http://dx.doi.org/10.1080/13697137.2019.1580258
A comparison of three outcome measures of the impact of vasomotor symptoms on women’s lives

Running title: Comparison of measures of impact of VMS

Myra S. Hunter, PhD,¹ Jacqui Nuttall, MSc,² Deborah Fenlon, PhD³

¹ Institute of Psychiatry, Psychology and Neuroscience, Kings College London, Guy’s Campus, London, UK.
² Southampton Clinical Trials Unit, University of Southampton, Southampton, UK
³ Dept of Nursing, Swansea University, Swansea, Wales

Funding/support: Menos 4 study was supported by project grant from Breast Cancer Now (ref: 2015CR_004).

Financial disclosure/conflicts of interest: The authors have no financial or conflicts of interest to disclose.

Address correspondence to: Prof Myra S. Hunter, Institute of Psychiatry, Psychology and Neuroscience, Kings College London, 5th Floor Bermondsey Wing, Guy’s Campus, London SE1 9RT. E-mail: myra.hunter@kcl.ac.uk
Abstract

Objective: Measures of the impact of vasomotor symptoms (VMS) have been used as outcomes in clinical trials but have not been compared. This study compares the Hot Flush Rating Scale (HFRS), the Hot Flash Related Daily Interference Scale (HFRDIS) and the shorter Hot Flash Interference (HFI) scale.

Methods: Baseline data from two studies included healthy women (menopause transition or postmenopause) and breast cancer patients, experiencing VMS. Participants completed questionnaires (sociodemographics, HFRS, HFRDIS, HFI, Work and Social Adjustment Scale (WSAS), depression (GAD7), anxiety (PHQ9) and use of medical services.

Results: 169 women (129 with history of breast cancer and 40 without), aged 54.47 (SD=9.11) took part. They had an average of 66 (SD=40.94) VMS per week, with mean HFRS problem rating of 6.53 (SD=1.99), HFRDIS score of 5.36 (SD=2.22) and HFI score of 6.13 (SD=2.30). HFRS problem-rating, HFRDIS and HFI were significantly associated (r=0.61-0.85), had good internal reliability (alpha=0.76-0.91) and significant concurrent validity with mood, WSAS and use of medical services. VMS frequency was not associated with mood, WSAS nor use of medical services.

Conclusion: The HFRS Problem-rating scale and the HFI are two brief, three-item measures that measure a similar concept of VMS interference/impact, with evidence of reliability and validity.

Key words: Hot flashes, flushes, vasomotor symptoms, impact, interference, measures, menopause
Health-related quality of life of women going through the menopause transition and post menopause, appears to be influenced by vasomotor symptoms (VMS), to a greater extent than stage of menopause per se.\textsuperscript{1-4} Moreover, it tends to be the impact of VMS, i.e. how much they interfere with life or how problematic they are perceived to be, rather than their frequency, that is associated with aspects of quality of life \textsuperscript{5,5-7}. Consequently, these variables have been considered as appropriate outcomes in clinical trials of treatments of VMS.\textsuperscript{8-18} These patient reported outcomes are particularly relevant to evaluation of behavioural treatments which tend to target coping strategies, cognitive appraisal and functioning\textsuperscript{9,12-18} For example, following a cognitive behaviour therapy (CBT) intervention for VMS, some participants reported that they still had VMS but they hardly noticed them.\textsuperscript{19} Measures of impact include: the \textit{Hot Flush Rating Scale} (HFRS)\textsuperscript{20}, the \textit{Hot Flash Related Daily Interference Scale} (HFRDIS)\textsuperscript{21} and the \textit{Hot Flash Interference} (HFI).\textsuperscript{22} HFRS is a self-report measure of VMS frequency and problem-rating over the past week. The \textit{Hot Flush Rating Scale} (HFRS) problem-rating is calculated as the mean of the scores on three Likert scales (scores range from 1–10) assessing the extent to which HFNS are problematic, distressing and causing interference in daily life. The \textit{Hot Flash Related Daily Interference Scale} (HFRDIS)\textsuperscript{21}, a 10-item questionnaire, assesses the impact of hot flushes on daily activities and quality of life in the past week, and the \textit{Hot Flash Interference} (HFI)\textsuperscript{22} measure is a 3 item shortened version of the HFRDIS. The HFRS problem-rating scale tends to be used in UK and European trials, while the HFRDIS tends to be used in the USA, but they have been not been directly compared.

