 0.93 at Time 2. Cronbach's alpha was 0.96 at both Times 1 and 2. Validity was supported through correlations with other hot flash variables, correlations with measures of affect and mood, significant differences between women with hot flashes and those without ($p<.001$), and demonstrated sensitivity to change over time. The HFRDIS was deemed "…a psychometrically sound measure for assessing the impact of hot flashes on

daily activities and overall quality of life in clinical practice or in research protocols." (Carpenter, 2001, p. 979).

Demographic Questionnaire

Participants were provided a questionnaire that ascertained a participant's age, self-identified ethnic group, marital status, education, body-mass index, smoking, alcohol use, number of months since onset of amenorrhea, and average daily hot flash frequency & severity.

Procedure

Participants who were deemed eligible came to the laboratory and completed baseline measures, including a demographic questionnaire and the Hot Flash Daily Interference Scale. Following the completion, participants were fitted with skin-conductance monitors. Sternal skin conductance monitors as shown in Figure 3 (Biolog ambulatory recorder model 3991; UFI, 2012) were activated and fitted in the morning (~0900 hours) by trained study staff.

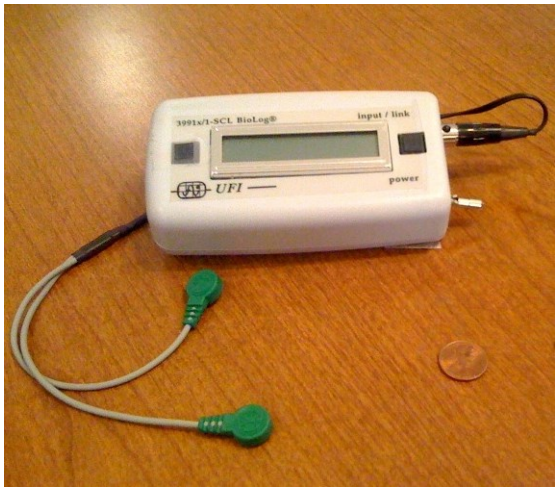


Figure 3. Biolog® Model 3991 SCL Monitor

After participants were fitted with their monitors, they were instructed in its' care and use. Participants were instructed to wear the monitors for 24-hours and activate the event marker at the onset of every hot flash and instructed on how to activate the marker and how to verify that the event was successfully recorded (the Biolog® monitor has a small LCD display that indicates that the signal was recorded).

Although 24-hours of data was collected from the participants, only one day's waking-hours data was used in the analyses in this study, in keeping with standard practice using self-reported hot flashes as a referent (M.-H. Savard et al., 2013). Participants were further instructed to avoid showering and bathing during the recording, if possible, but instructed on the removal and placement of new electrodes should removal be unavoidable. In order to minimize monitor drops, the participants were provided with a small pouch to place the monitor, where the electrodes could freely reach from the monitor to the electrodes on the participant, but the monitor could remain in the zippered pouch, which could be worn either belt-like or across the shoulder. Participants were not instructed to change their daily routine in any manner, simply to record their hot flashes on the event marker at every hot flash onset, and to return the monitor to the laboratory after 24-hours. After returning the monitor to the laboratory, data was downloaded to PC and customized software (FlashTrax Version 1.2; UFI, 2006) was used to evaluate hot flashes. After data were collected, every trace was visually scanned by a trained expert for artifact.

Using the proprietary software (FlashTrax Version 1.2; UFI, 2006), hot flashes will be detected if there is an increase of sternal skin conductance of 2 μmho (standard criterion), or 1.2 μmho increase within a 30 second period, with a 20-minute post-event

lock-out. Using established methodology (Sievert, 2013), indices of hot flash detection (true positives, false positives, false negatives, true negatives) will be calculated and *sensitivity, specificity, positive predictive value* and *negative predictive value* will be ascertained for both criteria.

Data Analysis

The type of inferential statistics for this study must take into account the nature of the performance metrics used to evaluate test criteria. These metrics were derived from signal detection theory, originally used to evaluate test criteria for radar installations (Abdi, 2007). The need was for the radar to have adequate sensitivity to detect true enemy aircraft, while maintaining good specificity to avoid falsely identifying friendly aircraft. Since improving sensitivity tends to decrease specificity, the performance metrics helped to evaluate the best criteria for the radar test so as to provide optimum test results.

These same performance metrics, which include sensitivity, specificity, positive predictive value, and negative predictive value, have been used extensively in medicine to evaluate medical tests (Zweig & Campbell, 1993). In medicine, the need is to have high sensitivity to detect a disease (e.g., cancer), while maintaining good specificity to avoid a false diagnosis. Each of these metrics represents a proportion (e.g., the proportion of hits, or the proportion of misses) as derived from what is called the "confusion matrix" that shows the following: the frequency of positive tests that are true (true positives), the frequency of positive tests that are false (false positives), the frequency of negative tests that are true (true negatives), and the frequency of negative tests that are false (false negatives). In medical tests, the test is either positive or negative

and the result is compared to a gold standard of diagnostic outcome, which is also either positive of negative, resulting in a 2 by 2 matrix.

In applying this approach to evaluate the two objective SCL criteria (standard and alternate) for hot flash, the gold standard would be the self-report of hot flash. So for each of the criteria being studied, standard and alternate, a confusion matrix like that in Table 1 would establish the proportions for each of the four metrics. For each SCL criterion, the four metrics are as follows: 1) *Sensitivity* is the proportion of true positives over all self-report positives (the sensitivity or power to find true positives; 2) *Specificity* is the proportion of true negatives over all self-report negatives (the specificity to correctly identify true negatives); 3) *Positive predictive value* is the proportion of true positives over all the SCL test positives (the precision of positive SCL results); (4) *Negative predictive value* is the proportion of true negatives over all the SCL test negatives (the precision of negative SCL results) .

Since all the performance metrics are proportions, the use of inferential statistics must take into account the fact that proportions do not satisfy the underlying assumptions of parametric statistics. Therefore, the analysis of significant differences in the current study will employ non-parametric tests.

Aim 1: Data Analyses

The objective for this aim is to compare, in ambulatory measurement, the results of the alternate SCL criterion for identifying hot flashes (≥ 1.2 μmho rise in a 30-s period) to the standard criterion (≥ 2 μmho rise in a 30-s period) using participant self-report of onset of hot flashes on the SCL monitor event marker as a referent to ascertain their relative *sensitivity* (TP/TP+FN)*, specificity* (TN/FP+TN) and *positive predictive value*

(TP/TP+FP) and *negative predictive value* (TN/TN+FN), in post-menopausal women. The dependent variable for this aim is the difference in metric proportions between two criteria from the same population (standard and alternate criteria), so this requires a test for correlated proportions. The non-parametric test that makes the fewest assumptions about distributions of correlated proportions is the Sign Test (Siegal, 1956). Another alternative would have been the Wilcoxon Signed Rank Test, but that test assumes that the distribution of scores is symmetric around the median, and that may not be the case in the current metrics. Therefore, the Sign Test was chosen as the inferential statistic for this aim.

In order to test the hypotheses of this aim, analyses of the SCL-trace will be performed using FlashTrax software (UFI, 2006). The software will record an event every time the skin conductance level of the participant meets the programmed criteria or after the participant depresses the event marker to signify a hot flash. After a hot flash has been registered the software will maintain a lock-out, where no further hot flashes will be registered for the following 20 minutes. Both the alternate SCL criterion ($\geq$ 1.2 μmho rise in a 30-s period to indicate a hot flash) and the standard SCL criterion ($\geq$ 2.0 μmho rise in a 30-s period to indicate a hot flash) will be employed and summations of *true positives, false positives, false negatives* and *true negatives* will be ascertained for each criterion. *Sensitivity, specificity, positive predictive value* and *negative predictive value* will be calculated for each criterion. The hypotheses will be supported if the results of a Sign Test (Siegel, 1956), shows that the alternate criterion sensitivity is significantly ($p < .05$) greater (save for *specificity* where hypothesis is supported if significantly lesser) than the standard criterion.

Aim 2: Data Analyses

The objective for this aim is to determine if there are racial/ethnic group differences in *sensitivity, specificity, positive predictive value and negative predictive value*), computed separately for the standard as opposed to the alternate SCL criteria. Though descriptive data will be provided for all racial/ethnic groups, there is insufficient power to test significance for all identified racial/ethnic groups. To allow for statistical comparisons, a bivariate White/non-White category will be employed. The dependent variable for this aim is the difference in metric proportions between racial/ethnic groups (White and non-White). As these are independent groups, this requires a test for independent proportions. The non-parametric test for independent proportions is the Mann Whitney U Test (H. B. Mann & Whitney, 1947). This test will be run for sensitivity, specificity, positive predictive value and negative predictive value individually, with a significant finding ($p < .05$), indicating a significant difference in SCL-indices between White and non-White racial/ethnic groups.

