Menopause: The Journal of The North American Menopause Society
Vol. 19, No. 6, pp. 664-671
DOI: 10.1097/gme.0b013e31823dbbe3

Laboratory and ambulatory evaluation of vasomotor symptom monitors from the Menopause Strategies Finding Lasting Answers for Symptoms and Health network

Janet S. Carpenter, PhD, RN, FAAN,[1] Katherine M. Newton, PhD,[2] Barbara Sternfeld, PhD,[3] Hadine Joffe, MD, MSc,[4] Susan D. Reed, MD, MPH,[5] Kristine E. Ensrud, MD, MPH, FACP,[6] and Jennifer L. Milata, MSN, RN[7]

Abstract

Objective: The aim of this study was to evaluate monitors for assessing vasomotor symptoms (VMS) in laboratory and ambulatory settings before use in the Menopause Strategies Finding Lasting Answers for Symptoms and Health network clinical trials testing VMS therapies.

Methods: This was a three-phase study. Phase 1 included laboratory testing of the Freedman and prototype Bahr Monitor, phase 2 included laboratory testing of the commercial Bahr Monitor and Biolog, and phase 3 included ambulatory testing of the commercial Bahr Monitor and Biolog. All phases enrolled midlife women with VMS, midlife women without VMS, and young women without VMS. The participants self-reported VMS by pressing event marker buttons. Questionnaires assessed demographics (all phases) and monitor acceptability (phases 2 and 3).

Results: Phase I testing was stopped because of sensitivity of the Freedman device to ambient humidity changes and lack of analytic software for the prototype Bahr Monitor. In phases 2 and 3, agreement between event-marked and commercial Bahr Monitor or Biolog-recorded VMS was higher in the laboratory than in the ambulatory setting; however, agreement between monitors was poor in two of three laboratory groups (midlife no VMS and young no VMS) and in all ambulatory groups. During ambulatory monitoring, the mean number of Bahr Monitor VMS was 16.33 in midlife women with VMS, 9.61 in midlife women without VMS, and 14.63 in young women without VMS (software version, March 2011). The Bahr Monitor was more acceptable than the larger Biolog, but feedback reflected annoyance at having to wear a device that itched and was visible under clothing.

Conclusions: The Bahr Monitor and Biolog seem suitable for use in controlled laboratory conditions during short periods of time. However, the current versions of these monitors may not be suitable for ambulatory clinical trials at this time.

Key Words: Menopause – Vasomotor symptoms – Sternal skin conductance – Hygrometer – Hot flashes – Symptom assessment.

Vasomotor symptoms (VMS; eg, hot flashes, night sweats) are the cardinal symptom of menopause. Careful selection and evaluation of VMS measures are integral in evaluating the efficacy/effectiveness of potential interventions for VMS. Objective measurement of VMS frequency using monitoring devices has been recommended as an adjunct to subjective measurement of frequency, severity, bother, and/or duration. The advantages of objective VMS monitoring are that results are thought to be unbiased by placebo effects,[1] sleep-wake cycles,[2,3] or reporting difficulties.[2,4] Disadvantages are the inability to objectively capture severity, bother, or duration,[5] the participant burden associated with

Received July 14, 2011; revised and accepted October 18, 2011.

From the [1]Department of Adult Health, School of Nursing, Indiana University, Indianapolis, IN; [2]Group Health Research Institute, Seattle, WA; [3]Division of Research, Kaiser Permanente, Oakland, CA; [4]Massachusetts General Hospital, Boston, MA; [5]University of Washington School of Medicine, Seattle, WA; [6]VA Medical Center/University of Minnesota, Minneapolis, MN; and [7]Center for Nursing Research and Scholarship, School of Nursing, Indiana University, Indianapolis, IN.

Funding/support: This study was supported by a cooperative agreement issued by the National Institute on Aging (NIA), in collaboration with the *Eunice Kennedy Shriver* National Institute of Child Health and Development (NICHD), the National Center for Complementary and Alternative Medicine (NCCAM), and the Office of Research and Women's Health (ORWH) and NIA grants U01AG032659, U01AG032669, U01AG032682, U01AG032699, and U01AG032700. In Indiana, the project was supported by the Indiana Clinical and Translational Sciences Institute and funded, in part, by grant UL1 RR025761 from the National Institutes of Health, National Center for Research Resources, Clinical and Translational Sciences Award.

Financial disclosure/conflicts of interest: Hadine Joffe provides research support (to the Center for Women's Mental Health at Massachusetts General Hospital where she participates in protocols as a co-investigator) to Bayer HealthCare Pharmaceuticals, Forest Laboratories, and GlaxoSmithKline. She also provides advisory/consulting services to Sanofi-Aventis/Sunovion.