The current study aims to compare these three measures and to examine their interrelationships and concurrent validity in relation to work and social adjustment, anxiety, depressed mood and use of medical services (doctor visits).
METHODS

Participants

Women with variable menstrual cycles (menopause transition) or who were more than one year from their last menstrual period (post-menopause) who were having VMS were recruited from two sources: (i) baseline data from women who took part in an unpublished student project investigating attentional bias amongst women with VMS recruited via online advertisement (referred to as ‘healthy women’ to distinguish from the breast cancer sample), and (ii) baseline data from a multicentre trial of Group CBT for women who had VMS following breast cancer treatment (MENOS4), recruited from breast cancer clinics. Inclusion criteria: For both samples, English speaking women, 16 years old or older, having problematic VMS for at least 1 month and minimum frequency of 7 flushes per week were included. For the breast cancer sample, women with primary breast cancer or ductal carcinoma in situ (DCIS) who had completed all primary treatment (surgery and/or radiotherapy and/or chemotherapy but may still be receiving adjuvant endocrine therapy or Herceptin) were included. Exclusion criteria: non-English speaking women and/or with history of medical or psychiatric conditions that would affect their ability to participate.

All women were offered a screening assessment and if eligible and interested, were sent information, a consent form and a baseline questionnaire, which they completed and returned to the research team.

Ethical approval for the student project was obtained from Kings College London Research Ethics Committee (Research Ethics Committee Reference Number: PNM/11/12-122) and for MENOS4 from National Research Ethics Service South Central - Hampshire A Research Ethics Committee and HRA (ref. [16]/SC/0364), University of Southampton Sponsored the study (sponsorship number: 19245). All participants provided written informed consent to participate and were free to withdraw at any time.
Measures

Sociodemographic information included age, ethnicity, relationship status (single, partner, married/cohabiting, divorced/widowed/separated), educational level (left school at 16, or 18 or degree/professional qualifications), employment (working fulltime/part-time or not working). Healthy women (n=40) were also asked: How many times in the last 6 months have you visited your General Practitioner (GP)? How many times in the last 6 months have you visited a hospital doctor? Have you ever sought help for menopause related problems (yes/no)?

*The Hot Flush Rating Scale* (HFRS)\(^20\) is a self-report measure of VMS frequency and problem-rating over the past week. *VMS Frequency* has significant correlations with diary recordings for hot flushes ($r=0.97$, $p<0.001$) and night sweats ($r=0.94$, $p<0.001$).

*HFRS Problem-rating* is calculated as the mean of the scores on three Likert scales (scores range from 1–10) assessing the extent to which HFNS are problematic, distressing and causing interference in daily life. Higher scores indicate more problematic VMS. HFRS *Problem-rating* has good test–retest reliability ($r=0.8$) and internal consistency (alpha=0.87).

The *Hot Flash Related Daily Interference Scale* (HFRDIS),\(^21\) a 10-item questionnaire, measures the impact of VMS on daily activities and quality of life in the past week. Items include: 1 Work (outside the home and housework), 2 Social activities (time with family/friends), 3 Leisure activities (time spent relaxing/doing hobbies), 4 Sleep, 5 Mood, 6 Concentration, 7 Relations with others, 8 Sexuality, 9 Enjoyment of life, 10 Overall quality of life, rated on scales from 0 (do not interfere) to 10 (completely interfere); responses are averaged to range from 0 to 10 with higher scores indicating greater interference (score range 0–10). Internal consistency has been reported as high, Cronbach alpha= 0.92.\(^21\)

The *Hot Flash Interference* (HFI)\(^22\) is a 3-item scale (including items 4,5, and 6 from the HFRDIS assessing interference with sleep, mood and concentration) which is a shortened
validated version of the HFRDIS. Internal consistency alpha=0.82.