Aim 3: Data Analyses

The objective for this aim is to determine and compare the relative correlations of the alternate SCL criteria versus the standard criteria to the components of Hot Flash Related Daily Interference Scale including the hot-flash related interference with *Work, Social Activities, Leisure Activities, Sleep, Mood, Concentration, Relationships, Sexuality, Enjoyment of Life* and *Overall Quality of Life in post-menopausal women.* The dependent variable for this aim is the relationship between the scores for true positive hot flash and the score on the components of the HFRDIS. For this, a Pearson Product Moment Correlations Coefficient (Rodgers & Nicewander, 1988) can be used to show

the extent of the relationship and determine if that relationship is significant. For this aim Pearson Product Movement Correlations will be performed for both the standard and alternate criteria true positive (criterion-derived hot flash in conjunction with self-report) hot flashes and all components of the HRFDIS, individually. Significance will be determined by t-test (.05 level of confidence). To control for family-wise error, Benjamini and Hochberg's False Discovery Rate will be employed (Benjamini & Hochberg, 1995). Should both criteria show significance on a construct of the HFRDIS, a Hotelling/William's procedure will be run to determine statistically significant (.05 level of confidence) differences exist between the two sets of dependent correlations.

Effect Size Estimates

With regard to the effect size, Cohen (1988) established guides for effect size for different test statistics. The most well-known of these effect size statistics is for the two group Student's t test which has an effect size (*d*) that represents the size of the difference over the standard error of the difference. For the difference between two correlated proportions (Aim 1), the effect size (*g*) is the absolute distance between the proportion difference and a proportion of *.50*. The standard is: small *g* = .05, medium *g* = .15, and large *g* = .25. For the difference between two independent proportions (Aim 2), the effect size (*h*) is based on an *arcsin* transformation ($\phi = \sqrt{P}$). The effect size *h,* then, is the absolute difference between ϕ1 and ϕ2. The standard is: *h* = .20 is a small effect, *h* = .50 is a medium effect and *h* = .80 is a large effect. The result of a Pearson Product Moment Correlation Coefficient is an effect size ranging from 0-1.0, with a correlation of .10 considered small, .30 considered moderate, and .50 considered a strong relationship (effect) (Cohen, 1988).

Justification for Non-Parametric Analyses

The dependent variables in the four performance metrics (sensitivity, specificity, positive predictive value, and negative predictive value) are proportions, and proportions violate assumptions of normality and homogeneity of variance in parametric tests, a problem that gets worse with unequal and small samples. Proportions have fixed limits (0 -1) and scores may cluster at the extremes, producing a skewed distribution (Bartlett, 1947). Also, the variance of smaller samples may be different than the variance of larger samples, and larger proportions tend to have lower variance while smaller proportions tend to have higher variance (Zubin, 1935). One common solution is to use a logit (logistic) transformation (logit p = log (p/1-p). Transformations may also be problematic for several reasons. First of all, data transformations alter the fundamental nature of the data, complicating the interpretation of the results. Also, proportions that are strongly clustered at the very extremes (0 or 1) can make transformations less effective (Carey, 2012) and this is evident in the current data (Figure 4) on racial differences. For this reason, rather than depend on transformations, non-parametric tests will be employed.

Frequency	*Stem*	*Leaf*
11	0	0 0 0 0 0 0 0 0 0 0 0
1	.1	2
4	.2	0 5 8 8
8	.3	3 3 3 3 3 3 7 7
11	.4	0 0 0 2 2 2 2 2 4 6 6
5	.5	0 0 5 5 8
10	.6	0 2 2 2 6 6 6 6 6 6
7	.7	1 1 1 5 5 5 6
14	.8	0 0 0 0 0 0 3 3 4 4 5 7 8 8
3	.9	0 0 2
18	1.0	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

Note: Stem width = .1, each leaf = 1 case.

Figure 4. Stem-and-Leaf Plot: Sensitivity proportions for White racial/ethnic group.

CHAPTER THREE

Results

Demographic Variables: Descriptive Statistics

The characteristics of the sample (n = 140) are described in Table 2. Since this is a study of post-menopausal women, the age range begins in middle age (range = 39-75), with a mean age of 54.66. With regard to race, the majority (72.9%) of the sample was White, with African American being second in percentage (17.9%). The percentage of American Indian and Hispanic was fairly small (2.1% and 7.1% respectively). The majority were married (65.7%), college educated (92.65%) and non-smokers (82%). The mean body mass index was 29.38 kg/m2 (*SD*= 6.17, range 20-47), and the mean number of months since onset of hot flashes was 90.30 (*SD*= 23.92, range 49-170). This population had a high rate of hot flashes (range = 30-170) with a mean weekly flash frequency of 70.14 (*SD* = 23.92, range 49-170). The hot flash composite score, based on the product of average daily hot flash frequency and ratings of severity (1 = mild, 2 = moderate, 3 = severe, 4 = very severe) showed a mean of 23.268 (*SD* = 10.54, range 7.00-70.71).

Comparing the Alternate and Standard Criteria in Sensitivity, Specificity, Positive and Negative Predictive Values.

The comparison of the standard criterion and the alternate criterion in all four values: *Sensitivity, Specificity, Positive Predictive Value*, and *Negative Predictive Value* is presented in Figure 5. Since these four values all represent correlated proportions, the Sign Test (Seigel, 1956) is employed (a two group test for related samples), using a

significance level of alpha .05 and 95% confidence interval. The sign test is distribution-free, with no assumptions of: normality, equal variances, or symmetric distributions around the median. The effect size of differences between correlated proportions was established by Cohen (Cohen, 1988) as the absolute value of *g* minus .50 (P = the proportion of the larger differences, and *g* = absolute value of P-.50). An effect size of *g* =. 05 would be a small effect size, *g* = .15 is a medium effect size, and *g* = .25 is a large effect size.

Table 2

Demographics

Characteristics	*n* = 140
Age group, *n* (%)	
35-44 *y*	10 (7.1%)
45-54 *y*	60 (42.9%)
55-65 *y*	57 (40.7%)
>65 *y*	13 (9.3%)
Age, mean (range)	54.66 (39-75)
Race, *n (*%)	
American Indian	3 (2.1%)
African American	25 (17.9%)
White	102 (72.9%)
Hispanic	10 (7.1%)
Marital status, *n* (%)	
Married	92 (65.7%)
Divorced	15 (10.7%)
Separated	7 (5%)
Single	23 (15.7%)
Widowed	4 (2.9%)
Education, *n* (%)	
Less than high school	13 (9.3%)
High school diploma	35 (25%)
College, non-degree	28 (20%)

(Continued)

Associate degree	21 (15%)
Bachelor's degree	28 (20%)
Graduate degree	15 (10.7%)
Body Mass Index, mean (SD), kg/m^2	29.38 (6.17)
<25	38 (27.1%)
25 to <30	45 (32.1%)
≥30	57 (40.7%)
Smoking, *n* (%)	
Never	114 (82%)
Once a week	1 (.7%)
Several times per week	4 (2.9%)
Daily	21 (15%)
Alcohol use, *n* (%)	
Never	98 (69.7%)
Once a week	21 (15%)
Several times per week	13 (9.3%)
Daily	7 (5%)

Note. Not all percentages will add to 100% due to rounding.

