HHS Public Access
Author manuscript
Menopause. Author manuscript; available in PMC 2018 August 01.

Published in final edited form as:
Menopause. 2017 August ; 24(8): 877–885. doi:10.1097/GME.0000000000000871.

Validity, Cutpoints, and Minimally Important Differences for Two Hot Flash Related Daily Interference Scales

Janet S. Carpenter, PhD, RN, FAAN[1], **Giorgos Bakoyannis, PhD**[1,2], **Julie L. Otte, PhD, RN, OCN**[1], **Chen X. Chen, PhD, RN**[1], **Kevin L. Rand, PhD**[3], **Nancy Woods, PhD, RN, FAAN**[4], **Katherine Newton, PhD**[5], **Hadine Joffe, MD, MSc**[6], **JoAnn E. Manson, MD, DrPH**[7], **Ellen W. Freeman, PhD**[8], and **Katherine A. Guthrie, PhD**[9]

[1]School of Nursing, Indiana University, Indianapolis, IN

[2]Department of Biostatistics, Fairbanks School of Public Health, School of Medicine, Indiana University, Indianapolis, IN

[3]Department of Psychology, School of Science, Indiana University-Purdue University Indianapolis

[4]University of Washington School of Nursing, Seattle, WA

[5]Group Health Research Institute, Seattle, WA

[6]Brigham and Women's Hospital, Dana Farber Cancer Institute, Harvard Medical School, Boston, MA

[7]Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, and Harvard School of Public Health, Boston, MA

[8]Departments of Obstetrics/Gynecology and Psychiatry, University of Pennsylvania School of Medicine, Philadelphia, PA

[9]MsFLASH Data Coordinating Center, Fred Hutchinson Cancer Research Center, Seattle, WA

Abstract

Objectives—To conduct psychometric analyses to (1) condense the Hot Flash Related Daily Interference Scale (HFRDIS) into a shorter form termed the Hot Flash Interference (HFI) scale, (2) evaluate cutpoints for both scales, and (3) establish minimally important differences (MIDs) for both scales.

Methods—We analyzed baseline and post-randomization patient-reported data pooled across three randomized trials aimed at reducing vasomotor symptoms (VMS) in 899 midlife women. Trials were conducted across five MsFLASH clinical sites between July 2009 and October 2012. We eliminated HFRDIS items based on experts' content validity ratings and confirmatory factor analysis and evaluated cutpoints and established MIDs by mapping HFRDIS and HFI to other measures.

Address for correspondence and reprints: Janet S. Carpenter, PhD, RN, FAAN, IU School of Nursing, 600 Barnhill Drive NU338, Indianapolis, IN 46260, phone 317-278-6093, fax none, carpentj@iu.edu.

Conflict of Interest: None.

Disclaimers (if any): None.

Results—The 3-item HFI (interference with sleep, mood, and concentration) demonstrated strong internal consistency (alphas of 0.830 and 0.856), showed good fit to the unidimensional "hot flash interference factor", and strong convergent validity with HFRDIS scores, diary VMS, and menopausal quality of life. For both scales, cutpoints of mild (0–3.9), moderate (4–6.9), and severe (7–10) interference were associated with increasing diary VMS ratings, sleep, and anxiety. The average MID was 1.66 for the HFRDIS and 2.34 for the HFI.

Conclusions—The HFI is a brief assessment of VMS interference and will be useful in busy clinics to standardize VMS assessment or in research studies where response burden may be an issue. The scale cutpoints and MIDs should prove useful in targeting those most in need of treatment, monitoring treatment response, and interpreting existing and future research findings.

Keywords

Menopause; Hot Flashes; Psychometrics; Quality of Life

INTRODUCTION

The Hot Flash Related Daily Interference scale (HFRDIS) is a widely used, psychometrically sound, 10-item, self-report questionnaire assessing the impact of hot flashes on a woman's life.[1] The scale was introduced in 2001 to capture a symptom dimension (interference) that had not been previously considered in vasomotor symptom research because there were no assessment tools.[1] The scale was based on similar pain interference[2] and fatigue interference[3] measures.[1] Since publication, the HFRDIS has been translated to 12 languages, cited over 175 times in journal articles and textbooks, and was included as an outcome measure within the National Cancer Institute Physician Data Query Cancer Information Summaries for Supportive and Palliative Care (Coping with Cancer).[4] The HFRDIS is psychometrically strong with demonstrated responsiveness to interventions such as pharmacologic treatments, dietary supplements, and behavioral therapies.[5–9]

Interpretation of HFRDIS scores in research and clinical practice is limited by a lack of investigation in two areas. Cutpoints delineating mild, moderate, and severe interference and a minimally important difference (MID) have never been evaluated. A MID is "the smallest difference in score in the domain of interest that patients perceived as important, either beneficial or harmful, and that would lead the clinician to consider a change in the patient's management" (p. 377).[10]

In addition, although the scale has only 10 items, a shorter scale could be more useful in busy clinical practices or in research where the HFRDIS is one of many measures to be completed by participants. Anecdotally, in using the scale over the past 15 years, the first author has noticed a pattern to women's responses where a few items are more likely to be rated as zero or no interference, suggesting they may not be important items. Thus, there may be potential to create a shorter scale. According to psychometric theory, short form scales can be developed through psychometric testing and can perform equally well as the longer scales from which they are derived.[11] Gaining expert opinion on the content validity of items, and performing psychometric analyses of women's responses to the scale could

help identify the most salient items to retain in a shorter version of the scale. Such analyses have not been reported.

Therefore, given the widespread use yet limited knowledge of HFRDIS cutpoints and MIDs and the desire to create a shorter scale, we conducted psychometric analyses to (1) condense the HFRDIS into a shorter Hot Flash Interference (HFI) scale, (2) establish cutpoints for both scales, and (3) establish MIDs for both scales.

METHODS

Design

This was an analysis of baseline and post-randomization data from 899 peri- and post-menopausal community-dwelling women who participated in MsFLASH trials 01, 02, and 03. The three MsFLASH studies used standardized methods.[12, 13] Common to all studies was a minimum eligibility criterion of 14 hot flashes per week. All studies were approved by the Institutional Review Boards at clinical sites and the Data Coordinating Center (DCC). Participants provided written informed consent and authorization to use protected health information.[14–18]

MsFLASH 01 was a randomized, placebo-controlled, double-blind trial of escitalopram 10 mg/day or placebo for 8 weeks.[14] If a reduction in VMS frequency of ≥ 50% or a decrease in VMS severity after 4 treatment weeks was not achieved, the dose was increased to 20 mg/day (or matched placebo) without revealing the randomization. Approximately equal numbers of African-American and White women were enrolled. Baseline and 8-week post-randomization data were used in this analysis.

MsFLASH 02 was a 3×2 factorial, randomized, controlled trial.[15–17] We randomized eligible participants in a 3:3:4 ratio to 12 weeks of yoga, exercise, or usual activity, and simultaneously randomized in a 1:1 ratio to 1.8 g/day of omega-3 fatty acid or placebo. The omega-3 component of the trial was double-blinded.[15] Baseline and 12-week post-randomization data were used in this analysis.