The *Work and Social Adjustment Scale* (WSAS)\(^{23}\) is a five-item scale that assesses functional impairment, i.e. an individual's ability to perform day-to-day activities including (1) work, (2) home management, (3) family and relationship interaction and (4) social and (5) private leisure activities, rated on an 8-point Likert scale (0=Not at all, 8=Very Severely). It provides the degree of impact of symptoms (VMS in this case) on a given activity. Score range between 0 (no impairment) to 40 (very severe impairment). The WSAS had good internal consistency (Cronbach alpha ranging between .70 and .90).

The *Generalized Anxiety Disorder 7* (GAD7) Spitzer et al, 2006) is a seven-item screening and severity measure validated for anxiety disorders. Responses were on a 4 point Likert scale from 0 ‘not at all’ to 3 ‘nearly every day’. A total score was calculated for each participant with scores ranging from 0-4 indicating no anxiety, 5-9 mild anxiety, 10-14 moderate anxiety and 15-21 as severe anxiety.

The *Patient Health Questionnaire 9* (PHQ9) (Kroenke et al, 2001) is a nine-item measure of depression with scores range from 0 to 27. Responses are on a 4 point Likert scales from 1 ‘not at all’ to 4 ‘nearly every day’. Depression severity is categorized as: 0-4 no depression, 5-9 mild depression, 10-14 moderate depression, 15-19 moderately severe and 20-27 severe (Kroenke et al, 2001).

The WSAS and help-seeking data were available from the healthy sub-sample only. Sociodemographic and questionnaire data were analyzed (descriptive statistics, independent sample t-tests and Pearson correlations) using the SPSS statistical software package (version 18.0).

RESULTS
One hundred and sixty-nine women took part; 40 from the healthy women sample and 129 from the MENOS4 breast cancer study. The total sample average age was 54.47 (SD=9.11) years; the majority (89.3%) identified as white ethnicity, 75.2% were married or cohabiting, 43% left school at 16 years, 43.6% at 18 years, while 13.3% had degree/professional qualifications; 62% were employed (38.6% fulltime and 23.5% part-time). The two samples did not differ in age or level of education (above or below age 16 years), depressed mood (PHQ9) nor anxiety (GAD7), but more of the breast cancer sample than the healthy sample were married (80% vs. 60%; Chi-squ=6.49, p<0.01) and of white ethnicity (96% vs. 72%; Chi-squ=19.49, p<0.000).

The total sample had mean VMS weekly frequency of 61.33 (SD=40.94), with HFRS problem rating of 6.53 (SD=1.99), HFRDIS total score of 5.36 (SD=2.22) and HFI mean score of 6.13 (SD=2.30). There were no significant differences (HFRS Problem rating p<0.10, HFRDIS p<0.67, HFI p<0.39) between healthy menopausal and breast cancer samples on interference measures; the breast cancer sample reported more frequent VMS (64 vs. 50) but the difference was of marginal significance t=-2.33, p<0.02, 95% CI = -29.95 to -2.14

The internal reliability alpha coefficients were as follows: HFNS Problem Rating 0.84, HFRDIS 0.91 and HFI 0.76. Correlations between HFNS frequency, the three interference measures and the associations between HFNS measures and measures of functioning (WSAS scale, medical help seeking, PHQ9 and GAD7) are shown in Table 1.

Table 1 about here
HFRS Problem-rating, HFRDIS and HFI were all significantly associated (r=0.61-0.85 p<0.001). HFRS Problem-rating was more strongly associated with the HFRDIS (r=0.74 p<0.001) than the shorter HFI (r=0.61, p<0.001). Interference measures were all significantly associated with VMS Frequency but to a lesser degree (r=0.22-0.39).
For concurrent validity, VMS Frequency was not significantly associated with any of the measures of functioning or medical help-seeking, whereas the HFRS Problem-rating Scale, HFRDIS and HFI all had significant associations with WSAS, depression (GAD7) and anxiety (PHQ9) and number of doctor visits (GP and hospital combined) in the past 6 months. The total number of visits to general practitioners and hospital doctors during the past 6 months was on average 3.65 (SD=3.40) with a range of 0 to 13. 42% (n=17) of women had sought medical help specifically for menopause related problems. Medical help-seeking for menopause related problems had non-significant associations with HFRS Problem-rating (Odds Ratio=1.42 CI 0.99-2.03), and to a lesser extent with HFRDIS (OR=1.16 CI 0.87-1.57) and HFI (OR=1.07 CI 0.79-1.45).