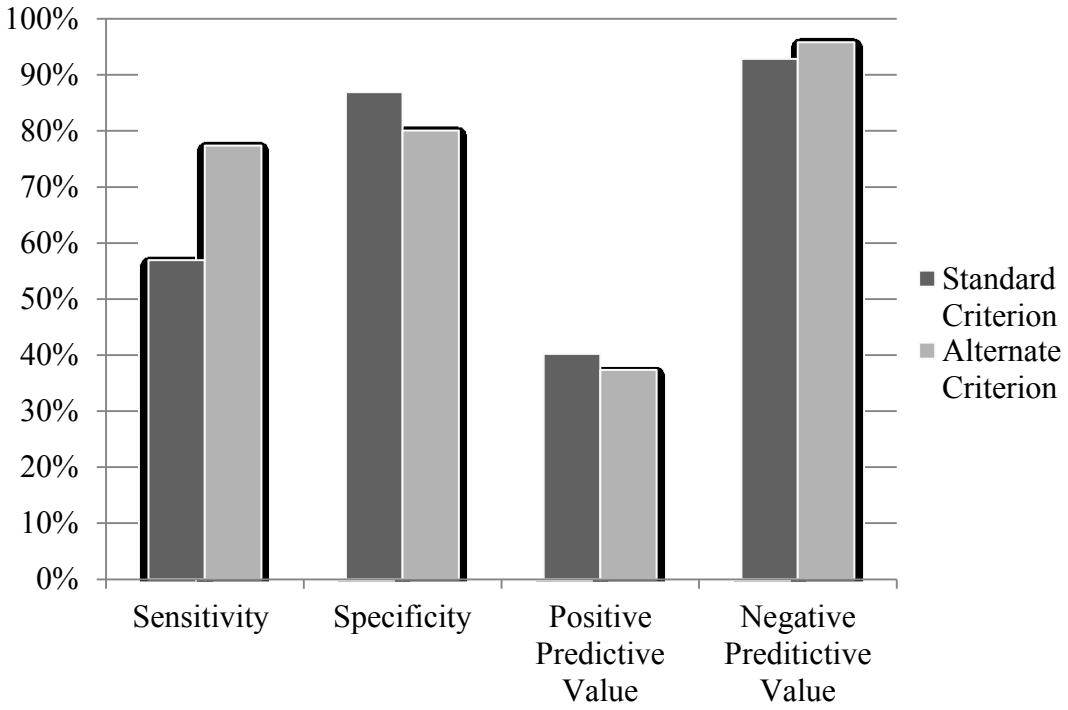
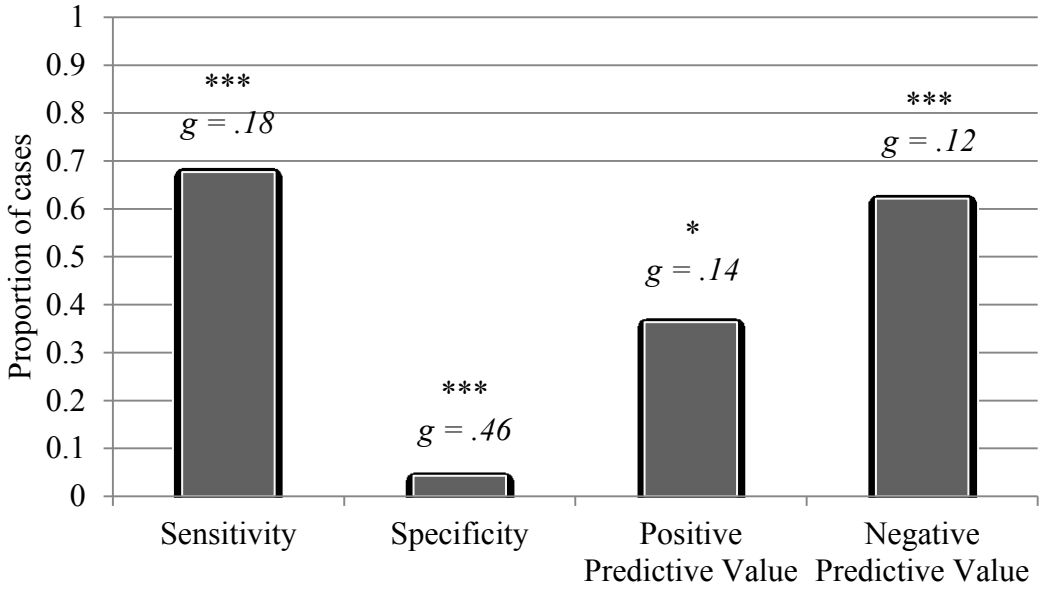


Figure 5. Sensitivity, specificity, positive predictive value and negative predictive value of the alternate and standard criteria.

Consistent with the prior study of cancer patients, this larger sample of post menopause women shows that the *sensitivity* of the alternate criterion (77.4%) is significantly better than the sensitivity of the standard criterion (57%). The alternate criterion was 35.8% more sensitive in detecting hot flashes, and about 68% of the individual cases showed a positive difference in favor of the alternate criterion (alternate criteria = 77.4%; standard criteria = 57%; positive differences = 90 of 133 = .6767 in favor of the alternate criterion, Sign Test significant, $p < .001$, standard error of differences = 4.899, $S_s = 90$; 95% CI for $S_s = 80.40\text{-}94.90$). In terms of sensitivity, the advantage in number of positive cases for the alternate criterion showed a medium effect size (*Cohen's g* = .18). The relative proportions and effect sizes are displayed in Figure 6.



Note: *Cohen's g* effect size (|P-.50|); * = *p <.05 by Sign Test, *** p <.001 by Sign Test.*

Figure 6. Proportion of cases in favor of the alternate criterion by sensitivity, specificity, positive predictive value and negative predictive value.

Also consistent with the prior study, the *specificity* of the alternate criterion (80.1%) was significantly reduced compared to the standard criterion (86.9%). The specificity of the alternate criterion was reduced by 8.5%, and only 4.3% of the cases showed a positive difference favoring the alternate criterion (alternate criterion = 80.1%, standard criterion = 86.9%, positive differences = 6 of 140 favoring standard; Sign Test significant, p< .001, number of positive differences = S_s = 6 of 140 = .0429, standard error of differences = 5.809, 95% CI for $S_{s,}$ = 0.1910-11.809). In terms of specificity, the advantage in number of positive differences for the standard criterion showed a large effect size (*Cohen's g* = .46). These two metrics show that the improvement in sensitivity for the alternate criterion was obtained at a cost of a decrement in specificity.

The *positive predictive values* show a value for the alternate criterion (37.4%) that was slightly lower than the standard criterion (40.2%). Looking at the number of positive cases, there were only 48 of 132 cases or 36% that favored the alternate criterion, showing that the majority of cases favored the standard criterion. Therefore it appears that the advantage is for the standard criterion (alternate = 37.4%; standard = 40.2%; positive differences favoring the alternate = 48 of 132 = .3636, Sign Test significant, p = < .05, S_s = 48, standard error of differences = 5.500, 95% CI for S_s = 42.50-53.50). For positive predictive value, the advantage in number of positive differences for the standard criterion showed a small effect size (*Cohen's g* = .14).

The *negative predictive values* showed the alternate (95.8%) was stronger than the standard criterion (92.9%) for a 3.0% advantage. Also, the proportion of positive differences (62%) favored the alternate criterion (alternate = 95.8%, standard = 92.9%, positive differences 87 of 140 = .6214, Sign Test significant, $p < .001$; S_s = 87, standard

error of differences = 5.220, 95% confidence interval for test statistic 81.7800 to 92.2200). The advantage in number of positive differences for the alternate criterion showed a small effect size (*Cohen's g* = .12).

Determination of Racial/Ethnic Differences in SCL-Results Using Standard and Alternate Criteria.

There were demonstrable differences in SCL-results (Figure 7, Figure 8), using the standard and alternate criteria, however, as sample size varied widely by group, performance indices should be interpreted with caution in racial/ethnic groups of very small sample size.

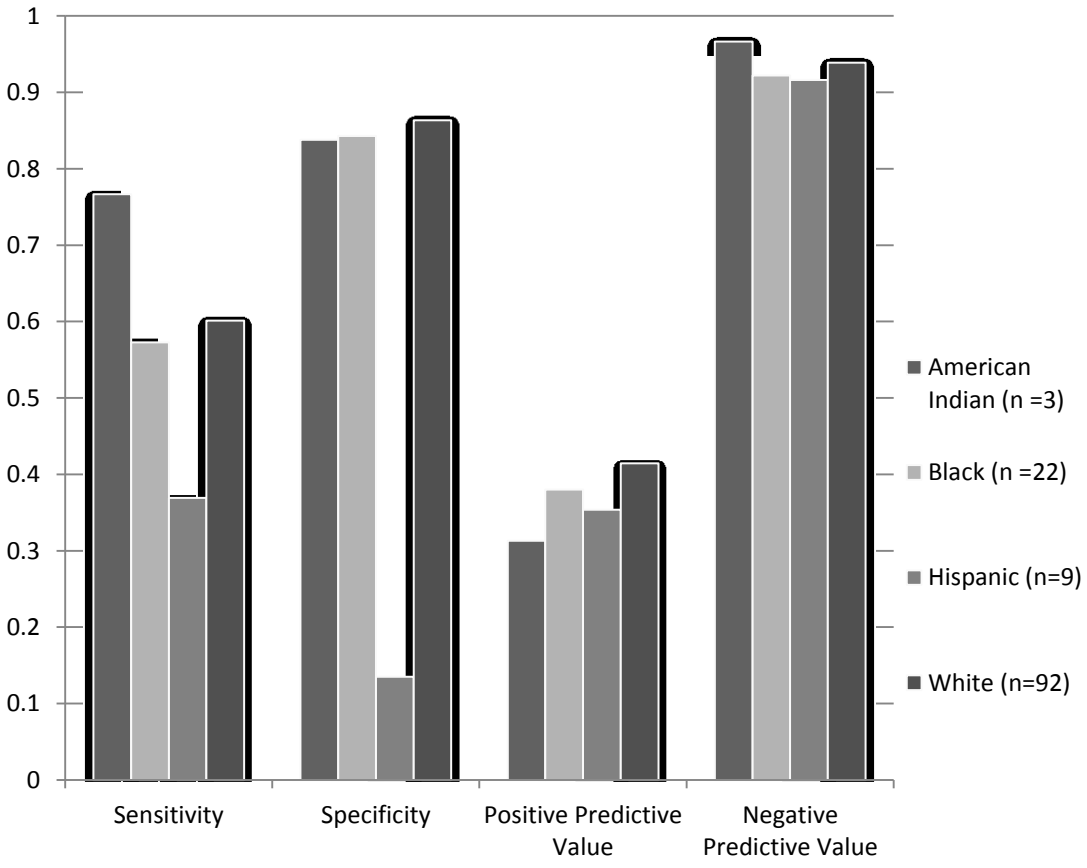


Figure 7. Racial/Ethnic differences in SCL indices using standard criterion.