Address correspondence to: Janet S. Carpenter, PhD, RN, FAAN, Professor and Sally Reahard Chair, Indiana University School of Nursing, 1111 Middle Drive, Indianapolis, IN, USA 46202. E-mail: carpentj@iupui.edu

wearing such devices, and the resources required for data analysis. Although some investigators believe that subjective ratings are the only measure that should be used to assess VMS because women seek treatment based on their subjective experiences, others believe that objective measures can give important information about the physiological effects of an intervention. Therefore, objective monitoring is generally viewed as an adjunct to, not as a replacement for, subjective measurement.[6]

Currently available VMS monitors are limited. Sternal skin conductance is the most widely accepted objective measure of VMS. Sternal skin conductance rises rapidly and transiently during VMS events, even in response to very small amounts of sweating. One skin conductance monitor (Biolog; UFI, Morro Bay, CA) has been used in menopause studies in the past 2 decades.[1,2,4,7-9] However, widespread use has been limited by its cost, size, and weight, relatively short monitoring capacity of 24 hours, and the need for customized conductive paste and electrodes. Further evaluation of the Biolog's performance is warranted because the customized electrodes were recently re-engineered. Two additional miniaturized monitors have been developed to overcome the limitations of the Biolog. One is a sternal skin conductance monitor (Bahr Monitor; Simplex Scientific, Middleton, WI) that attaches directly to the skin via a specially designed self-adhesive electrode patch. It has an event marker button and will record for 7 days. The other is a hygrometric (humidity) monitor (Freedman monitor; Kolar Engineering, Royal Oak, MI) that attaches directly to the sternal skin via adhesive tape without electrodes or gel.[10] It records for 1 month but does not have an event marker button.

The purpose of this three-phase study was to evaluate VMS monitors before use in randomized controlled trials of VMS therapies within the Menopause Strategies Finding Lasting Answers for Symptoms and Health research network. In this study, each monitor was compared with event-marked (self-reported) VMS and one other monitor. Phase 1 evaluated the feasibility of the Freedman and prototype Bahr Monitor. The Bahr Monitor and Biolog were then evaluated under laboratory (phase 2) and ambulatory conditions (phase 3).

To be acceptable for use in the Menopause Strategies Finding Lasting Answers for Symptoms and Health clinical trials, a monitor had to show (1) high agreement between event-marked and monitor-recorded VMS in the laboratory, including the absence of monitor-recorded VMS in midlife and young women not reporting VMS; (2) high agreement between event-marked and monitor-recorded VMS in ambulatory settings, including the absence of monitor-recorded VMS in asymptomatic midlife and young women; (3) high participant acceptability ratings; and (4) high data capture rates (eg, minimal loss of data).

METHODS

Sample inclusion and exclusion criteria

Each testing phase included three groups of participants: midlife women with self-reported VMS, midlife women without VMS, and young women without VMS. Inclusion criteria were (1) female sex; (2) ability to read, write, and speak English; (3) willingness to wear monitors and provide feedback; and (4) report taking no hot flash treatments during the past month (eg, hormones, antidepressants, gabapentin, clonidine, bellargal). Exclusion criteria were (1) pregnant or lactating, (2) diagnosis of Sjogren syndrome (causes decreased sweating and altered skin conductance), (3) reported allergy to tape/adhesives or other skin sensitivities, and (4) inability to follow the study protocol.

In addition, the midlife VMS group (1) was aged 40 to 62 years, (2) reported being in the late menopausal transition (>2 skipped cycles and an interval of amenorrhea >60 days) or postmenopausal (at least 12 mo amenorrhea), and (3) reported four or more VMS per day (28 or more per week). The midlife no VMS group (1) was aged 40 to 62 years, (2) reported being in the late menopausal transition or postmenopausal as previously defined, and (3) denied having any VMS in the past 3 months. The young no VMS group (1) were aged 18 to 35 years old, (2) reported having monthly menstrual cycles, and (3) denied having any VMS.

Measures

Participants activated event markers when they felt VMS. Women waved a magnet over the monitor (Bahr Monitor, phase 1) or pressed one to two small buttons on the monitors (Bahr Monitor and Biolog, phases 2 and 3). The Freedman monitor does not have an integrated event marker, so women pressed buttons on a small wrist actigraph (Actiwatch 2; Respironics, Murrysville, PA) that was time-synchronized to the VMS monitor. Self-reported or event-marked VMS have been the comparison standard for monitor-recorded VMS in other similar device testing studies because no other validated objective measure of VMS is available.[2,7,10-12]

The Freedman monitor continuously measures relative humidity and is a lightweight (14 g) small round disk (3.8-cm diameter × 1 cm) that attaches to the sternal skin via double-sided adhesive tape. It is powered by a small hearing aid–type battery to allow for up to 1 month of recording. VMS are defined as a 3% per minute increase in relative humidity.

The Bahr Monitor in phase 1 was a preproduction 6.2 × 5.7 × 0.1–cm prototype resembling a small green circuit board. The prototype circuitry is identical to the commercially marketed 7.0 × 3.5 × 1.2–cm Bahr Monitor tested in phases 2 and 3. The Bahr Monitor applies a pulsed alternating current to the sternal skin (0.5 V for 250 ms alternating with −0.5 V for 250 ms) every 10 seconds. During the pulse, the current that passes between a single-unit electrode skin patch containing two snaps is measured in siemens. The monitor is powered by a lithium button battery that lasts for approximately 1 week of recording. VMS are defined as a sudden spike in skin conductance of 2 units or more (phase 1) or using a five-pattern recognition algorithm (phases 2 and 3).