MsFLASH 03 was a randomized, placebo-controlled, double blind, 8-week trial of low-dose oral 17-beta-estradiol 0.5 mg/day, venlafaxine XR (37.5 mg/day the first week, then 75 mg/day), or placebo in a 2:2:3 ratio.[18] Baseline and 8-week post-randomization data were used in this analysis.

Setting and Participants

Participants were recruited from July 2009 to October 2012, primarily by mass mailings to age-eligible women using purchased mailing lists and health-plan enrollment files. Common inclusion criteria for all trials included: women aged 40-62 years, in the menopause transition (amenorrhea ≥ 60 days in the past year), or postmenopausal (≥ 12 months since last menstrual period or bi-lateral oophorectomy), or had a hysterectomy with one or both ovaries remaining and FSH > 20 mIU/mL and estradiol ≤ 50 pg/mL, and in general good health as determined by medical history, physical exam, and blood tests. Hot flashes had to be rated as bothersome or severe on at least 4 days or nights per week, and the frequency in

screening week 3 could not decrease > 50% from the mean weekly levels in screening weeks 1 and 2. Common exclusion criteria were: use of prescription or over-the-counter treatments for hot flashes (past 30 days), use of hormones or hormonal contraceptives (past 2 months), pregnancy or breastfeeding, any current severe or unstable medical conditions, drug or alcohol abuse (past year), history of myocardial infarction, angina or cerebrovascular events, or a major depressive episode (past 3 months).

Procedures

Interested women completed telephone screening and if eligible, completed a 2-week VMS diary and questionnaire with more detailed screening questions. Those who remained eligible completed a clinic visit for baseline assessment. They continued with the diary and completed a second clinic visit to confirm eligibility (second baseline visit) and complete randomization. To encourage protocol adherence and assess adverse events, telephone calls were made one or two weeks after randomization, and then again midway through the intervention periods. Follow-up clinic visits were conducted at midpoint and end of study: 4 and 8 weeks post-randomization (MsFLASH 01 and 03) or 6 and 12 weeks post-randomization (MsFLASH 02).

Measures

Baseline demographic characteristics collected from all women included age, race, ethnicity, menopausal status, education, and income. Height and weight were assessed by study staff in clinics for calculating body mass index.

VMS frequency, severity, and bother were collected in the morning and at bedtime on a paper-based daily diary. Ratings of severity were 0 (mild) to 2 (severe) and of bother were 0 (not at all) to 3 (a lot). Ratings were used to calculate daytime, nighttime, and total VMS frequency, severity, and bother. Higher scores indicated worse outcomes.

The following measures were completed at a baseline clinic visit pre-randomization and again post-randomization (8 weeks for MsFLASH 01 and 03; 12 weeks for MsFLASH 02): the HFRDIS, Menopause Quality of Life Scale (MENQOL), two sleep measures (Pittsburgh Sleep Quality Index (PSQI) and Insomnia Severity Index (ISI), and the Generalized Anxiety Disorders-7 (GAD-7) questionnaire.

The 10-item HFRDIS measures how much hot flashes have interfered with the following 10 aspects of life in the past week: interference with work, social activities, leisure activities, sleep, mood, concentration, relations with others, sexuality, enjoyment of life, and life satisfaction. Each item is rated from 0 (none) to 10 (extremely). Responses were averaged to range from 0 to 10 with higher scores indicating greater interference (worse outcomes).

The 29-item MENQOL[19] assesses menopause-related quality-of-life over the past 4 weeks. The scale has four domains (vasomotor, physical, psychosocial, and sexual). Subscores on each domain range from 1 (none) to 8 (worst) quality of life. Total MENQOL scores are calculated as the average of the 4 domain subscores. The MENQOL has good internal consistency (Cronbach's alpha of 0.81 to 0.87 for each domain), adequate test-retest reliability, good validity, and demonstrated sensitivity to change with treatment.[20]

The 19-item PSQI assesses global sleep as well as 7 components of sleep over the past 4 weeks: sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbance, use of sleep medications, and daytime dysfunction.[21, 22] Items use varying response categories and scores are calculated using a computational algorithm. Global scores greater than 5.0 are indicative of poor sleep quality and high sleep disturbances, and global scores ≥ 8 have been linked also to fatigue.[23] Studies have documented one, two, and three factor structures for the PSQI, with a three factor model best representing sleep in midlife women with hot flashes.[24] The scale has demonstrated sensitivity to detect changes in sleep during treatment for primary insomnia[25] and sleep problems concurrent with vasomotor symptoms.[5]

The 7-item ISI assesses insomnia severity over the past 2 weeks. Participants rate each item on a 0 to 4 point scale (totals range from 0 to 28). Cutpoint scores for severity of insomnia are none (0–7), subthreshold (8–14), moderate (15–21), and severe (22–28).[26] The ISI has the following psychometric properties: (1) good internal consistency with Cronbach's alphas ≥ 0.90 in people with and without insomnia;[26] (2) concurrent validity with measures of fatigue, quality of life, and mood;[26] and (3) responsiveness to treatments for primary insomnia[26] and insomnia concurrent with vasomotor symptoms.[27]

The GAD-7 is a 7-item screening and severity measure validated for the four most common anxiety disorders in primary care: generalized anxiety, panic, social anxiety, and posttraumatic stress disorder. Responses to each item are for the past 2 weeks and range from 0 (not at all) to 4 (extremely) with total scores of 0 to 21. Reliability and validity have been established among more than 2700 general medical outpatients.[28] GAD-7 cutpoints of 5, 10, and 15 represent mild, moderate, and severe anxiety symptoms, respectively.[28, 29]

Statistical Analysis

Demographics were compared across the three trials using Pearson's chi-squared and Kruskal-Wallis tests as appropriate.

Condensing the Scale—Following recommendations that five to ten experts rate items for content validity,[30] we had seven menopause experts holding MD or PhD degrees and actively involved in menopause research and/or menopause clinical practice rate each HFRDIS item as: a) essential, b) useful but not essential, or c) not necessary. Ratings were done based on their own expertise and experience without knowledge of participants' ratings or scores. We then calculated the content validity ratio (CVR) for each item and selected items to retain in the shorter scale based on a CVR cutpoint value ≥ 0.622.[31] CVRs range from −1 (none of the experts rated the item as essential) to 1 (all of the experts rated the item as essential), with 0 indicating that 50% of the experts rated the item as essential. A CVR of 0.622 corresponded to seven experts and a one-sided test of the null hypotheses that ≤ 50% of the experts would rate the item as essential at an alpha level of 0.05.[31] Any item below this cutpoint was removed. We also examined item distributions. Total scores on both the HFRDIS and HFI were calculated as the average of the items. Since the possible response range for each item was 0 to 10, the average total score on both scales had a possible range of 0 to 10.