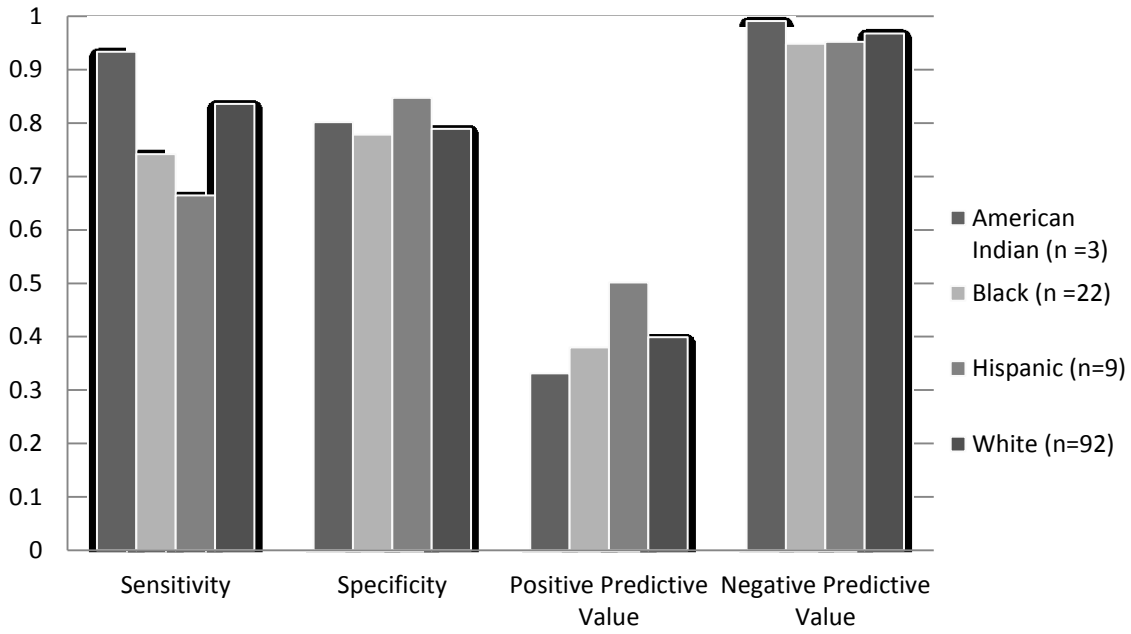


Figure 8. Racial/Ethnic differences in SCL indices using alternate criterion.

To have adequate sample size to compare the alternative versus standard criteria across racial groups the smaller race groups were combined into a single *non-White* group for comparison to the *White* group. A Mann-Whitney U test was performed for each of the performance metrics (*sensitivity, specificity*, *positive predictive value* and *negative predictive value*) to compare the alternative versus standard criteria's ability to find a difference between *White* and *non-White* groups. The Mann-Whitney U test is a non-parametric test for two independent groups *(White* and *non-White*), and it was computed separately for sensitivity, specificity, and positive predictive value for both standard and alternative criteria (for a total of six tests). At alpha of .05, the alternative criterion showed a significant difference in sensitivity between *white* and *non-white groups,* in favor of the *white* group. None of the other Mann-Whitney U test results were significant (Table 3).

Table 3

Racial/Ethnic Group Results and p-values from Mann-Whitney U Test.

Comparison	*White*	*Non-White*	*p value*
Standard Criterion Sensitivity	.6667	.5000	$p = .283$, n. s.
Standard Criterion Specificity	.8807	.8706	$p = .994$, n. s.
Standard Criterion Positive Predictive Value	.4143	.3485	$p = .419$, n. s.
Standard Criterion Negative Predictive Value	.9391	.9241	$p = .604$, n. s.
Alternative Criterion Sensitivity	.8358	.7381	$p = .043$*
Alternative Criterion Specificity	.8068	.8045	$p = .663$, n. s.
Alternative Criterion Positive Predictive Value	.3660	.4083	$p = .674$, n. s
Alternative Criterion Negative Predictive Value	.9676	.9528	$p = .604$, n. s.

* *Significant at the $p < .05$ level.*

Determination of the Relationship of SCL-Derived Hot Flashes Using Standard and Alternate Criteria to Components of the Hot Flash Related Daily Interference Scale.

To compare the two SCL criteria to the components of Hot Flash Related Daily Interference Scale (HFRDIS), the proportion of true positives for each criterion was correlated with the HFRDIS components as shown in Table 4. For both the standard criterion and the alternative criterion the Pearson Product Moment correlations were all small and non-significant ($p>.05$).

Summary of Hypothesis Testing

A summary of all tested hypotheses are provided in Table 5.

Table 4

Correlations of standard and alternate criteria-derived true positive hot flashes to constructs of the Hot Flash Daily Interference Scale.

	Standard Criterion True Positive Hot Flashes		Alternate Criterion True Positive Hot Flashes	
HFRDIS	*r*	*p*	*r*	*p*
Work interference	-.097	.259	-.088	.304
Social activities	-.113	.182	-.036	.669
Leisure interference	-.011	.902	-.001	.994
Sleep interference	.140	.099	.107	.207
Mood interference	-.016	.851	.042	.625
Concentration interference	-.025	.771	.013	.878
Relationships with others	.024	.777	.128	.133
Sexuality interference	.023	.786	.018	.836
Enjoyment of life interference	.025	.770	.109	.201
Overall quality of life interference	.033	.701	.088	.303

Note: All correlations non-significant at alpha of .05.

Table 5

Summary of Results

Hypotheses	*Description*	*Hypothesis supported?*
	Aim 1	
Hypothesis 1.1	The *sensitivity* (TP/TP+FN), of the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show a significantly greater proportion in detecting an objective hot flash when one has been self-reported than the standard criterion (≥ 2 μmho rise in a 30-s period).	yes
Hypothesis 1.2	The *specificity* (TN/FP+TN) of the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show, as was shown in previous study of breast cancer survivors, a significantly lesser proportion in recording the absence of a hot flash when none has been self-reported than the standard criterion (≥ 2 μmho rise in a 30-s period).	yes
Hypothesis 1.3	The *positive predictive value* (TP/TP+FP) of the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show a significantly greater proportion of criterion-indicated hot flashes with corresponding self-report than the standard criterion (≥ 2 μmho rise in a 30-s period).	yes
Hypothesis 1.4	The *negative predictive value* (TN/TN+FN) of the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show a significantly greater proportion of correctly identified negative results than the standard criterion (≥ 2 μmho rise in a 30-s period).	yes

(Continued)

Aim 2

Hypothesis 2.1	The *sensitivity* of the standard SCL criterion for identifying hot flashes (≥ 2.0 μmho rise in a 30-s period) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show sensitivity differs significantly among White and non-White racial/ethnic groups.	no
Hypothesis 2.2	The *specificity* of the standard SCL criterion for identifying hot flashes (≥ 2.0 μmho rise in a 30-s period) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show specificity differs significantly among White and non-White racial/ethnic groups.	no
Hypothesis 2.3	The *positive predictive value* of the standard SCL criterion for identifying hot flashes (≥ 2.0 μmho rise in a 30-s period) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show positive predictive value differs significantly among White and non-White racial/ethnic groups.	no
Hypothesis 2.4	The *negative predictive value* of the standard SCL criterion for identifying hot flashes (≥ 2.0 μmho rise in a 30-s period) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show the *negative predictive value* differs significantly among White and non-White racial/ethnic groups.	no
Hypothesis 2.5	The *sensitivity* of the alternate SCL criterion for identifying hot flashes (≥ 1.2 μmho rise in a 30-s period) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show *sensitivity* differs significantly among White and non-White racial/ethnic groups.	yes
Hypothesis 2.6	The *specificity* of the alternate SCL criterion for identifying hot flashes (≥ 1.2 μmho rise in a 30-s period) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show *specificity* differs significantly among White and non-White racial/ethnic groups.	no

(Continued)