The Biolog uses two custom electrode skin patches and a 0.5-V constant voltage circuit. The electrode skin patches are filled with a custom gel. Lead wires snap onto the electrodes

and are then linked to the 12.5 × 6.7 × 3.5–cm monitor. The Biolog contains a microprocessor that samples skin conductance for each second. It is powered by a 9-V battery that can record for up to 7 days. VMS are defined as an increase in skin conductance of 2.0 μmho or more within a 30-second period (FlashTrax; UFI).

Questionnaires were used to assess basic demographic information, record height and weight measured by study staff, and obtain feedback on the feasibility and acceptability of wearing each monitor. The latter included a series of open-ended questions (data available from the author) as well as structured questions for rating various aspects of wearing the devices on a 5-point scale from very satisfied to very dissatisfied.

Procedures

The staff posted study information in newsletters and on bulletin boards in the study buildings, local universities, and within the community. Interested women were instructed to call the study staff. The study staff provided more information about the study and determined eligibility. Interested and eligible women were scheduled for a study visit.

Laboratory visits took place in a private room at a clinical research center. Written informed consent and authorization to use health information was obtained before the start of data collection. Women were connected to the monitors, instructed on their use and care, and instructed to sit or lie quietly (reading, watching television) for the 4-hour session with cellular phones turned off. After 4 hours, the staff disconnected the monitors. The participants received a $25 gift card as compensation. Phase 1 was conducted in February 2009. Phase 2 was done between July 2, 2010, and September 16, 2010.

Phase 3 ambulatory testing was done between July 2, 2010, and September 16, 2010. Eligible and interested participants were scheduled for two private study visits 1 week apart. Written informed consent and authorization to use health information were obtained. At the first study visit, the study staff measured each participant's height and weight and provided instructions on event markers and use/care of monitors. The participants provided demographic information, were connected to the Bahr Monitor, and were then discharged to wear the monitor during their normal activities for 1 week. The participants who wore the Biolog in addition to the Bahr Monitor (Indianapolis site only) removed the Biolog after 24 hours because of its limited recording capacity. No participants reapplied the Biolog at any time during this phase of testing. At the end of the week, they returned for the second study visit. The staff collected the monitors and assessed skin integrity at the electrode sites. The participants received $70.

Data analysis

Monitor-specific software was used to generate reports for monitor-recorded VMS events. The prototype Bahr Monitor did not have any analysis software, so VMS were defined by a sudden spike in skin conductance of 2 units or more. For the commercially available Bahr Monitor, a five-pattern recognition algorithm was used by the manufacturer to generate reports for each individual participant, listing the date and time of each monitor-recorded VMS and the total number of VMS (March 2011 version). These results could not be independently verified or checked for artifact because the analysis was based on a pattern recognition algorithm rather than on absolute change in skin conductance value. The Freedman monitor software (FlashMarkPro; Kolar Engineering) applied an analysis program to identify monitor-recorded VMS defined as a 3% per minute increase in humidity. Software reports listed the time of each monitor-recorded VMS and the summed total number of VMS for each participant. The Biolog software (FlashTrax) generated similar results for monitor-recorded VMS defined as a 2-μmho increase in skin conductance within 30 seconds.

Data capture for the commercially available Bahr Monitor was evaluated using software output and quality ratings (March 2011 version). The software generated the following summary statistics: (1) number of evaluable monitoring hours, (2) percentage evaluable monitor recording (ie, evaluable monitoring hours/total number of hours), (3) total VMS detected, and (4) average VMS per 24 hours. Evaluable data were defined by the software program as ranging from 0.05 to 30 siemens. In addition, each Bahr Monitor data file was visually examined by two individuals who rated the quality of the recordings as good, fair, or poor. Good recordings were defined as those with stable baselines, clear increases in conductance with subsequent return to baseline, no artifact, and more than 6 days (85%) of recording. Fair recordings were defined as those with 1 to 2 days of poor data capture (unstable or elevated baselines, data displayed upside down because of the monitor being placed on the chest upside down, noisy signals with multiple episodes of artifact, or lost data signal [sustained recordings at 0 μmho]). Poor recordings were defined in terms of more than 2 days of poor data capture. Interrater agreement was 97% with a discrepancy resolved through discussion to reach 100% agreement.

Demographics and phase 3 acceptability data were analyzed using descriptive statistics and frequency tables. Differences in acceptability between day 3 and day 7 assessments were analyzed using χ^2 tests. Responses to open-ended questions were qualitatively examined and grouped into common themes (data available by request from authors).