We conducted several analyses to evaluate the shortened scale. First, we calculated Cronbach's alphas for the full and shortened scales using data from the total cohort and each trial at baseline and post-randomization.

Next and because the HFI was conceptualized as measuring a single hot flash interference construct, we sought to verify that a one-factor model showed good fit to the original 10-item HFRDIS data. To assess the fit of the unidimensional hot flash interference construct on the data we performed four confirmatory factor analyses based on unweighted least squares estimation of the model parameters: total sample and the three trial samples. Good fit of the models was defined as a Comparative Fit Index > 0.95 and a Tucker-Lewis Fit Index > 0.95.[32] We also assessed Standardized Root Mean Square Residual, a reverse measure of fit with a value of 0 indicating perfect fit and values < 0.08 indicating good fit.[32]

Third, we correlated the full and shortened scales with one another, diary outcomes, and MENQOL scores at each time point using data from the total cohort. These associations were estimated using the nonparametric Spearman's rank correlation coefficient (Spearman's rho).

Evaluating Cutpoints—We used a theoretically driven approach to evaluate cutpoints for the HFRDIS and HFI in relation to other menopausal symptoms. We applied numeric bounds of mild (0–3.9), moderate (4–6.9) and severe (7–10). We then compared VMS diary, sleep and anxiety scores among the mild, moderate, and severe HFRDIS groups and among the mild, moderate, and severe HFI groups. For diary ratings, we used Kruskall-Wallis tests. For sleep and anxiety we used Kendall's tau coefficients.

The numeric bounds were based on the authors' experiences with how 0-10 point numeric rating scales are commonly interpreted in clinical practice and the only published study of symptom interference cutpoints we could locate.[33] That study also used symptom severity ratings to establish interference cutpoints. Our theoretically driven approach is advantageous over a data driven approach for two reasons. First, it is less dependent on the idiosyncrasies of a particular sample and distribution of data within that sample. Second, it ensures there are common rather than unique numeric bounds across different symptoms or symptom dimensions. Data-driven approaches to establishing severity and interference cutpoints have resulted in equivocal findings across samples and across symptoms.[34–37] Using common cutpoints ensures easier translation into practice.

Establishing MIDs—We calculated HFRDIS and HFI MIDs using anchor-based methods. First, we defined the MID as the mean change in these scales from baseline to post-randomization (8 or 12 weeks) for women who reported a 40%-60% reduction in VMS frequency during a trial. We used this cutpoint because a 50% reduction in hot flashes has been suggested as the minimally clinically important difference.[5, 38] We chose ± 10% as a reasonable boundary to ensure there was an adequate sample size of women for the MID analyses. Second, we calculated the MID as the mean change in scores for those that reported a 0.5 to 1.5 decrease in MENQOL total score, which is equal to the percent changing one or more response categories within a domain. We also conducted post hoc analyses to calculate the MID as the median change in scores and found similar results (i.e.

MIDs within 0.25 points of those found in the mean analysis). Results confirmed the mean analysis was not influenced by outliers or skewness, thus, only the mean MID analysis results are presented.

Sample Size Justification and Software—We evaluated statistical power post hoc using our available sample and our main hypothesis of a one-factor model for the HFRDIS. Power in confirmatory factor analysis using structural equation modeling is still a developing area of inquiry, so sample size guidelines are based on expert recommendations in two areas. First, for testing overall fit of a model, experts recommend a minimum sample size of 200[39] which we exceeded with our overall sample size (N = 899) and each of our subsamples (n = 205, n = 355, n = 339). Second, for accurately estimating model parameters, Jackson[40] recommends using an optimal participant to parameter ratio of 20:1, although ratios as low as 5:1 are considered acceptable.[41] With a 10-item scale modeled as 1 factor, there are 20 parameters being estimated (e.g., 1 factor loading plus 1 residual/error parameter per item). Thus, the participant to parameter ratio for the HFRDIS in our study was 45:1 for the overall sample, and > 10:1 for each of the subsamples.

All statistical analyses except for the confirmatory factor analysis were performed using STATA 14 for Windows. The lavaan[42] statistical package in R was used for the confirmatory factor analyses.

RESULTS

Demographics of the 899 participants in the three trials are shown in Table 1. Most participants were in their middle 50's, White or African-American, educated beyond high school, never smokers, overweight or obese, and postmenopausal. In all trials, median VMS frequency at baseline was more than 6 per day with moderate severity and bother, and median hot flash interference was around 3.

Condensed Scale

Shown in Table 2 are the experts' CVRs and participants' pooled descriptive data by HFRDIS item. Based on the pre-specified CVR cutpoint value, we eliminated all items except sleep, mood, and concentration in the HFI. Eliminated items were more frequently endorsed by participants' as zero or no interference (22% to 37.9%) compared to the three retained items (2.2% to 18.1%).

Cronbach's alphas are shown in Table 3. Alphas for the HFRDIS exceeded 0.92 for all time points and samples. Alphas for the HFI exceeded 0.82 for all time points and samples.

The confirmatory factor analyses suggested that a one-factor model of the HFRDIS fit the data well in the pooled sample as well as in each trial sample (see Table 4). Because the one-factor model of the HFI was saturated (i.e., just identified with three indicators), all indices showed perfect fit in each sample. In addition, all items for both scales had large factor loadings.

Correlations showing good convergent validity of the HFI are shown in Table 5. The HFI exhibited strong and positive associations with the HFRDIS (rho ≥0.907, $p < 0.001$). The pattern of correlations between the HFRDIS and diary and MENQOL was similar in direction, strength, and significance to the HFI and those outcomes.

Cutpoints

Additional associations are shown in Table 6. Cutpoints of mild (0 to 3.9), moderate (4 to 6.9) and severe (7 to 10) on both the HFRDIS and HFI were significantly and positively associated with symptom measures. More hot flash interference reflected worse VMS diary frequency, severity and bother as well as worse PSQI sleep, insomnia severity, and anxiety. The associations supported the cutpoint values for both scales.

MIDs

Table 7 displays MID results for the HFRDIS and HFI. Overall, the average MID for the HFRDIS was 1.66 and for the HFI was 2.34.

DISCUSSION

Hot flash related daily interference is an important patient-reported outcome with the HFRDIS measure having been widely used internationally. This is the first report to evaluate a shortened version of the HFRDIS, to evaluate cutpoints, and establish MIDs for both the HFRDIS and condensed HFI. Data from the 10-item HFRDIS showed good fit to a single hot flash interference factor, suggesting that the single-factor conceptualization of the 3-item HFI was reasonable. The HFI had high internal consistency reliability (Cronbach's alphas) and demonstrated strong construct and convergent validity (correlations). In addition, cutpoints for mild, moderate, and severe interference on both the HFRDIS and HFI were associated with VMS diary, sleep, and anxiety symptoms. Finally, we identified MIDs for the HFRDIS and HFI, which may be used in future power analyses to determine appropriate sample sizes for clinical trials and in interpreting responsiveness to change and sensitivity to clinical treatments and other intervention effects over time.