Hypothesis 2.7	The *positive predictive value* of the alternate SCL criterion for identifying hot flashes (≥ 1.2 μmho rise in a 30-s period) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show positive predictive value differs significantly among White and non-White racial/ethnic groups.	no
Hypothesis 2.8	The *negative predictive value* of the alternate SCL criterion for identifying hot flashes (≥ 1.2 μmho rise in a 30-s period) in ambulatory hot flash detection over one day's waking hours in post-menopausal women will show the negative predictive value differs significantly among White and non-White racial/ethnic groups.	no
	Aim 3	
Hypothesis 3.1	Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the standard SCL criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with work as indicated on scores on the *Work* construct of the Hot Flash Related Daily Interference Scale in post-menopausal women.	no
Hypothesis 3.2	Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the standard SCL criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with social activities as indicated on scores on the *Social Activities* construct of the Hot Flash Related Daily Interference Scale in post-menopausal women.	no
Hypothesis 3.3	Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the standard SCL criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with leisure activity as indicated on scores on the *Leisure* construct of the Hot Flash Related Daily Interference Scale in post-menopausal women.	no

(Continued)

Hypothesis 3.4	Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the standard SCL criterion (≥ 2.0 µmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with sleep as indicated on scores on the *Sleep* construct of the Hot Flash Related Daily Interference Scale in post-menopausal women.	no
Hypothesis 3.5	Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the standard SCL criterion (≥ 2.0 µmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related effect on mood as indicated on scores on the *Mood* construct of the Hot Flash Related Daily Interference Scale in post-menopausal women.	no
Hypothesis 3.6	Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the standard SCL criterion (≥ 2.0 µmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with concentration as indicated on scores on the *Concentration* construct of the Hot Flash Related Daily Interference Scale in post-menopausal women.	no
Hypothesis 3.7	Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the standard SCL criterion (≥ 2.0 µmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with relationships as indicated on scores on the *Relationships* construct of the Hot Flash Related Daily Interference Scale in post-menopausal women.	no
Hypothesis 3.8	Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the standard SCL criterion (≥ 2.0 µmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related effect on sexuality as indicated on scores on the *Sexuality* construct of the Hot Flash Related Daily Interference Scale in post-menopausal women.	no

(Continued)

Hypothesis 3.9	Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the standard SCL criterion (≥ 2.0 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with overall quality of life as indicated on scores on the *Overall Quality of Life* construct of the Hot Flash Related Daily Interference Scale in post-menopausal women.	no
Hypothesis 3.10	Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with work as indicated on scores on the *Work* construct of the Hot Flash Related Daily Interference Scale in post-menopausal women.	no
Hypothesis 3.11	Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with social activities as indicated on scores on the *Social Activities* construct of the Hot Flash Related Daily Interference Scale in post-menopausal women.	no
Hypothesis 3.12	Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with leisure activity as indicated on scores on the *Leisure* construct of the Hot Flash Related Daily Interference Scale in post-menopausal women.	no
Hypothesis 3.13	Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with sleep as indicated on scores on the *Sleep* construct of the Hot Flash Related Daily Interference Scale in post-menopausal women.	no

(Continued)

Hypothesis 3.14	Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related effect on mood as indicated on scores on the *Mood* construct of the Hot Flash Related Daily Interference Scale in post-menopausal women.	no
Hypothesis 3.15	Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with concentration as indicated on scores on the *Concentration* construct of the Hot Flash Related Daily Interference Scale in post-menopausal women.	no
Hypothesis 3.16	Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with relationships as indicated on scores on the *Relationships* construct of the Hot Flash Related Daily Interference Scale in post-menopausal women.	no
Hypothesis 3.17	Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related effect on sexuality as indicated on scores on the *Sexuality* construct of the Hot Flash Related Daily Interference Scale in post-menopausal women.	no
Hypothesis 3.18	Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the alternate SCL criterion (≥ 1.2 μmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with enjoyment of life as indicated on scores on the *Enjoyment of Life* construct of the Hot Flash Related Daily Interference Scale in post-menopausal women.	no

(Continued)

Hypothesis 3.19	Monitor-verified hot flashes (meeting criteria with corresponding self-report) using the alternate SCL criterion (≥ 1.2 µmho rise in a 30-s period to indicate a hot flash) will demonstrate a significant positive correlation to the hot flash-related interference with overall quality of life as indicated on scores on the *Overall Quality of Life* construct of the Hot Flash Related Daily Interference Scale in post-menopausal women.	no
Hypothesis 3.20	The correlations between the monitor-verified hot flashes using the alternate SCL criterion (≥ 1.2 µmho rise in a 30-s period to indicate a hot flash) and the standard criterion (≥ 2.0 µmho rise in a 30-s period to indicate a hot flash) will differ significantly in their relationship to hot flash-related interference with work as indicated on scores on the *Work* construct of the Hot Flash Related Daily Interference Scale.	no
Hypothesis 3.21	The correlations between the monitor-verified hot flashes using the alternate SCL criterion (≥ 1.2 µmho rise in a 30-s period to indicate a hot flash) and the standard criterion (≥ 2.0 µmho rise in a 30-s period to indicate a hot flash) will differ significantly in their relationship to hot flash-related interference with social activities as indicated on scores on the *Social Activities* construct of the Hot Flash Related Daily Interference Scale.	no*
Hypothesis 3.22	The correlations between the monitor-verified hot flashes using the alternate SCL criterion (≥ 1.2 µmho rise in a 30-s period to indicate a hot flash) and the standard criterion (≥ 2.0 µmho rise in a 30-s period to indicate a hot flash) will differ significantly in their relationship to hot flash-related interference with leisure activities as indicated on scores on the *Leisure* construct of the Hot Flash Related Daily Interference Scale.	no*
Hypothesis 3.23	The correlations between the monitor-verified hot flashes using the alternate SCL criterion (≥ 1.2 µmho rise in a 30-s period to indicate a hot flash) and the standard criterion (≥ 2.0 µmho rise in a 30-s period to indicate a hot flash) will differ significantly in their relationship to hot flash-related interference with sleep as indicated on scores on the *Sleep* construct of the Hot Flash Related Daily Interference Scale.	no*

(Continued)

Hypothesis 3.24	The correlations between the monitor-verified hot flashes using the alternate SCL criterion (≥ 1.2 µmho rise in a 30-s period to indicate a hot flash) and the standard criterion (≥ 2.0 µmho rise in a 30-s period to indicate a hot flash) will differ significantly in their relationship to hot flash-related impact on mood as indicated on scores on the *Mood* construct of the Hot Flash Related Daily Interference Scale.	no*
Hypothesis 3.25	The correlations between the monitor-verified hot flashes using the alternate SCL criterion (≥ 1.2 µmho rise in a 30-s period to indicate a hot flash) and the standard criterion (≥ 2.0 µmho rise in a 30-s period to indicate a hot flash) will differ significantly in their relationship to hot flash-related interference with concentration as indicated on scores on the *Concentration* construct of the Hot Flash Related Daily Interference Scale.	no*
Hypothesis 3.26	The correlations between the monitor-verified hot flashes using the alternate SCL criterion (≥ 1.2 µmho rise in a 30-s period to indicate a hot flash) and the standard criterion (≥ 2.0 µmho rise in a 30-s period to indicate a hot flash) will differ significantly in their relationship to hot flash-related interference with relationships as indicated on scores on the *Relationships* construct of the Hot Flash Related Daily Interference Scale.	no*
Hypothesis 3.27	The correlations between the monitor-verified hot flashes using the alternate SCL criterion (≥ 1.2 µmho rise in a 30-s period to indicate a hot flash) and the standard criterion (≥ 2.0 µmho rise in a 30-s period to indicate a hot flash) will differ significantly in their relationship to hot flash-related interference with sexuality as indicated on scores on the *Sexuality* construct of the Hot Flash Related Daily Interference Scale.	no*
Hypothesis 3.28	The correlations between the monitor-verified hot flashes using the alternate SCL criterion (≥ 1.2 µmho rise in a 30-s period to indicate a hot flash) and the standard criterion (≥ 2.0 µmho rise in a 30-s period to indicate a hot flash) will differ significantly in their relationship to hot flash-related impact on the enjoyment of life as indicated on scores on the *Enjoyment of Life* construct of the Hot Flash Related Daily Interference Scale.	no*

(Continued)

Hypothesis 3.29	The correlations between the monitor-verified hot flashes using the alternate SCL criterion (≥ 1.2 µmho rise in a 30-s period to indicate a hot flash) and the standard criterion (≥ 2.0 µmho rise in a 30-s period to indicate a hot flash) will differ significantly in their relationship to hot flash-related impact on quality of life as indicated on scores on the *Overall Quality of Life* construct of the Hot Flash Related Daily Interference Scale	no*
Hypothesis 3.30	The correlations between the monitor-verified hot flashes using the alternate SCL criterion (≥ 1.2 µmho rise in a 30-s period to indicate a hot flash) and the standard criterion (≥ 2.0 µmho rise in a 30-s period to indicate a hot flash) will differ significantly in their relationship to hot flash-related impact on quality of life as indicated on scores on the *Overall Quality of Life* construct of the Hot Flash Related Daily Interference Scale.	no*

Note. Correlations were not significant for either variable; no tests for statistically significant differences between correlations were performed.