RESULTS

Phase 1

Phase 1 laboratory testing was stopped after the accrual of six participants (two per group) because of problems with both monitors. On the Freedman monitor, there were 4 event-marked and 25 monitor-recorded VMS (Table 1). Monitor-recorded VMS occurred when the monitor was placed on or removed from the skin (12 of 25 events, 48%) or when the monitor was on a desktop immediately before placement on participants (3 of 25 events, 12%). The investigative group stopped further testing because of concerns that the lack of sensitivity for humidity changes as markers of VMS in the controlled laboratory environment would be exacerbated

TABLE 1. *Total number of events and percentage agreement for phases 1 and 2 laboratory studies by subgroup and measure*

	Phase 1 laboratory evaluation			Phase 2 laboratory evaluation		
	Midlife VMS (n = 2)	Midlife no VMS (n = 2)	Young no VMS (n = 2)	Midlife VMS (n = 5)	Midlife no VMS (n = 5)	Young no VMS (n = 5)
Bahr Monitor						
Total no. event marks[a]	5	0	0	13	0	1
Total no. monitor VMS[a]	5	0	1	18	0	1
Bahr Monitor to event mark						
Event mark + monitor	5	0	0	13	0	1
Event mark only	0	0	0	0	0	0
Monitor only	0	0	1	5	0	0
% Event marks with monitor-recorded VMS	100	NA	NA	100	NA	100
% Monitor-recorded VMS with event mark	100	NA	0	72	NA	100
Freedman Monitor				NA	NA	NA
Total no. event marks[a]	4	0	0			
Total no. monitor VMS[a]	12	7	6			
Freedman Monitor to event mark				NA	NA	NA
Event mark + monitor	4	0	0			
Event mark only	0	0	0			
Monitor only	8	7	6			
% Event marks with monitor-recorded VMS	100	NA	NA			
% Monitor-recorded VMS with event mark	67	0	0			
Biolog Monitor	NA	NA	NA			
Total event marks[a]				14	0	0
Total monitor VMS[a]				15	4	0
Biolog Monitor to event mark	NA	NA	NA			
Event mark + monitor				12	0	0
Event mark only				2	0	0
Monitor only				3	4	0
% Event marks with monitor-recorded VMS				86	NA	NA
% Monitor-recorded VMS with event mark VMS				80	0	NA
Bahr Monitor to Freedman				NA	NA	NA
Bahr + Freedman	5	0	1			
Bahr only	0	0	0			
Freedman only	7	7	5			
% Agreement	42	0	17			
Bahr Monitor to Biolog	NA	NA	NA			
Bahr + Biolog				14	0	0
Bahr only				4	0	1
Biolog only				1	4	0
% Agreement				74	0	0

VMS, vasomotor symptoms (hot flash, night sweat); NA, not applicable, monitor not tested.
[a]Total event marks and total monitor VMS equal the number of each after collating to remove duplicate events. Events marks in non-VMS groups can be interpreted as false button presses.

during ambulatory monitoring, thus rendering it impossible to distinguish VMS events from ambient humidity changes. On the Bahr Monitor, there seemed to be five event-marked and six monitor-recorded VMS (see Table 1). On the Bahr Monitor, 100% of monitor-recorded VMS were event-marked in midlife women with VMS, but the one monitor-recorded VMS that occurred in the young no VMS group was not event-marked. Further testing of the Bahr Monitor prototype was stopped because there was no analytic software for summarizing events, and the investigative group was told that the commercially available production version of the monitor would soon be available.

Phase 2

Fifteen participants completed phase 2; five midlife women with VMS, five midlife women without VMS, and five young women without VMS. Demographics are shown in Table 2.

On the Bahr Monitor, there were 14 event-marked and 19 monitor-recorded VMS (Table 1). On the Biolog, there were 14 event-marked and 19 monitor-recorded VMS. For event-marked VMS, 100% on the Bahr Monitor and 86% on the Biolog occurred in conjunction with a monitor-recorded VMS. For monitor-recorded VMS in the midlife VMS group, 72% on the Bahr Monitor and 80% on the Biolog were event-marked. In the no VMS groups, there was one Bahr Monitor–recorded VMS (with an event mark) and four Biolog-recorded VMS, none of which were event-marked. There was modest agreement between the Bahr and Biolog in midlife women with VMS (74%) but no agreement in midlife women without VMS (0%) or young women without VMS (0%).

Phase 3

Thirty-five participants completed phase 3; 15 at the Indianapolis site, 10 at the Seattle site, and 10 at the Oakland site. Demographics are shown in Table 2.

Bahr Monitor data from the full 7-day monitoring period showed 721 event-marked and 3,471 monitor-recorded VMS or 4.8 times more Bahr Monitor-recorded VMS than