We did not find a single item that was adequate for measuring interference, despite this being a possibility at the start of our analyses. We did not make a priori assumptions regarding the number of items to be retained in the shortened version. Our analytic methods also did not preclude this possibility. Sloan et al.[43] present an excellent discussion of the relative advantages of single item vs. multi item scales. Single-item scales are easy to administer, reduce response burden, and can be psychometrically sound with ability to demonstrate validity and change over time, including in response to treatment effects.[43] Multi-item scales can be useful for multidimensional latent constructs, can reduce measurement error, can be scored to handle missing item data, and can improve reliability, validity, responsiveness to change over time, and sensitivity to intervention effects.[43] Our three-item HFI has the advantages of a multi-item scale and because of its very small number of items, also has other advantages similar to single item scales such as being easy to administer, with minimal response burden, and demonstrated reliability, validity, and responsiveness to intervention effects.

The three retained HFI items received the highest ratings from experts and were the least likely to receive participant ratings of zero or no interference. The links between frequent, severe, or bothersome VMS and sleep problems,[44, 45] mood problems,[46–49] and concentration difficulties[50, 51] have all been previously documented and are consistent with two previously published MsFLASH reports: MsFLASH trial participants' symptom priorities[52] and symptom clusters.[53] In MsFLASH 02, a companion study was added and it required trial participants to complete a card sort task to indicate which three of twelve possible symptoms they would most like to alleviate.[52] The four most highly rated symptoms were VMS, sleep disruption, concentration impairment, and fatigue. Mood was endorsed much less frequently in this analysis possibly because women with mood disturbances were excluded from the trial. Similarly, using pooled data across the MsFLASH trials (n=899), five classes of symptom clusters were found.[53] The classes were: (1) hot flash interference, sleep problems, and pain, (2) hot flash interference, sleep problems, mood problems, and pain, (3) hot flash interference and sleep problems, (4) hot flash interference and mood problems, and (5) low severity of all symptoms. Problems with concentration were not measured as a MsFLASH outcome for use in the clusters analysis. In general, these studies substantiate the inclusion of sleep, mood, and concentration problems as the symptoms most closely associated with hot flashes or hot flash interference.

Published findings from one study support the exclusion of the 'relationships with others' item but not the interference with 'work' item.[54] In an analysis of data from the Seattle Midlife Women's Health study, multilevel modeling was used to test correlates of symptom interference. Interference with relationships and work were each measured with diary questions over time and 0 (not at all) to 6 (extremely or a lot) response options. Hot flash severity was measured over time and rated from 0 (not present) to 4 (extreme). Interference with relationships was not significantly associated with hot flash severity, but was associated with other variables including mood, sleep symptoms, and forgetfulness/difficulty concentrating. Interference with work was significantly associated with hot flash severity, depressed mood, difficulty getting to sleep, forgetfulness/difficulty concentrating, and other variables. Differences between our study and this published study might have been related to the longitudinal nature of the published study or differences between measurement tools. However, in both our study and this published study, hot flashes, sleep, and difficulty concentrating did cluster together.

The values we identified for the MIDs can be compared to another MID for a similarly scored symptom interference scale. Using pooled data from over 2000 patients with nerve pain from diabetic neuropathy or postherpetic neuralgia, a 1 to 2 point change on a 0 to 10 point sleep interference mean score was found to be an appropriate MID.[55] This MID range for sleep interference is very similar to our MIDs of 1.56 to 2.57 for hot flash interference.

Study findings should be considered in light of some limitations. First, although MsFLASH 01 trial participants were diverse in terms of race/ethnicity, participants of the remaining two trials' were less diverse. All participants were American and English speaking. Therefore, whether our findings will hold true if more diverse samples or translated versions of the scale are included remains to be determined. Second, MsFLASH trials excluded women with mood disturbances, which may have affected the results reported here. Third, we

acknowledge the data driven approach to establishing cutpoints as an alternative method that might have yielded different results.[33–37] Fourth, MIDs can vary among different populations[56] and additional studies may yield different findings.

Future research implications are based on study findings and limitations and include the following. It will be important to replicate our findings in other populations and using different translations of the HFRDIS to determine whether the 3-item solution and MIDs hold across populations, including those with mood disturbances, and across different instrument translations. While the three retained items for the HFI were those most relevant to this American sample, it is possible that different items or a different number of items might perform best in non-American or non-English speaking populations. It is also possible that MIDs may vary by population or with different translated versions. In addition, a future study could use a data-driven approach to establish severity-based cutpoints for hot flash interference. Methods could follow those outlined by Jeon et al.[33] Using data from two studies involving cancer patients assessed at multiple time points, they used 16 different symptom severity ratings to evaluate cutpoints for the corresponding 16 different symptom interference ratings. Interference cutpoints varied across the 16 symptoms. The most common cutpoints (4 symptoms, 25%) followed the categories we used (1-3, 4-6, and 7-10) and the second most common cutpoints (3 symptoms, 19%) used categories of 1, 2-4, and 5-10. All other symptoms used different combinations of cutpoints. We encourage additional analyses by other investigators and clinicians who have used the HFRDIS.

CONCLUSION

In conclusion, we successfully created a shorter HFI scale that will be useful in busy clinical practices to effectively monitor hot flash interference and response to treatment over time. Moreover, our findings may help research studies where response burden may be an issue. The scale cutpoints and MIDs for the HFRDIS and HFI should prove useful in interpreting the existing as well as future research. We urge clinicians and researchers to consider assessing hot flash interference in midlife women as an important patient-reported outcome and to conduct similar psychometric analyses using existing data from more diverse populations and/or with translated versions of the HFRDIS.

Acknowledgments

The network sites that participated in these studies included Boston, MA (Massachusetts General Hospital and Brigham and Women's Hospital; Principal Investigators: Lee Cohen, MD and Hadine Joffe, MD, MSc); Indianapolis, IN (Indiana University; Principal Investigator: Janet S Carpenter, PhD, RN, FAAN); Oakland, CA (Kaiser Permanente Division of Research; Principal Investigators: Barbara Sternfeld, PhD and Bette Caan, PhD; Philadelphia, PA (University of Pennsylvania; Principal Investigator: Ellen W. Freeman, PhD); Seattle, WA (Group Health Research Institute; Principal Investigators: Katherine Newton, PhD and Susan Reed, MD). The Data Coordinating Center of the network is based at the Fred Hutchinson Cancer Research Center; Principal Investigators: Andrea LaCroix, PhD and Katherine A. Guthrie, PhD (Garnet L. Anderson, PhD, was previously a PI) and Statistician: Joseph C. Larson, MS. The chairperson is Kris E. Ensrud, MD, University of Minnesota. The project coordinator is Judy Hannah, PhD: National Institute on Aging/US National Institutes of Health, Bethesda, MD. We thank Nirmala Xavier, PhD for her editorial assistance.