CHAPTER FOUR

Discussion and Conclusions

Hot flashes are a highly prevalent and serious symptom for millions, globally. The study of hot flashes is vibrant, multidisciplinary and ongoing, but nascent. Effective measures of hot flashes are critical, both in laboratory study and also for use in ambulatory clinical trials. Skin conductance monitors have been found to be the most sensitive measure of hot flashes in laboratory study, but the results of ambulatory trials have shown the standard criteria for detecting hot flashes to be inadequately sensitive. The results of a study of breast cancer survivors suggest that a more liberal SCL criterion provides a substantial improvement in sensitivity, at only a small cost in specificity. In clinical utility, a more sensitive instrument is optimal, with over-reporting, if consistent, being less confounding than a criterion that fails to identify existent hot flashes.

This study examined whether an alternate criterion, identified as superior in the detection of hot flashes in breast cancer patients, would also be superior in healthy, post-menopausal women. The results from 140 post-menopausal women showed that the alternative criterion was significantly (35.8%) more sensitive in detecting hot flashes than the standard criterion. Consistent with the findings in the Savard study, this improvement in sensitivity came at a cost to specificity (M.-H. Savard et al., 2013). This cost, though statistically significant, was only a reduction in specificity of 8.5% from the standard criterion. Positive predictive value for the alternate criterion was significantly (7.5%), inferior to the standard criterion, though neither criterion was particularly strong on this

metric (37.4% vs. 40.2%) suggesting that both criteria produce a high number of false positive hot flashes. Negative predictive values however, were excellent for both criterion (92.9% standard, 95.8% alternate), the alternate criterion showing a statistically significant ($p < .001$), but perhaps only marginally, (3.1%) meaningful advantage. Taken together, these findings suggest that the alternate criterion, given a strong gain in sensitivity with only minor costs to specificity, would be recommended over the standard criterion in ambulatory studies of SCL hot flash detection in healthy, post-menopausal women.

This study also examined racial/ethnic differences in the standard and alternate criteria, though small sample sizes only allowed for the statistical examination of white versus non-whites. A significant difference was found using the alternate criterion in this bivariate examination with white participants showing significantly greater sensitivity than non-whites (83.58% vs. 73.81%, $p < .001$). Though differences were observed in other metrics (see Figure 3 & Figure 4), they were not found significant using the Mann-Whitney U test. The reason for the difference in sensitivity between White and non-White racial/ethnic groups is unknown; however, there are a few possibilities.

The difference in sensitivity between White and non-White racial/ethnic groups could be the result of differing skin properties. There is some support for this hypothesis as previous study has shown eccrine sweat gland activity, as well as resting skin conductance, differs among White and Black participants (Johnson & Landon, 1965; Wesley & Maibach, 2003). Another possibility lies in cultural differences in reporting of hot flashes. Previous research has demonstrated differences in the subjective report of

hot flashes between White and Black participants (Grisso, Freeman, Maurin, Garcia-Espana, & Berlin, 1999).

It is also possible that the expression of hot flashes themselves differs between racial/ethnic groups. Whether subjective report of hot flashes, artifacts of skin conductance, or differing physiological expression of hot flashes differs in racial/ethnic groups remains an empirical question. It may be advantageous to investigate the optimal SCL criterion for differing racial/ethnic groups. Future study should investigate this issue in order to best measure hot flashes in a diverse world.

The third area that this study examined was the determination if there was a relationship between criteria-derived hot flashes (standard and alternate), and indices of hot flash related daily interference. It was hypothesized that the frequency of physiologically determined hot flashes should correspond to impact on daily interference and quality of life (Carpenter, 2001). The results failed to show any significant relationship between SCL-derived hot flashes and *Work, Social Activities, Leisure Activities, Sleep, Mood, Concentration, Relationships, Sexuality, Enjoyment of Life* and *Overall Quality of Life* scores on the HFRDIS. Though this finding is seemingly counter-intuitive, it is not unique. Previous research has shown that, rather than the frequency of hot flashes, it is the problem-rating of hot flashes that correspond to impact on health-related quality of life (Ayers & Hunter, 2013). Further, as SCL-derived hot flashes do not correspond with hot flash severity (Carpenter, Azzouz, Monahan, Storniolo, & Ridner, 2005), the physiological measurement of hot flashes via skin conductance may not be a suitable measure to determine hot flash related impact. Investigation into

alternate physiological measures that adequately captures not only frequency but also some indication of the severity of hot flashes is warranted.

In a study of 150 post-menopausal women, participants with varying frequency of hot flashes underwent 24-hour electrocardiographic recording (Hoikkala et al., 2010). Though the researchers did not find cardiovascular risk factors to be associated with the frequency and severity of hot flashes, there were signs of altered autonomic control of heart rate, which the authors speculate may be involved in the regulatory mechanisms of hot flashes. Specifically, the very-low-frequency spectral component of heart rate variability increased by 72% during a hot flash compared with a control period and was accompanied by a small but significant 3% increase in heart rate. It is possible that sampled epochs of very-low-frequency spectral component of heart rate variability could be used to detect the frequency and severity of hot flashes, though this remains an empirical question.

Further, it is possible that using a second physiological measure, such as ECG for heart rate variability, could be used to reduce false positive findings using current skin conductance monitors. Previous study has shown that heart rate increases markedly during a hot flash (Freedman, 1989), however, to-date, no study has attempted an ambulatory study of hot flashes involving SCL and heart-rate variability.

There were a number of limitations to the current study. As the study was a retrospective analysis of previously collected SCL-monitor data, the addition of additional measures (such as alternate quality of life measures) was not possible. The study was also limited in that it was a largely white, college-educated sample. Though the data used in this study came from a more diverse sample than the majority of the

published literature (which comes from an overwhelmingly white population), there was insufficient sample size in all racial/ethnic groups for statistical analyses of differences. Only by making a bivariate, white/non-white comparison, could sufficient power for statistical analyses be achieved. This finding, although underpowered for specific analyses between all sampled racial/ethnic groups, is rare in the literature, with the sample size used in this study one of the largest datasets of SCL-monitor data published to-date.

This study was the first study to investigate an alternate SCL criterion ($\geq$ 1.2 μmho rise in a 30-s period) versus the standard SCL criterion ($\geq$ 2 μmho rise in a 30-s period) to detect, in ambulatory measurement, hot flashes in healthy, post-menopausal women. This study found that the alternate criterion has a distinct advantage in terms of sensitivity, with only minor deficit to specificity, suggesting that the alternate criterion is advised for future ambulatory study of hot flashes in post-menopausal women. This finding is important as it is an easily diseminable addition to ongoing hot flash research using the current gold-standard ambulatory monitors and software. No new or additional equipment or software will be required to implement this criterion. This study also investigated ethnic differences, finding that the sensitivity of both criteria for detecting hot flashes was significantly greater for white participants than non-white participants, suggesting the need for a study of racial/ethnic differences in SCL-monitor hot flash detection to determine optimal criteria for detecting hot flashes in diverse populations. Finally, this study examined the SCL-data determining that SCL-criterion derived hot flashes showed no relationship to hot flash-related daily interference.

Future Directions

The next logical step in the study of SCL hot flash detection is a study that is designed to achieve four aims: 1) Replicate the findings of this study 2) Examine SCL-criteria differences between White, Black and Hispanic participants, 3) Perform receiver operating characteristics (ROC) analyses of White, Black and Hispanic participants collectively and individually in order to ascertain if a racial/ethnic-specific SCL-criterion is optimal, and 4) Determine if the data collected from the addition of ambulatory heart-rate variability to SCL monitors can be used to improve the specificity of SCL-derived hot flashes.