TABLE 2. *Demographics for participants in phases 2 and 3*

	Phase 2 laboratory evaluation			Phase 3 ambulatory evaluation		
	Midlife VMS (n = 5)	Midlife no VMS (n = 5)	Young no VMS (n = 5)	Midlife VMS (n = 20)	Midlife no VMS (n = 10)	Young no VMS (n = 5)
Age, mean (SD), y	56.40 (2.97)	49.00 (4.85)	24.60 (3.44)	53.55 (4.79)	50.20 (4.47)	25.40 (4.22)
Body mass index, mean (SD), kg/m^2	32.86 (4.72)	26.67 (2.46)	29.65 (7.93)	30.83 (8.18)	30.85 (11.47)	27.37 (8.34)
Ethnicity, n (%)						
White, non-Hispanic	2 (40)	4 (80)	3 (60)	10 (50)	8 (80)	2 (40)
Black, non-Hispanic	3 (60)	1 (20)	0 (0)	8 (40)	2 (20)	0 (0)
Other, non-Hispanic	0 (0)	0 (0)	2 (40)	2 (10)	0 (0)	3 (60)
Overall health, n (%)						
Excellent	0 (0)	3 (60)	1 (20)	5 (25)	4 (40)	1 (20)
Very good	0 (0)	2 (40)	3 (60)	6 (30)	5 (50)	3 (60)
Good/fair	5 (100)	0 (0)	1 (20)	9 (45)	1 (10)	1 (20)
Marital status, n (%)						
Not married or partnered	4 (80)	1 (20)	5 (100)	12 (60)	4 (40)	5 (100)
Currently married or partnered	1 (20)	4 (80)	0 (0)	8 (40)	6 (60)	0 (0)
Education, n (%)						
High school diploma/GED	1 (20)	0 (0)	1 (20)	2 (10)	0 (0)	1 (20)
Some training after high school	3 (60)	2 (40)	0 (0)	9 (45)	4 (40)	0 (0)
College graduate	0 (0)	1 (20)	2 (40)	4 (20)	2 (20)	3 (60)
Graduate schooling/degree	2 (40)	2 (40)	2 (40)	5 (25)	4 (40)	1 (20)
Employment						
Full time	3 (60)	4 (80)	2 (40)	14 (70)	7 (70)	3 (60)
Part time	0 (0)	0 (0)	3 (60)	4 (20)	2 (20)	2 (40)
Not currently working	2 (40)	1 (20)	0 (0)	2 (10)	1 (10)	0 (0)
Menopause status						
Premenopause	0 (0)	0 (0)	5 (100)	0 (0)	0 (0)	5 (100)
Perimenopause	0 (0)	3 (60)	0 (0)	4 (20)	5 (50)	0 (0)
Postmenopause	5 (100)	2 (40)	0 (0)	16 (80)	5 (50)	0 (0)

VMS, vasomotor symptom (hot flash, night sweat).

event-marked VMS (Table 3). Event-marked VMS in the midlife no VMS and young no VMS groups were false button presses because these groups verified that they did not have VMS. The percentage of event-marked VMS occurring in conjunction with Bahr Monitor–recorded VMS was 59% in the midlife VMS group, 20% in midlife no VMS, and 0%

TABLE 3. *Total number of events and percentage agreement for phase 3 ambulatory study by subgroup and measure*

	Phase 3 ambulatory evaluation 7 days			Phase 3 ambulatory evaluation 1st 24 hours		
	Midlife VMS (n = 20)	Midlife no VMS (n = 10)	Young no VMS (n = 5)	Midlife VMS (n = 5)	Midlife no VMS (n = 5)	Young no VMS (n = 5)
Bahr Monitor						
Total no. event marks[a]	705	10	6	32	2	5
Total no. monitor VMS[a]	2,286	673	512	111	71	92
Bahr Monitor to event mark						
Event mark + monitor	439	2	0	25	0	0
Event mark only	266	8	6	7	2	5
Monitor only	1,847	671	512	86	71	92
% Event marks with monitor-recorded VMS	59%	20%	0%	78%	0%	0%
% Monitor-recorded VMS with event mark	19%	<1%	0%	23%	0%	0%
Biolog Monitor	NA	NA	NA			
Total event marks[a]				50	14	14
Total monitor VMS[a]				49	55	42
Biolog Monitor to event mark	NA	NA	NA			
Event mark + monitor				22	3	3
Event mark only				28	11	11
Monitor only				27	52	39
% Event marks with monitor-recorded VMS				44%	21%	21%
% Monitor-recorded VMS with event mark VMS				45%	5%	7%
Bahr Monitor to Biolog Monitor	NA	NA	NA			
Bahr + Biolog				39	23	29
Bahr only				72	48	63
Biolog only				10	32	13
% Agreement				32%	22%	28%

Biolog worn only by Indianapolis site field participants.
VMS, vasomotor symptoms (hot flash, night sweat); NA, not applicable.
[a]Total event marks and total monitor VMS represent the number of each after collating to remove duplicate events. Events marks in non-VMS groups can be interpreted as false button presses.

TABLE 4. *Bahr Monitor data capture information*

	Study 2 field evaluation 7 days		
	Midlife VMS (n = 20)	Midlife no VMS (n = 10)	Young no VMS (n = 5)
Quality ratings, %[a]			
Good	20%	40%	0%
Fair	45%	20%	40%
Poor	35%	40%	60%
Percentage evaluable data, %[b]	79%	81%	83%
Time monitored, mean (SD), h:min[c]	168:45 ± 00:10	168:44 ± 00:25	168:38 ± 00:07
Evaluable time, mean (SD), h:min[d]	133:15 ± 00:31	136:16 ± 00:12	139:38 ± 00:02
Evaluable 24-h periods, mean (SD)[e]	5.55 (1.35)	5.68 (0.88)	5.82 (0.96)
Average VMS detected, mean(SD)[f]	26.39 (8.46)	13.97 (7.61)	22.98 (12.28)
Total power events, mean (SD)[g]	4.10 (1.25)	4.90 (3.48)	5.00 (0.71)