DISCUSSION

This study provides evidence to support the use of all three measures of the impact of VMS – a concept which includes impact and interference with daily living, and an appraisal of how distressing and problematic they are perceived to be. The HFRDIS includes a range of situations and contexts that VMS might impact, whereas the HFI focuses on three specific items: VMS impact upon sleep, mood and concentration. The HFRS problem-rating assesses women’s appraisal of symptoms as problematic, distressing and interfering with daily life. The internal reliability and concurrent validity and the correlations between measures suggest that, overall, they are reliable and are measuring similar concepts. The term ‘impact of VMS’ might capture the concept assessed by these measures. While the HFRS problem-rating is more strongly associated with the longer HFRDIS than the shorter HFI, in other respects the shorter HFI performs generally as well as the HRDIS, which supports findings of Carpenter et al.22
Interestingly HFRS Frequency was associated with measures of VMS impact (r=0.22-0.39), but not with work and social adjustment, mood nor medical help-seeking. These findings provide further evidence that considering frequency of VMS as the only outcome measure in clinical trials may fail to capture the important clinically relevant concept of impact of the symptoms on women’s lives. Some treatments, such as CBT, specifically target symptom appraisal and management. Mediation analyses of two trials suggests that CBT works mainly by changing symptom perceptions and cognitive appraisals as well as using helpful behavioural strategies, i.e. changes in these variables mediated the improvement in VMS problem-rating.

Limitations include the small sample size particularly for the WSAS and use of medical services. It would have been difficult to obtain a general estimate of medical help seeking from the breast cancer sample, as this would have been strongly influenced by breast cancer appointments. Future research could assess the relative sensitivity to change of the three measures in response to treatments and their associations with the single item of ‘bother’ which has also been found to be highly correlated with the HFRDIS.

In conclusion, the results suggest that the three measures appear to measure a similar concept of impact of VMS. Short measures tend to be preferred in clinical trials; HFRS problem-rating and the HFI are both short reliable measures of the impact of vasomotor symptoms, with evidence of concurrent validity.
References


behavioural therapy is an effective non-hormonal intervention for managing
vasomotor symptoms. *BMJ* 2015;351:h6434

27. Norton S, Chilcot J, Hunter MS. Cognitive behaviour therapy for menopausal
symptoms (hot flushes and night sweats): moderators and mediators of treatment
effects. *Menopause* 2014;21(6):574-78.

symptoms following breast cancer treatment: who benefits and how does it work?
*Maturitas* 2014;78:56-61.
Table 1 Correlations between HFRS (Frequency and Problem-Rating), HFRDIS, HFI, PHQ9 and GAD7 (n=169), WSAS and use of medical services visits to GP or hospital doctor in general (n=40)

<table>
<thead>
<tr>
<th></th>
<th>HFRS Problem Rating</th>
<th>HFRDIS</th>
<th>HFI</th>
<th>WSAS</th>
<th>GAD7</th>
<th>PHQ9</th>
<th>Doctor visits past 6 months</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HFRS Frequency</strong></td>
<td>0.39 **</td>
<td>0.34 **</td>
<td>0.22 **</td>
<td>0.04 Ns</td>
<td>0.08 ns</td>
<td>0.12 ns</td>
<td>-0.19 ns</td>
</tr>
<tr>
<td><strong>HFRS Problem Rating</strong></td>
<td>0.74 **</td>
<td>0.61 **</td>
<td>0.42 **</td>
<td>0.32 **</td>
<td>0.40 **</td>
<td>0.40 **</td>
<td></td>
</tr>
<tr>
<td><strong>HFRDIS</strong></td>
<td></td>
<td>0.85 **</td>
<td>0.69 **</td>
<td>0.55 **</td>
<td>0.55 **</td>
<td>0.35 *</td>
<td></td>
</tr>
<tr>
<td><strong>HFI</strong></td>
<td></td>
<td></td>
<td>0.65 **</td>
<td>0.48 **</td>
<td>0.53 **</td>
<td>0.35 *</td>
<td></td>
</tr>
</tbody>
</table>

*Sign at the 0.05 level (2-tailed)

**Sign at the 0.01 level (2-tailed)