Sources of support: The MsFLASH studies were supported by a cooperative agreement issued by the National Institute of Aging (NIA), in collaboration with the Eunice Kennedy Shriver National Institute of Child Health and Development (NICHD), the National Center for Complementary and Alternative Medicine (NCCAM) and the Office of Research and Women's Health (ORWH), and NIA grants U01AG032659, U01AG032669,

U01AG032682, U01AG032699, and U01AG032700. In Indiana, the project was supported by the Indiana Clinical and Translational Sciences Institute, funded in part by grant UL1 RR025761 from the National Institutes of Health, National Center for Research Resources, Clinical, and Translational Sciences Award. One author (CXC) was supported by the National Institute of Nursing Research of the National Institutes of Health under Award Number T32NR007066. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Carpenter JS. The Hot Flash Related Daily Interference Scale: a tool for assessing the impact of hot flashes on quality of life following breast cancer. J Pain Symptom Manage. 2001; 22:979–989. [PubMed: 11738160]
2. Daut RL, Cleeland CS, Flanery RC. Development of the Wisconsin brief pain questionnaire to assess pain in cancer and other diseases. Pain. 1983; 17:197–210. [PubMed: 6646795]
3. Hann DM, Jacobsen PB, Azzarello LM, Martin SC, Curran SL, Fields KK, Greenberg H, Lyman G. Measurement of fatigue in cancer patients: development and validation of the Fatigue Symptom Inventory. Qual Life Res. 1998; 7:301–310. [PubMed: 9610214]
4. PDQ® Supportive and Palliative Care Editorial Board. PDQ Hot Flashes and Night Sweats. Bethesda, MD: National Cancer Institute; Available at: https://www.cancer.gov/about-cancer/treatment/side-effects/sexuality-fertility-women/hot-flashes-hp-pdq#section/_209. Accessed November 11, 2016
5. Carpenter JS, Storniolo AM, Johns S, Monahan PO, Azzouz F, Elam JL, Johnson CS, Shelton RC. Randomized, double-blind, placebo-controlled crossover trials of venlafaxine for hot flashes after breast cancer. Oncologist. 2007; 12:124–135. [PubMed: 17227907]
6. Elkins G, Marcus J, Stearns V, Perfect M, Rajab MH, Ruud C, Palamara L, Keith T. Randomized trial of a hypnosis intervention for treatment of hot flashes among breast cancer survivors. J Clin Oncol. 2008; 26:5022–5026. [PubMed: 18809612]
7. Barton DL, LaVasseur BI, Sloan JA, Stawis AN, Flynn KA, Dyar M, Johnson DB, Atherton PJ, Diekmann B, Loprinzi CL. Phase III, placebo-controlled trial of three doses of citalopram for the treatment of hot flashes: NCCTG trial N05C9. J Clin Oncol. 2010; 28:3278–3283. [PubMed: 20498389]
8. Soares CN, Frey BN, Haber E, Steiner M. A pilot, 8-week, placebo lead-in trial of quetiapine extended release for depression in midlife women: Impact on mood and menopause-related symptoms. J Clin Psychopharmacol. 2010; 30:612–615. [PubMed: 20814317]
9. Freeman MP, Hibbeln JR, Silver M, Hirschberg AM, Wang B, Yule AM, Petrillo LF, Pascuillo E, Economou NI, Joffe H, Cohen LS. Omega-3 fatty acids for major depressive disorder associated with the menopausal transition: A preliminary open trial. Menopause. 2011; 18:279–84. [PubMed: 21037490]
10. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR. Methods to explain the clinical significance of health status measures. Mayo Clin Proc. 2002; 77:371–83. [PubMed: 11936935]
11. Nunnaly, JC., Bernstein, IH., editors. Psychometric Theory. Third. New York, NY: McGraw-Hill; 1994.
12. Newton KM, Carpenter JS, Guthrie KA, Anderson GL, Caan B, Cohen LS, Ensrud KE, Freeman EW, Joffe H, Sternfeld B, Reed SD, Sherman S, Sammel MD, Kroenke K, Larson JC, Lacroix AZ. Methods for the design of vasomotor symptom trials: the Menopausal Strategies: Finding Lasting Answers to Symptoms and Health network. Menopause. 2014; 21:45–58. [PubMed: 23760428]
13. Sternfeld B, Lacroix A, Caan BJ, Dunn AL, Newton KM, Reed SD, Guthrie KA, Booth-Laforce C, Sherman KJ, Cohen L, Freeman MP, Carpenter JS, Hunt JR, Roberts M, Ensrud KE. Design and methods of a multi-site, multi-behavioral treatment trial for menopausal symptoms: The MsFLASH experience. Contemp Clin Trials. 2013; 35:25–34. [PubMed: 23462342]
14. Freeman EW, Guthrie KA, Caan B, Sternfeld B, Cohen LS, Joffe H, Carpenter JS, Anderson GL, Larson JC, Ensrud KE, Reed SD, Newton KM, Sherman S, Sammel MD, LaCroix AZ. Efficacy of escitalopram for hot flashes in healthy menopausal women: A randomized controlled trial. JAMA. 2011; 305:267–274. [PubMed: 21245182]