Based on the lessons learned from this study, the following study will be designed to specifically accrue a minimum of 30 Black, 30 White and 30 Hispanic participants. In order to determine whether the alternate criterion performs across the menopausal spectrum, inclusion criteria will be altered from the current study to include symptomatic women of all ages and cancer status. Thus, comparative analyses will be possible to determine differences in SCL-criteria efficacy across menopausal stage and cancer-status. Cancer-survivors are to be included as no study to-date has provided convincing evidence that the physiological expression and the corresponding physiological detection of hot flashes differs as a result of cancer-status. This study will involve participants completing two sessions, one a 4-hour laboratory session with a 30-minute resting baseline, a stress condition (L. C. Swartzman et al., 1990) and a heating paradigm (Freedman, 1989) to illicit hot flashes. The second session will be a 24-hour, outside of the laboratory, ambulatory recording session using a SCL ambulatory monitor, an ambulatory ECG, an event marker to indicate onset of hot flashes and a waking-hour

hourly activity log. The results of this subsequent study will provide critical data on the efficacy of alternate SCL-criterion, whether individual criterion are needed for White, Black or Hispanic racial/ethnic groups, and whether an alternate criterion is equally effective in participants at differing stages of menopause and cancer status. The data derived from the proposed study will provide critically important information on the measurement of hot flashes, and thus the future study of the etiology and treatment of hot flashes.

WORKS CITED

Abdi, H. (2007). Signal detection theory (SDT). In N. J. Salkind (Ed.), *Encyclopedia of measurement and statistics* (pp. 886-889). Thousand Oaks, CA: Sage, Inc.

Anderson, D., Yoshizawa, T., Gollschewski, S., Atogami, F., & Courtney, M. (2004). Menopause in Australia and Japan: effects of country of residence on menopausal status and menopausal symptoms. *Climacteric, 7*(2), 165-174.

Avis, N. E., Stellato, R., Crawford, S., Bromberger, J., Ganz, P., Cain, V., & Kagawa-Singer, M. (2001). Is there a menopausal syndrome? Menopausal status and symptoms across racial/ethnic groups. *Social Science and Medicine, 52*(3), 345-356.

Ayers, B., & Hunter, M. (2013). Health-related quality of life of women with menopausal hot flushes and night sweats. *Climacteric, 16*(2), 235-239.

Bartlett, M. S. (1947). The use of transformations. *Biometrics, 3*(1), 39-52.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289-300.

BIOPAC Systems, I. EL-507 Disposable EDA Electrode (Vol. Accessed at http://www.biopac.com/disposable-electrodermal-electrode-100. Accessed February 12, 2012.). Goletta, CA

Carey, G. (2012). *Quantitative Methods in Neuroscience.* (Vol. E-textbook, available at http://psych.colorado.edu/~carey/qmin/qmin.php): University of Colorado.

Carpenter, J. S. (2001). The Hot Flash Related Daily Interference Scale: a tool for assessing the impact of hot flashes on quality of life following breast cancer. *Journal of Pain Symptom Management, 22*(6), 979-989.

Carpenter, J. S. (2005). State of the science: hot flashes and cancer. Part 1: definition, scope, impact, physiology, and measurement. *Oncology nursing forum, 32*(5), 959-968. doi: 10.1188/05.ONF.959-968

Carpenter, J. S., Andrykowski, M. A., Freedman, R. R., & Munn, R. (1999). Feasibility and psychometrics of an ambulatory hot flash monitoring device. *Menopause, 6*(3), 209-215.

Carpenter, J. S., Azzouz, F., Monahan, P. O., Storniolo, A. M., & Ridner, S. H. (2005). Is sternal skin conductance monitoring a valid measure of hot flash intensity or distress? *Menopause, 12*(5), 512-519.

Carpenter, J. S., Gilchrist, J. M., Chen, K., Gautam, S., & Freedman, R. R. (2004). Hot flashes, core body temperature, and metabolic parameters in breast cancer survivors. *Menopause, 11*(4), 375-381.

Carpenter, J. S., Newton, K. M., Sternfeld, B., Joffe, H., Reed, S. D., Ensrud, K. E., & Milata, J. L. (2012). Laboratory and ambulatory evaluation of vasomotor symptom monitors from the Menopause Strategies Finding Lasting Answers for Symptoms and Health network. *Menopause, 19*(6), 664-671. doi: 10.1097/gme.0b013e31823dbbe3

Carpenter, J. S., & Rand, K. L. (2008). Modeling the hot flash experience in breast cancer survivors. *Menopause, 15*(3), 469-475. doi: 10.1097/gme.0b013e3181591db7

Chedraui, P., Pérez-López, F. R., Sánchez, H., Aguirre, W., Martínez, N., Miranda, O., . . . Zambrano, B. (2012). Assessment of sexual function of mid-aged Ecuadorian women with the 6-item Female Sexual Function Index. *Maturitas, 71*(4), 407-412. doi: http://dx.doi.org/10.1016/j.maturitas.2012.01.013

Cohen, J. (1988). *Statistical power analysis for the behavioral sciencies*: Routledge.

Daut, R., Cleeland, C., & Flannery, R. (1993). Development of the Wisconsin Brief Pain Inventory to assess pain in cancer and other diseases. *Pain, 17*, 197-210. doi: 10/1983.

de Bakker, I. P., & Everaerd, W. (1996). Measurement of menopausal hot flushes: validation and cross-validation. *Maturitas, 25*(2), 87-98.

Elkins, G. R., Fisher, W. I., Johnson, A. K., Carpenter, J. S., & Keith, T. Z. (2012). Clinical hypnosis in the treatment of postmenopausal hot flashes: a randomized controlled trial. *Menopause*. doi: 10.1097/gme.0b013e31826ce3ed

Erlik, Y., Tataryn, I. V., Meldrum, D. R., Lomax, P., Bajorek, J. G., & Judd, H. L. (1981). Association of waking episodes with menopausal hot flushes. *JAMA, 245*(17), 1741-1744.

Freedman, R. R. (1989). Laboratory and ambulatory monitoring of menopausal hot flashes. *Psychophysiology, 26*(5), 573-579.

Freedman, R. R. (1998). Biochemical, metabolic, and vascular mechanisms in menopausal hot flashes. *Fertility and sterility, 70*(2), 332-337.

Freedman, R. R. (2000). Menopausal hot flashes. In R. A. Lobo, J. Kelsey & R. Marcus (Eds.), *Menopause: Biology and Pathobiology* (pp. 215-227). San Diego: Academic Press.

Freedman, R. R., & Wasson, S. (2007). Miniature hygrometric hot flash recorder. *Fertility and sterility, 88*(2), 494-496. doi: http://dx.doi.org/10.1016/j.fertnstert.2006.11.184

Freeman, E. W., & Sherif, K. (2007). Prevalence of hot flushes and night sweats around the world: a systematic review. *Climacteric, 10*(3), 197-214. doi: 10.1080/13697130601181486

Frödin, T., Ålund, G., & Varenhorst, E. (1985). Measurement of skin blood-flow and water evaporation as a means of objectively assessing hot flushes after orchidectomy in patients with prostatic cancer. *The Prostate, 7*(2), 203-208. doi: 10.1002/pros.2990070210

Gold, E. B., Sternfeld, B., Kelsey, J. L., Brown, C., Mouton, C., Reame, N., . . . Stellato, R. (2000). Relation of demographic and lifestyle factors to symptoms in a multi-racial/ethnic population of women 40-55 years of age. *American Journal of Epidemiology, 152*(5), 463-473.

Goldberg, R. M., Loprinzi, C. L., O'Fallon, J. R., Veeder, M. H., Miser, A. W., Mailliard, J., . . . Burnham, N. (1994). Transdermal clonidine for ameliorating tamoxifen-induced hot flashes. *Journal of Clinical Oncology, 12*(1), 155-158.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1): Wiley New York.

Grisso, J. A., Freeman, E. W., Maurin, E., Garcia-Espana, B., & Berlin, J. A. (1999). Racial differences in menopause information and the experience of hot flashes. *Journal of general internal medicine, 14*(2), 98-103.

Hanisch, L. J., Palmer, S., Marcus, S., Hantsoo, L., Vaughn, D., & Coyne, J. (2009). Comparison of objective and patient-reported hot flash measures in men with prostate cancer. *Journal of Supportive Oncology, 7*(4), 1-5.

Hanisch, L. J., Palmer, S. C., Donahue, A., & Coyne, J. C. (2007). Validation of sternal skin conductance for detection of hot flashes in prostate cancer survivors. *Psychophysiology, 44*(2), 189-193.

Hann, D., Jacobsen, P., Azzarello, L., Martin, S., Curran, S., Fields, K., . . . Lyman, G. (1998). Measurement of fatigue in cancer patients: development and validation of the Fatigue Symptom Inventory. *Quality of Life Research, 7*(4), 301-310.

Hoikkala, H., Haapalahti, P., Viitasalo, M., Väänänen, H., Sovijärvi, A. R. A., Ylikorkala, O., & Mikkola, T. S. (2010). Association between vasomotor hot flashes and heart rate variability in recently postmenopausal women. *Menopause, 17*(2), 315-320. doi: 10.1097/gme.0b013e3181c2bb6d

Hunter, M. S., & Liao, K. (1995). A psychological analysis of menopausal hot flushes. *British Journal of Clinical Psychology, 34*(4), 589-599.