VMS, vasomotor symptoms (hot flash, night sweat).
[a]Quality ratings were defined through visual inspection by two raters. Ideal is 100% good quality.
[b]Ideal percentage is 100%.
[c]Ideal time monitored is 168 hours (7 days × 24 hours per day).
[d]Ideal evaluable time is 168 hours (7 days × 24 hours per day).
[e]Ideal number of periods is 7.00 (7 days of monitoring).
[f]Ideal number in midlife no VMS and young no VMS is 0 (no hot flashes or night sweats).
[g]Number of times the unit was powered off and on during the data collection: ideal is ≤1.

in the young no VMS group. Conversely, 19%, less than 1%, and 0% of Bahr Monitor-recorded VMS occurred in conjunction with an event marker in the three groups, respectively. The average number of Bahr Monitor–recorded VMS per day (number of VMS per group/number of women per group for 7 days) was 16.33 in midlife women with VMS, 9.61 in midlife women without VMS, and 14.63 in young women without VMS.

Data for the first 24 hours, when both the Bahr Monitor and Biolog were worn by the Indianapolis participants, are also in Table 3. For the Bahr Monitor, there were 39 event-marked VMS and 274 monitor-recorded VMS. The percentage of event-marked VMS that occurred with a Bahr Monitor-recorded VMS was 78% in the midlife VMS group and 0% in the midlife no VMS and young no VMS groups. The percentage of monitor-recorded VMS that were event-marked was 23% in the midlife VMS group and 0% in the no VMS groups. For the Biolog, there were 78 event-marked VMS and 146 monitor-recorded VMS, with slightly higher agreement rates than the Bahr Monitor. The percentage of event-marked VMS that occurred with a Biolog monitor–recorded VMS was 44% in the midlife VMS group, 21% in midlife no VMS, and 21% in young no VMS groups. The percentage of Biolog monitor-recorded VMS that were event-marked was 45% in the midlife VMS group, 5% in the midlife no VMS group, and 7% in the young no VMS group. Bahr Monitor to Biolog agreement rates were poor in all three groups (32% in midlife VMS group, 22% in midlife no VMS group, 28% in young no VMS group), with the highest proportion of VMS being recorded only by the Bahr Monitor. In calculating the average number of monitor-recorded VMS per group, we noted that approximately twice as many VMS were recorded by the Bahr Monitor as compared with the Biolog. The average number of Bahr Monitor–recorded versus Biolog-recorded VMS (number of VMS per group/number of women per group) was 22.2 versus 9.8 in midlife women with VMS, 14.2 versus 11.0 in midlife women without VMS, and 18.4 versus 8.4 in young women without VMS.

Data capture rates indicated that Bahr Monitor data were not evaluable approximately 20% of the time, for a total of about 1.5 days of missing data per 7 days of monitoring per participant (Table 4). The Biolog data capture is limited to 24 hours or a single day, and all data were available for that period of observation.

There were no statistically significant differences between acceptability ratings on days 3 and 7; thus, only day 7 ratings are shown in Figure 1. Ratings for the Bahr Monitor were more positive in terms of ease of use, freedom of movement, ability to sleep, adherence to skin, and comfort and weight. Complaints of itchiness from the electrodes were similar for both monitors. Responses to open-ended questions for the

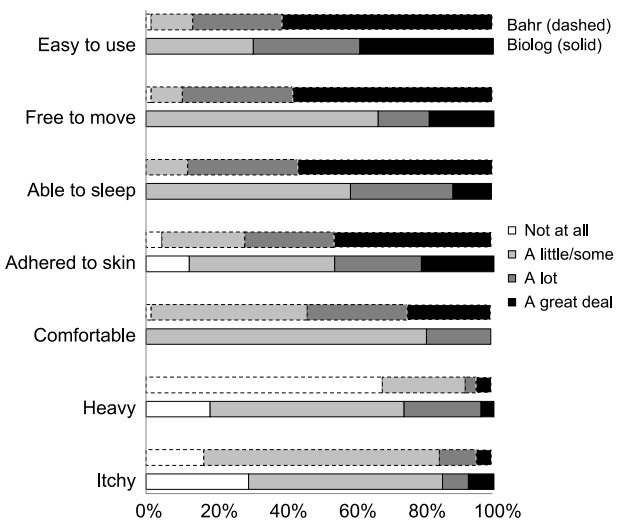

FIG. 1. Field study day 7 monitor acceptability ratings. The responses shown are valid percentages of answers to each question accounting for missing data. On day 7, 35 of 35 participants rated the Biolog monitor items and 11 to 14 of 15 participants rated the Bahr Monitor. The Biolog was worn in the ambulatory study only for the first 24 hours by Indianapolis site participants. Solid lines are responses for Biolog. Dashed lines are responses for Bahr Monitor.

Bahr Monitor reflected annoyance at wearing a device that itched and interfered with clothing choices.