15. Cohen LS, Joffe H, Guthrie KA, Ensrud KE, Freeman M, Carpenter JS, Learman LA, Newton KM, Reed SD, Manson JE, Sternfeld B, Caan B, Freeman EW, Lacroix AZ, Tinker LF, Laforce CB, Larson JC, Anderson GL. Efficacy of omega-3 for vasomotor symptoms treatment: A randomized controlled trial. Menopause. 2014; 21:347–354. [PubMed: 23982113]
16. Newton KM, Reed SD, Guthrie KA, Sherman KJ, Booth-Laforce C, Caan B, Sternfeld B, Carpenter JS, Learman LA, Freeman EW, Cohen LS, Joffe H, Anderson GL, Larson JC, Hunt JR, Ensrud KE, Lacroix AZ. Efficacy of yoga for vasomotor symptoms: a randomized controlled trial. Menopause. 2014; 21:339–346. [PubMed: 24045673]
17. Sternfeld B, Guthrie KA, Ensrud KE, Lacroix AZ, Larson JC, Dunn AL, Anderson GL, Seguin RA, Carpenter JS, Newton KM, Reed SD, Freeman EW, Cohen LS, Joffe H, Roberts M, Caan BJ. Efficacy of exercise for menopausal symptoms: a randomized controlled trial. Menopause. 2014; 21:330–338. [PubMed: 23899828]
18. Joffe H, Guthrie KA, LaCroix AZ, Reed SD, Ensrud KE, Manson JE, Newton KM, Freeman EW, Anderson GL, Larson JC, Hunt J, Shifren J, Rexrode KM, Caan B, Sternfeld B, Carpenter JS, Cohen L. Low-dose estradiol and the serotonin-norepinephrine reuptake inhibitor venlafaxine for vasomotor symptoms: a randomized clinical trial. JAMA Intern Med. 2014; 174:1058–1066. [PubMed: 24861828]
19. Hilditch JR, Lewis J, Peter A, van Maris B, Ross A, Franssen E, Guyatt GH, Norton PG, Dunn E. A menopause-specific quality of life questionnaire: development and psychometric properties. Maturitas. 1996; 24:161–175. [PubMed: 8844630]
20. LaCroix AZ, Freeman EW, Larson J, Carpenter JS, Joffe H, Reed SD, Newton KM, Seguin RA, Sternfeld B, Cohen L, Ensrud KE. Effects of escitalopram on menopause-specific quality of life and pain in healthy menopausal women with hot flashes: a randomized controlled trial. Maturitas. 2012; 73:361–368. [PubMed: 23031421]
21. Buysse DJ, Reynolds CF 3rd, Monk TH, Berman SR, Kupfer DJ. The pittsburgh sleep quality index: A new instrument for psychiatric practice and research. Psychiatry Res. 1989; 28:193–213. [PubMed: 2748771]
22. Buysse DJ, Reynolds CF 3rd, Monk TH, Hoch CC, Yeager AL, Kupfer DJ. Quantification of subjective sleep quality in healthy elderly men and women using the pittsburgh sleep quality index (PSQI). Sleep. 1991; 14:331–338. [PubMed: 1947597]
23. Carpenter JS, Andrykowski MA. Psychometric evaluation of the pittsburgh sleep quality index. J Psychosom Res. 1998; 45:5–13. [PubMed: 9720850]
24. Otte JL, Rand KL, Landis CA, Paudel ML, Newton KM, Woods N, Carpenter JS. Confirmatory factor analysis of the pittsburgh sleep quality index in women with hot flashes. Menopause. 2015; 22:1190–1196. [PubMed: 25944520]
25. Buysse DJ, Germain A, Moul DE, Franzen PL, Brar LK, Fletcher ME, Begley A, Houck PR, Mazumdar S, Reynolds CF, Monk TH. Efficacy of brief behavioral treatment for chronic insomnia in older adults. Arch Intern Med. 2011; 171:887–895. [PubMed: 21263078]
26. Morin CM, Belleville G, Belanger L, Ivers H. The insomnia severity index: psychometric indicators to detect insomnia cases and evaluate treatment response. Sleep. 2011; 34:601–608. [PubMed: 21532953]
27. Ensrud KE, Guthrie KA, Hohensee C, Caan B, Carpenter JS, Freeman EW, LaCroix AZ, Landis CA, Manson J, Newton KM, Otte J, Reed SD, Shifren JL, Sternfeld B, Woods NF, Joffe H. Effects of estradiol and venlafaxine on insomnia symptoms and sleep quality in women with hot flashes. Sleep. 2015; 38:97–108. [PubMed: 25325454]
28. Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med. 2006; 166:1092–1097. [PubMed: 16717171]
29. Kroenke K, Spitzer RL, Williams JB, Monahan PO, Lowe B. Anxiety disorders in primary care: Prevalence, impairment, comorbidity, and detection. Ann Intern Med. 2007; 146:317–25. [PubMed: 17339617]
30. Wynd CA, Schmidt B, Schaefer MA. Two quantitative approaches for estimating content validity. West J Nurs Res. 2003; 25:508–18. [PubMed: 12955968]
31. Wilson FR, Pan W, Schumsky DA. Recalculation of the critical values for lawshe's content validity ratio. Measure Eval Counsel Devel. 2012; 45:197–210.

32. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct Equat Mod. 1999; 6:1–55.
33. Jeon S, Given CW, Sikorskii A, Given B. Do interference-based cut-points differentiate mild, moderate, and severe levels of 16 cancer-related symptoms over time? J Pain Symptom Manage. 2009; 37:220–232. [PubMed: 18619769]
34. Boonstra AM, Stewart RE, Koke AJ, Oosterwijk RF, Swaan JL, Schreurs KM, Schiphorst Preuper HR. Cut-Off Points for Mild, Moderate, and Severe Pain on the Numeric Rating Scale for Pain in Patients with Chronic Musculoskeletal Pain: Variability and Influence of Sex and Catastrophizing. Front Psychol. 2016; 7:1466–1474. [PubMed: 27746750]
35. Li KK, Harris K, Hadi S, Chow E. What should be the optimal cut points for mild, moderate, and severe pain? J Palliat Med. 2007; 10:1338–1346. [PubMed: 18095813]
36. Serlin RC, Mendoza TR, Nakamura Y, Edwards KR, Cleeland CS. When is cancer pain mild, moderate or severe? Grading pain severity by its interference with function. Pain. 1995; 61:277–284. [PubMed: 7659438]
37. Wang XS, Zhao F, Fisch MJ, O'Mara AM, Cella D, Mendoza TR, Cleeland CS. Prevalence and characteristics of moderate to severe fatigue: a multicenter study in cancer patients and survivors. Cancer. 2014; 120:425–432. [PubMed: 24436136]
38. Butt DA, Deng LYR, Lewis JE, Lock M. Minimal decrease in hot flashes desired by postmenopausal women in family practice. Menopause. 2007; 14:203–207. [PubMed: 17099324]
39. Kline, RB. Principles and practice of structural equation modeling. Third. New York, NY: Guilford; 2011.
40. Jackson DL. Revisiting sample size and number of parameter estimates: some support for the N:q hypothesis. Struct Equat Model. 2003; 10:128–141.
41. Bentler PM, Chou CP. Practical issues in structural equation modeling. Socio Methods & Res. 1987; 16:78–117.
42. Rosseel Y. Lavaan: an R package for structural equation modeling. J Stat Software. 2012; 48:1–36.
43. Sloan JA, Aaronson N, Cappelleri JC, Fairclough DL, Varricchio C, Clinical Significance Consensus Meeting G. Assessing the clinical significance of single items relative to summated scores. Mayo Clin Proc. 2002; 77:479–487. [PubMed: 12004998]
44. Xu H, Thurston RC, Matthews KA, Bryce CL, Hays RD, Kapoor WN, Ness RB, Hess R. Are hot flashes associated with sleep disturbance during midlife? Results from the STRIDE cohort study. Maturitas. 2012; 71:34–38. [PubMed: 22051577]
45. Cray LA, Woods NF, Herting JR, Mitchell ES. Symptom clusters during the late reproductive stage through the early postmenopause: observations from the Seattle Midlife Women's Health Study. Menopause. 2012; 19:864–869. [PubMed: 22643229]
46. Seritan AL, Iosif AM, Park JH, DeatherageHand D, Sweet RL, Gold EB. Self-reported anxiety, depressive, and vasomotor symptoms: a study of perimenopausal women presenting to a specialized midlife assessment center. Menopause. 2010; 17:410–415. [PubMed: 20216277]
47. Freeman EW. Associations of depression with the transition to menopause. Menopause. 2010; 17:823–827. [PubMed: 20531231]
48. Cray L, Woods NF, Mitchell ES. Symptom clusters during the late menopausal transition stage: observations from the Seattle Midlife Women's Health Study. Menopause. 2010; 17:972–977. [PubMed: 20628322]
49. Burleson MH, Todd M, Trevathan WR. Daily vasomotor symptoms, sleep problems, and mood: using daily data to evaluate the domino hypothesis in middle-aged women. Menopause. 2010; 17:87–95. [PubMed: 19675506]
50. Maki PM, Drogos LL, Rubin LH, Banuvar S, Shulman LP, Geller SE. Objective hot flashes are negatively related to verbal memory performance in midlife women. Menopause. 2008; 15:848–856. [PubMed: 18562950]
51. Sliwinski JR, Johnson AK, Elkins GR. Memory decline in peri- and post-menopausal women: the potential of mind-body medicine to improve cognitive performance. Integr Med Insights. 2014; 9:17–23. [PubMed: 25125972]