Isidori, A., Pozza, C., Esposito, K., Giugliano, D., Morano, S., Vignozzi, L., . . . Jannini, E. (2010). Development and validation of a 6-item version of the female sexual function index (FSFI) as a diagnostic tool for female sexual dysfunction. *The journal of sexual medicine, 7*(3), 1139-1146.

Johnson, L. C., & Landon, M. M. (1965). Eccrine sweat gland activity and racial differences in resting skin conductance. *Psychophysiology, 1*(4), 322-329. doi: 10.1111/j.1469-8986.1965.tb03264.x

Kronenberg, F. (1990). Hot flashes: epidemiology and physiology. *Annals of the New York Academy of Science, 592*, 52-86; discussion 123-133.

Kronenberg, F. (1994). Hot flashes: phenomenology, quality of life, and search for treatment options. *Experimental Gerontology, 29*(3-4), 319-336.

Kronenberg, F., Cote, L. J., Linkie, D. M., Dyrenfurth, I., & Downey, J. A. (1984). Menopausal hot flashes: thermoregulatory, cardiovascular, and circulating catecholamine and LH changes. *Maturitas, 6*(1), 31-43.

Lambrinoudaki, I., Augoulea, A., Armeni, E., Rizos, D., Alexandrou, A., Creatsa, M., . . . Stamatelopoulos, K. (2012). Menopausal symptoms are associated with subclinical atherosclerosis in healthy recently postmenopausal women. *Climacteric: The Journal Of The International Menopause Society, 15*(4), 350-357. doi: 10.3109/13697137.2011.618564

Lock, M. (1986). Ambiquities of aging: Japanese experience and perceptions of menopause. *Culture, Medicine and Psychiatry, 10*(1), 23-46. doi: 10.1007/BF00053261

Loprinzi, C. L., Kugler, J. W., Sloan, J. A., Mailliard, J. A., LaVasseur, B. I., Barton, D. L., . . . Christensen, B. J. (2000). Venlafaxine in management of hot flashes in survivors of breast cancer: a randomised controlled trial. *Lancet, 356*(9247), 2059-2063.

Mann, E., & Hunter, M. S. (2011). Concordance between self-reported and sternal skin conductance measures of hot flushes in symptomatic perimenopausal and postmenopausal women: a systematic review. *Menopause, 18*(6), 709-722. doi: 10.1097/gme.0b013e318204a1fb

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics, 18*(1), 50-60.

McCoy, N., Cutler, W., & Davidson, J. (1985). Relationships among sexual behavior, hot flashes, and hormone levels in perimenopausal women. *Archives of Sexual Behavior, 14*(5), 385-394. doi: 10.1007/BF01542000

Melby, M. K., Lock, M., & Kaufert, P. (2005). Culture and symptom reporting at menopause. *Human Reproduction Update, 11*(5), 495-512. doi: 10.1093/humupd/dmi018

Negata, C., Shimizu, C., Takami, R., Hayashi, M., Takeda, N., & Yasuda, K. (1999). Hot flushes and other menopausal symptoms in relation to soy product intake in Japanese women. *Climacteric, 2*(1), 6-12. doi: doi:10.3109/13697139909025557

Otte, J. L., Flockhart, D., Storniolo, A. M., Schneider, B., Azzouz, F., Lemler, S., . . . Stearns, V. (2009). Comparison of subjective and objective hot flash measures over time among breast cancer survivors initiating aromatase inhibitor therapy. *Menopause, 16*(4), 653. doi: 10.1097/gme.0b013e3181a5d0d6

Pandya, K. J., Raubertas, R. F., Flynn, P. J., Hynes, H. E., Rosenbluth, R. J., Kirshner, J. J., . . . Morrow, G. R. (2000). Oral clonidine in postmenopausal patients with breast cancer experiencing tamoxifen-induced hot flashes: A University of Rochester Cancer Center Community Clinical Oncology Program study. *Annals of Internal Medicine, 132*(10), 788-793.

Richardson, A. (1994). The health diary: an examination of its use as a data collection method. *Journal of Advanced Nursing, 19*(4), 782-791.

Rickles, W., & Day, J. (1968). Electrodermal activity in non-palmer skin sites. *Psychophysiology, 4*(4), 421-435.

Rodgers, J., & Nicewander, W. (1988). Thirteen ways to look at the correlation coefficient. *The American Statistician, 42*(1), 59-66.

Savard, J., Davidson, J. R., Ivers, H., Quesnel, C., Rioux, D., Dupéré, V., . . . Morin, C. M. (2004). The association between nocturnal hot flashes and sleep in breast cancer survivors. *Journal of Pain and Symptom Management, 27*(6), 513-522. doi: http://dx.doi.org/10.1016/j.jpainsymman.2003.10.013

Savard, M.-H., Savard, J., & Ivers, H. (2013). Validation of a sternal skin conductance criterion for measuring hot flashes in breast cancer patients. *Menopause: The Journal of The North American Menopause Society, 21*(2), Epub ahead of print.

Scoutellas, V., O'Neill, T. W., Lunt, M., Reeve, J., & Silman, A. J. (1999). Does the presence of postmenopausal symptoms influence susceptibility to vertebral deformity? *Maturitas, 32*(3), 179-187. doi: http://dx.doi.org/10.1016/S0378-5122(99)00025-0

Siegel, S. (1956). *Nonparametric statistics for the behavioural sciences*: McGraw-Hill.

Sievert, L. L. (2013). Subjective and objective measures of hot flashes. *American Journal of Human Biology, 25*(5), 573-580. doi: 10.1002/ajhb.22415

Sievert, L. L., Begum, K., Sharmeen, T., Chowdhury, O., Muttukrishna, S., & Bentley, G. (2008). Patterns of occurrence and concordance between subjective and objective hot flashes among Muslim and Hindu women in Sylhet, Bangladesh. *American Journal of Human Biology, 20*(5), 598-604.

Sloan, J. A., Loprinzi, C. L., Novotny, P. J., Barton, D. L., Lavasseur, B. I., & Windschitl, H. (2001). Methodologic lessons learned from hot flash studies. *Journal of Clin Oncology, 19*(23), 4280-4290.

Stein, K. D., Jacobsen, P. B., Hann, D. M., Greenberg, H., & Lyman, G. (2000). Impact of hot flashes on quality of life among postmenopausal women being treated for breast cancer. *Journal of Pain and Symptom Management, 19*(6), 436-445.

Swartzman, L. C., Edelberg, R., & Kemmann, E. (1990). Impact of stress on objectively recorded menopausal hot flushes and on flush report bias. *Health Psychology, 9*(5), 529.

Swartzman, L. C., Edelberg, R., & Kemmann, E. (1990). The menopausal hot flush: symptom reports and concomitant physiological changes. *Journal of Behavioral Medicine, 13*(1), 15-30.

Tataryn, I., Lomax, P., Meldrum, D., Bajorek, J., Chesarek, W., & Judd, H. (1981). Objective techniques for the assessment of postmenopausal hot flashes. *Obstetrics & Gynecology, 57*(3), 340-344.

Thurston, R. C., Blumenthal, J. A., Babyak, M. A., & Sherwood, A. (2005). Emotional antecedents of hot flashes during daily life. *Psychosomatic Medicine, 67*(1), 137-146.

Thurston, R. C., Hernandez, J., Del Rio, J. M., & De La Torre, F. (2011). Support Vector Machines to improve physiologic hot flash measures: application to the ambulatory setting. *Psychophysiology, 48*(7), 1015-1021. doi: 10.1111/j.1469-8986.2010.01155.x

Thurston, R. C., Sutton-Tyrrell, K., Everson-Rose, S. A., Hess, R., & Matthews, K. A. (2008). Hot flashes and subclinical cardiovascular disease: findings from the Study of Women's Health Across the Nation Heart Study. *Circulation, 118*(12), 1234-1240. doi: 10.1161/CIRCULATIONAHA.108.776823

UFI. (2006). FlashTrax (Version 1.2). Morro Bay, California.

UFI. (2012). Biolog Model 3991 SCL. Retrieved June 12, 2012, from http://www.UFIservingscience.com/biologs.html

Wesley, N. O., & Maibach, H. I. (2003). Racial (ethnic) differences in skin properties: the objective data. *American Journal Of Clinical Dermatology, 4*(12), 843-860.

Woodward, S., Greville, H., & Freedman, R. R. (1995). Ventilatory response during menopausal hot flashes. *Menopause, 2*(2), 81-88.

Zeserson, J. M. (2001). How Japanese women talk about hot flushes: implications for menopause research. *Medical anthropology quarterly, 15*(2), 189-205.

Zubin, J. (1935). Note on a transformation function for proportions and percentages. *Journal of Applied Psychology, 19*(2), 213.

Zweig, M. H., & Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical chemistry, 39*(4), 561-577.