DISCUSSION

This report resembles other device evaluation studies in terms of sample size, higher concordance during laboratory versus ambulatory monitoring, and capture of VMS events in women who do not report VMS. VMS monitors have been evaluated in samples of 20 women or less with VMS and 5 or fewer women without VMS[2,10-13], with fewer studies including 30 women or more (Fig. 2).[7,14] Although VMS monitors typically perform with near-perfect correspondence to event-marked VMS in the laboratory when women are resting and paying attention to VMS,[10,11] correspondence decreases in the ambulatory setting when women are more physically active and potentially distracted from pushing event markers.[15] Our testing was done in both environments; thus, the variation in estimates across phases reflects the setting rather than the low sample sizes. In addition, our data showed that VMS monitors record changes in skin conductance that meet criteria for VMS events in younger premenopausal women without VMS, consistent with previous research.[2]

This report differs from previous reports in the inclusion of the Bahr Monitor, use of different Biolog electrodes, and Freedman monitor findings. This seems to be the first published report of the Bahr Monitor. The Biolog electrodes used here were manufactured differently than ones used in the past and did not perform well in the ambulatory setting. The importance of the electrode and gel configuration for the Biolog has been previously noted.[13] In previous ambulatory studies using various electrodes for the Biolog, 18% to 45% of VMS are both monitor-recorded and self-reported, 19% to 42% are monitor-recorded only, and 22% to 23% are self-reported only.[1,2,4,7,8,14,15] The results we report here may differ from previous reports because of use of different electrodes. In addition, correspondence rates for the Freedman monitor were lower than a previous report showing 100% agreement between 20 event-marked and 20 monitor-recorded VMS recorded among 10 women in a laboratory.[10] We could find no published reports of Freedman monitor testing in midlife women without VMS or young women without VMS; thus, our 0% agreement between monitors in these groups represents a new finding. Our findings were partially caused by the Freedman monitor's sensitivity in detecting humidity changes from staff handling during placement/removal/ambient changes. Although some of these events could be edited out with careful review by a trained rater (eg, VMS occurring at placement/removal), it would be quite difficult to discern menopausal VMS from ambient humidity changes during a laboratory or ambulatory monitoring session. Further improvements in device hardware, scoring algorithms, and/or software may help decrease noise caused by other physiological processes and may improve the detection of menopausal VMS.

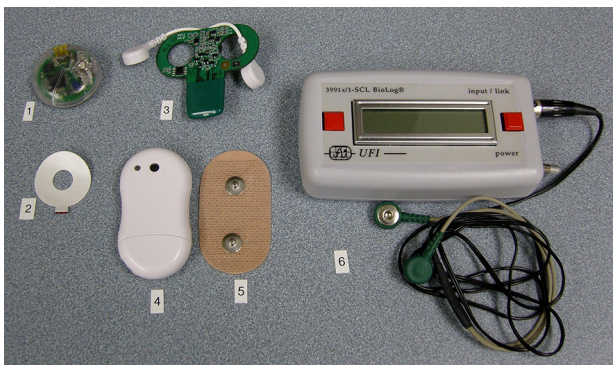

FIG. 2. VMS monitoring systems. The photo shows the monitoring systems used in the study. From left to right are the (1) Freedman FlashMark Pro minitiarized hygrometer with (2) double sided tape ring to adhere the device to the skin; (3) prototype Bahr Monitor resembling a small green circuit board, (4) production version Bahr Monitor, (5) electrode used to adhere the Bahr Monitor to the skin; and (6) Biolog with electrode lead wires.

Recommendations

Three recommendations emerge from our data. First, the Bahr Monitor and Biolog seem appropriate for laboratory-based studies such as those evaluating VMS etiology. This recommendation is based on high agreement between event-marked and monitor-recorded VMS in midlife women with VMS and the relative absence of monitor-recorded VMS in non-VMS groups. However, the results from either monitor should be interpreted cautiously because both recorded at least some VMS in non-VMS groups and the monitors were not perfectly correlated with one another. The lack of monitor-to-monitor agreement may be attributable to differences in electrical systems, sampling rates, electrode size, configuration, or gel. When used in the laboratory, it may be difficult to interpret which monitor is more accurate because monitors corresponded well to event marks but not to each other.

Second, currently available VMS monitors may not be suitable for use in clinical trials evaluating the efficacy of interventions for VMS in ambulatory participants until the systems can be further developed and their present limitations addressed. This recommendation is based on (1) relatively modest agreement between event-marked and monitor-recorded VMS in women with VMS, (2) the large number of monitor-recorded VMS in non-VMS groups, (3) the presence of event markers in non-VMS groups, (4) participant feedback about device acceptability, (5) loss of data during a 7-day monitoring with the Bahr Monitor, and (6) the lack of a commercially available analysis program or algorithm to independently verify results for the Bahr Monitor. Although it was expected that ambulatory monitors might record more VMS than self-report, both Bahr Monitor– and Biolog-recorded VMS far exceeded women's event-marked VMS. The number of monitor-recorded events in the two no VMS groups is of particular concern because it seems that the monitors are not specific to detecting menopausal VMS and may instead be picking up noise or artifacts related to perspiration or other undetermined physiological processes. These findings raise concerns about whether current monitors

are useful adjuncts to subjective reports in ambulatory studies or a new source of VMS measurement error.