52. Carpenter JS, Woods NF, Otte JL, Guthrie KA, Hohensee C, Newton KM, Joffe H, Cohen L, Sternfeld B, Lau RJ, Reed SD, LaCroix AZ. MsFLASH participants' priorities for alleviating menopausal symptoms. Climacteric. 2015; 18:859–866. [PubMed: 26517583]

53. Woods NF, Hohensee C, Carpenter JS, Cohen L, Ensrud K, Freeman EW, Guthrie KA, Joffe H, LaCroix AZ, Otte JL. Symptom clusters among MsFLASH clinical trial participants. Menopause. 2016; 23:158–165. [PubMed: 26506500]

54. Woods NF, Mitchell ES. Symptom interference with work and relationships during the menopausal transition and early postmenopause: observations from the Seattle Midlife Women's Health Study. Menopause. 2011; 18:654–661. [PubMed: 21317821]

55. Vernon MK, Brandenburg NA, Alvir JM, Griesing T, Revicki DA. Reliability, validity, and responsiveness of the daily sleep interference scale among diabetic peripheral neuropathy and postherpetic neuralgia patients. J Pain Symptom Manage. 2008; 36:54–68. [PubMed: 18411009]

56. Revicki DA, Cella D, Hays RD, Sloan JA, Lenderking WR, Aaronson NK. Responsiveness and minimal important differences for patient reported outcomes. Health Qual Life Outcomes. 2006; 4:70–75. [PubMed: 17005038]

Table 1

Baseline Demographic and Clinical Characteristics of the three MsFLASH Trials

	MsFLASH 01 (n=205) n (%)	MsFLASH 02 (n=355) n (%)	MsFLASH 03 (n=339) n (%)	p
Age at screening, median (IQR)	54.0 (51.0, 57.0)	54.0 (52.0, 57.0)	54.0 (52.0, 57.0)	0.062
<50	24 (11.7)	19 (5.4)	30 (8.8)	0.189
50-54	95 (46.3)	162 (45.6)	147 (43.4)	
55-59	66 (32.2)	130 (36.6)	123 (36.3)	
≥60	20 (9.8)	44 (12.4)	39 (11.5)	
Ethnicity				<0.001
African-American	89 (43.4)	88 (24.9)	111 (32.8)	
White	99 (48.3)	228 (64.6)	203 (60.1)	
Other	17 (8.3)	37 (10.5)	24 (7.1)	
Education				<0.001
≤High school diploma or GED	38 (18.5)	21 (5.9)	55 (16.3)	
Post-high school	87 (42.4)	112 (31.6)	111 (32.8)	
College graduate	80 (39.0)	221 (62.4)	172 (50.9)	
Smoking				<0.001
Never	99 (48.3)	233 (65.8)	175 (51.9)	
Past	59 (28.8)	89 (25.1)	107 (31.8)	
Current	47 (22.9)	32 (9.0)	55 (16.3)	
Body mass index, median (IQR)	28.3 (24.5, 32.2)	26.9 (23.6, 29.9)	27.0 (23.7, 31.2)	0.007
<25	54 (26.5)	123 (34.6)	118 (35.5)	0.005
25-29.9	72 (35.3)	144 (40.6)	107 (32.2)	
≥30	78 (38.2)	88 (24.8)	107 (32.2)	
Menopause status				0.27
Postmenopausal	142 (69)	266 (75)	256 (76)	
Perimenopausal	41 (20)	65 (18)	52 (15)	
Indeterminate	22 (11)	24 (7)	31 (9)	
Clinical Site				<0.001
Boston	43 (21.0)	0 (0.0)	100 (29.5)	
Indianapolis	35 (17.1)	118 (33.2)	0 (0.0)	
Oakland	57 (27.8)	110 (31.0)	0 (0.0)	
Philadelphia	70 (34.1)	0 (0.0)	121 (35.7)	
Seattle	0 (0.0)	127 (35.8)	118 (34.8)	
VMS[a] per day, median (IQR)				
Frequency	8.4 (6.4, 11.1)	6.9 (4.9, 9.9)	6.5 (4.9, 9.8)	<0.001
Severity	1.1 (0.8, 1.5)	1.0 (0.7, 1.3)	1.0 (0.7, 1.3)	<0.001
Bother[b]	2.1 (1.8, 2.5)	2.0 (1.6, 2.3)	2.0 (1.6, 2.3)	<0.001
HFRDIS, median (IQR)	3.5 (2.0, 5.3)	2.7 (1.5, 4.9)	2.9 (1.7, 5.2)	0.037

IQR, interquartile range.

[a]VMS, vasomotor symptoms, VMS severity was measured from 0 (mild), 1 (moderate) or 2 (severe).

[b]VMS bother was measured from 0 (not at all), 1 (a little), 2 (moderately), or 3 (a lot).

Table 2

MsFLASH Experts' Content Validity Ratios (CVRs) and MsFLASH Participants' HFRDIS Item Distributions

	MsFLASH Experts CVRs[a] (n=7)	MsFLASH Participants (n=899) n (%) Endorsing Zero (No) Interference[b]	Median (IQR)
Work	−0.143	242 (27.3)	2.0 (0.0, 4.0)
Social activities	−0.429	311 (34.8)	1.0 (0.0, 3.0)
Leisure activities	−1.000	274 (30.8)	1.0 (0.0, 4.0)
Sleep[c]	1.000	20 (2.2)	5.0 (2.0, 8.0)
Mood[c]	1.000	157 (17.6)	2.0 (0.0, 5.0)
Concentration[c]	0.714	161 (18.1)	2.0 (0.0, 5.0)
Relations with others	−0.143	337 (37.9)	1.0 (0.0, 3.0)
Sexuality	−0.143	288 (32.8)	1.0 (0.0, 5.0)
Enjoyment of life	0.429	221 (24.8)	2.0 (0.0, 4.0)
Life satisfaction	0.429	196 (22.0)	2.0 (0.0, 4.0)

[a]CVR, content validity ratio, cutpoint of 0.622 was used to identify items to retain in HFI.