Third, further development and refinement of VMS monitoring systems may overcome present limitations. Objective VMS monitors may provide useful information about physiological occurrences of VMS, and their ability to capture the time of VMS onset may be useful in etiological studies. However, the excess occurrence of monitor-recorded VMS and the relative lack of agreement between monitor-recorded VMS and event-marked VMS in the ambulatory study suggest that the events measured and defined by the monitors as VMS are not necessarily the same events perceived and reported by respondents as VMS. In certain respects, this is analogous to the lack of congruence between self-reported physical activity and device-based measures of physical activity, such as accelerometry or heart rate monitoring.[16] In the case of physical activity, the physiological processes measured by various devices may not adequately capture the full range and complexity of physical activity behavior, although in the case of VMS, the physiological processes may not always reflect the experience of VMS as perceived by the individual. Continued development and testing of VMS monitors may address this issue and lead to improved ability in measuring VMS.

Limitations

The limitations of this study include the following. Freedman device testing was done in a small sample of women. Although we intended to study this device in a larger sample, the initial results showing lack of correspondence between event-marked and monitor-recorded VMS led us to discontinue testing the device. The Bahr Monitor that was tested evolved from a preproduction prototype (phase 1) to a commercially available version (phase 2). Although the internal circuitry and electrodes were the same in both versions, the event mark differed from a magnet passing over the monitor to a button on the face of the monitor. This change could have affected results. Bahr Monitor results were based on a proprietary algorithm (March 2011) and could not be independently verified. The algorithm remains under development, and subsequent improvements may lead to different results.

CONCLUSIONS

We noted limitations in all of the monitoring systems we tested. The recommendations are to (1) use either the Bahr Monitor or Biolog in short-term laboratory studies with subsequent caution in interpreting findings and (2) carefully consider the development status and limitations of VMS monitoring systems before using any in ambulatory studies, such as those evaluating the efficacy of VMS therapies.

Acknowledgments: We thank John Peterman for his assistance in applying the five-pattern recognition algorithm to the Bahr Monitor files (March 2011).

REFERENCES

1. Carpenter JS, Storniolo AM, Johns S, et al. Randomized, double-blind, placebo-controlled crossover trials of venlafaxine for hot flashes after breast cancer. *Oncologist* 2007;12:124-135.
2. Carpenter JS, Andrykowski MA, Freedman RR, Munn R. Feasibility and psychometrics of an ambulatory hot flash monitoring device. *Menopause* 1999;6:209-215.
3. Carpenter JS, Monahan PO, Azzouz F. Accuracy of subjective hot flush reports compared with continuous sternal skin conductance monitoring. *Obstet Gynecol* 2004;104:1322-1326.
4. Thurston RC, Blumenthal JA, Babyak MA, Sherwood A. Emotional antecedents of hot flashes during daily life. *Psychosom Med* 2005;67: 137-146.
5. Carpenter JS, Azzouz F, Monahan PO, Storniolo AM, Ridner SH. Is sternal skin conductance monitoring a valid measure of hot flash intensity or distress? *Menopause* 2005;12:512-519.
6. Loprinzi CL, Barton DL, Sloan JA. Whose opinion counts? *J Clin Oncol* 2006;24:5183-5185.
7. Sievert LL, Freedman RR, Garcia JZ, et al. Measurement of hot flashes by sternal skin conductance and subjective hot flash report in Puebla, Mexico. *Menopause* 2002;9:367-376.
8. Sievert LL, Reza A, Mills P, et al. Diurnal rhythm and concordance between objective and subjective hot flashes: the Hilo Women's Health Study. *Menopause* 2010;17:471-479.
9. Joffe H, Partridge A, Giobbie-Hurder A, et al. Augmentation of venlafaxine and selective serotonin reuptake inhibitors with zolpidem improves sleep and quality of life in breast cancer patients with hot flashes: a randomized, double-blind, placebo-controlled trial. *Menopause* 2010; 17:908-916.
10. Freedman RR, Wasson S. Miniature hygrometric hot flash recorder. *Fertil Steril* 2007;88:494-496.
11. Freedman RR. Laboratory and ambulatory monitoring of menopausal hot flashes. *Psychophysiology* 1989;26:573-579.
12. Tataryn IV, Lomax P, Meldrum DR, Bajorek JG, Chesarek W, Judd HL. Objective techniques for the assessment of postmenopausal hot flashes. *Obstet Gynecol* 1981;57:340-344.
13. Dormire SL, Carpenter JS. An alternative to Unibase/glycol as an effective nonhydrating electrolyte medium for the measurement of electrodermal activity. *Psychophysiology* 2002;39:423-426.
14. Sievert LL, Begum K, Sharmeen T, Chowdhury O, Muttukrishna S, Bentley G. Patterns of occurrence and concordance between subjective and objective hot flashes among Muslim and Hindu women in Sylhet, Bangladesh. *Am J Hum Biol* 2008;20:598-604.
15. Mann E, Hunter MS. Concordance between self-reported and sternal skin conductance measures of hot flushes in symptomatic perimenopausal and postmenopausal women: a systematic review. *Menopause* 2011;18: 709-722.
16. Troiano RP, Berrigan D, Dodd KW, Masse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. *Med Sci Sports Exerc* 2008;40:181-188.