[b]Actual and possible range for all items was 0 (no) to 10 (completely) interference.

[c]Item was retained in the short scale.

Table 3

HFRDIS and HFI Cronbach's Alpha Coefficients by Study Population and Time Points

	Combined Sample (n=899)	MsFLASH 01 (n=205)	MsFLASH 02 (n=355)	MsFLASH 03 (n=339)
HFRDIS				
Baseline	0.932	0.938	0.929	0.930
Post-randomization	0.942	0.945	0.945	0.927
HFI				
Baseline	0.830	0.827	0.837	0.823
Post-randomization	0.856	0.859	0.830	0.882

Table 4

HFRDIS Confirmatory Factor Analyses Fit Indices by Study Population

	Combined Sample (n=899)	MsFLASH 01 (n=205)	MsFLASH 02 (n=355)	MsFLASH 03 (n=339)
HFRDIS				
Comparative Fit Index	0.988	0.989	0.987	0.986
Tucker-Lewis Index	0.985	0.986	0.983	0.983
SRMR[a]	0.061	0.057	0.066	0.063

[a]SRMR, Standardized Root Mean Square Residual.

Table 5

Spearman's Rank Correlation Coefficients among HFRDIS, HFI, and Other VMS Measures

	HFRDIS		HFI	
	Baseline	**Post-randomization**	**Baseline**	**Post-randomization**
HFRDIS[a]	1	0.944*	0.907*	1
HFI[b]	0.907*	1	1	0.944*
VMS[c] frequency	0.179*	0.287*	0.183*	0.299*
Daytime	0.132*	0.480*	0.118*	0.504*
Nighttime	0.218*	0.479*	0.243*	0.465*
VMS[c] severity	0.369*	0.469*	0.363*	0.471*
Daytime	0.345*	0.437*	0.309*	0.466*
Nighttime	0.333*	0.448*	0.354*	0.438*
VMS[c] bother	0.397*	0.485*	0.401*	0.483*
Daytime	0.368*	0.439*	0.342*	0.480*
Nighttime	0.360*	0.454*	0.392*	0.445*
MENQOL[d]	0.546*	0.588*	0.540*	0.582*

*p-value < 0.001.

[a]HFRDIS, Hot Flash Related Daily Interference Scale.

[b]HFI, Hot Flash Interference scale.

[c]VMS, vasomotor symptoms.

[d]MENQOL, Menopausal Quality of Life Scale.

Table 6

HFRDIS and HFI Cutpoints in Relation to Other Symptom Measures

	HFRDIS[a]				HFI[b]			
VMS diary (Median, IQR)[c]	**Mild (0-3.9)**	**Moderate (4-6.9)**	**Severe (7-10)**	**p**	**Mild (0-3.9)**	**Moderate (4-6.9)**	**Severe (7-10)**	**p**
VMS frequency	6.8 (4.8, 9.6)	7.4 (5.6, 10.5)	9.1 (5.6, 12.6)	<0.001	6.6 (4.7, 9.2)	7.4 (5.6, 10.2)	7.6 (5.5, 11.5)	<0.001
VMS severity	0.9 (0.6, 1.1)	1.1 (0.9, 1.4)	1.6 (1.2, 1.9)	<0.001	0.9 (0.6, 1.1)	1.0 (0.7, 1.3)	1.3 (1.0, 1.7)	<0.001
VMS bother	1.8 (1.5, 2.1)	2.1 (1.8, 2.5)	2.6 (2.2, 3.0)	<0.001	1.8 (1.5, 2.1)	2.0 (1.7, 2.3)	2.3 (2.0, 2.7)	<0.001
Sleep and Anxiety (n, %)[d]								
PSQI[e] Sleep Score				<0.001 (0.255)				<0.001 (0.308)
No problems (0-4)	124 (83.8)	22 (14.9)	2 (1.4)		105 (67.3)	41 (26.3)	10 (6.4)	
Sleep problems (5-7)	184 (69.2)	66 (24.8)	16 (6.0)		128 (47.4)	102 (37.8)	40 (14.8)	
Sleep and fatigue (8+)	206 (50.2)	150 (36.6)	54 (13.2)		116 (27.1)	171 (40.0)	141 (32.9)	
ISI[f] Sleep Score				<0.001 (0.327)				<0.001 (0.441)
No insomnia (0-7)	202 (83.8)	34 (14.1)	5 (2.1)		175 (70.6)	60 (24.2)	13 (5.2)	
Mild (8-14)	226 (66.3)	96 (28.2)	19 (5.6)		139 (39.4)	153 (43.3)	61 (17.3)	
Moderate (15-21)	89 (40.1)	100 (45.0)	33 (14.9)		40 (17.0)	101 (43.0)	94 (40.0)	
Severe (22-28)	6 (15.8)	16 (42.1)	16 (42.1)		2 (5.1)	5 (12.8)	32 (82.1)	
GAD[g]-7 Anxiety Score				<0.001 (0.213)				<0.001 (0.236)
Mild (0-5)	412 (68.3)	150 (24.9)	41 (6.8)		295 (47.5)	217 (34.9)	109 (17.6)	
Moderate (6-10)	28 (63.6)	11 (25.0)	5 (11.4)		20 (43.5)	14 (30.4)	12 (26.1)	
Severe (11+)	86 (42.0)	88 (42.9)	31 (15.1)		44 (20.2)	91 (41.7)	83 (38.1)	

For all scales, higher scores equal poorer outcomes.

[a]HFRDIS, Hot Flash Related Daily Interference Scale.

[b]HFI, Hot Flash Interference scale.

[c]IQR, Interquartile range, p values based on Kruskall-Wallis tests.

[d]p values based on Kendall's tau coefficients, Tau shown in parentheses in p column.

[e]PSQI, Pittsburgh Sleep Quality Index.

[f] ISI, Insomnia Severity Index.

[g] GAD, Generalized Anxiety Disorder.

Table 7

Minimally Important Differences (MID) for HFRDIS and HFI Scores

	N (%)	HFRDIS[a] MID Mean (SD)	HFI[b] MID Mean (SD)
40%-60% reduction in total VMS[c] frequency	153 (17.35)	−1.76 (1.97)	−2.57 (2.29)
40%-60% reduction in daytime VMS frequency	132 (14.97)	−1.58 (1.67)	−2.24 (2.33)
40%-60% reduction in nighttime VMS frequency	112 (12.70)	−1.76 (1.90)	−2.38 (2.19)
0.5-1.5 reduction in total MENQOL[d] score	291 (32.99)	−1.60 (1.85)	−2.26 (2.24)
Average of all of the column MIDs		−1.66	−2.34

[a]HFRDIS, Hot Flash Related Daily Interference Scale.

[b]HFI, Hot Flash Interference scale.

[c]VMS, vasomotor symptoms.

[d]MENQOL, Menopausal Quality of Life Scale.