

## REVIEW ARTICLE

# A critical appraisal of vasomotor symptom assessment tools used in clinical trials evaluating hormone therapy compared to placebo

Marie K. Christakis, MD, FRCSC, NCMP, MPH,<sup>1,2,3</sup> Donna M. Strobino, PhD,<sup>2</sup>  
and Wen Shen, MD, MPH, FACOG<sup>4,5</sup>

---

### Abstract

**Objective:** Vasomotor symptoms (VMS) have been consistently reported as the leading predictor of health-related quality of life (HRQOL) among menopausal women, and the strongest indication for treatment. The North American Menopause Society endorses the use of oral estrogen for the treatment of VMS based on a Cochrane meta-analysis. The Cochrane review concludes that oral hormone therapy reduces the frequency and severity of VMS. The objective of this review is to critically appraise the outcome measures used in these clinical trials to evaluate whether there is adequate evidence that oral hormone therapy improves HRQOL.

**Methods:** Each trial in the 2004 Cochrane review of oral hormone therapy for the management of VMS was evaluated with respect to study design, outcome measures, and method of analysis.

**Results:** Twenty-four randomized, double-blind, placebo-controlled clinical trials were appraised. Six trials were excluded from the Cochrane meta-analysis due to inadequate reporting of outcome measures. Of the remaining trials, 15 trials assessed only symptom frequency and/or severity. One trial used a subscale of the General Health Questionnaire. Two trials used the Greene Climacteric Scale, a validated outcome measure in menopausal women, to directly assess the impact of hormone therapy on HRQOL. Both studies showed an improvement in HRQOL in the hormone-treated group, although the sample size was small (n = 118) and the effect was modest.

**Conclusion:** Although oral hormone therapy improves VMS scores, there is a paucity of evidence on whether it improves HRQOL in menopausal women. Future studies using validated, patient-reported outcome measures that directly assess HRQOL are needed.

**Key Words:** Hormone therapy – Hot flashes – Quality of life – Randomized controlled trials.

---

**M**enopause is a developmental transition resulting from reduced ovarian function which signifies the end of a woman's reproductive years.<sup>1</sup> Menopause

is diagnosed after 12 consecutive months without menstruation.<sup>1</sup> Perimenopause commences with the onset of changes to the menstrual cycle and their associated symptoms and concludes 12 months after the final menstrual period.<sup>1</sup> Nonetheless, moderate to severe symptoms persist, on average, 5 years beyond menopause.<sup>2</sup> How a woman experiences menopause is influenced by biological, psychological, and social factors, which ultimately impact her physical, mental, and sexual well-being.

Among women from high-income countries, the median age at menopause varies between 50 and 52 years, whereas the median age of onset of perimenopause is 48 years old.<sup>3</sup> In 2010, an estimated 45 million women in the United States were older than 52 years.<sup>4</sup> By 2020, this figure is expected to reach more than 50 million.<sup>4</sup> Approximately 6,000 American women reach menopause every day and more than 2 million each year.<sup>4</sup> By 2025, an estimated 45% of American women will be older than 45 years of age.<sup>5</sup>

The most widely regarded symptoms of menopause are vasomotor symptoms (VMS). Also referred to as hot flashes and night sweats, VMS are common among postmenopausal women. Several studies based on nationwide samples report that 60% to 80% of American women of menopausal age experience VMS.<sup>6,7</sup> The results of a 2005 multiethnic cohort study of 4,402 US women aged 40 to 65 years old assessing

---

Received April 11, 2019; revised and accepted May 28, 2019.

From the <sup>1</sup>Department of Obstetrics and Gynaecology, St. Michael's Hospital, Toronto, Canada; <sup>2</sup>Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland; <sup>3</sup>University of Toronto, Toronto, Canada; <sup>4</sup>The Johns Hopkins University School of Medicine, Baltimore, Maryland; and <sup>5</sup>Johns Hopkins School of Nursing, Baltimore, Maryland.

This work was previously presented as an oral abstract entitled "A critical appraisal of vasomotor symptom assessment tools used in clinical trials evaluating estrogen therapy compared to placebo" at The North American Menopause Society Annual Meeting in San Diego, USA, October 2018.

Funding/support: None reported.

Financial disclosure/conflicts of interest: Dr Shen has received a CDC grant "Increase awareness and support among women diagnosed with breast cancer"; FOA DP14-1408; CDC/ONDIEH/NCCDPHP. Dr Shen also has received grant funding from Pfizer. The other authors have nothing to disclose.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Website ([www.menopause.org](http://www.menopause.org)).

Address correspondence to: Marie K. Christakis, MD, FRCSC, NCMP, MPH, Department of Obstetrics & Gynaecology, St. Michael's Hospital, 61 Queen St East – 5th Floor, Toronto, M5C 2T2 Ontario, Canada.  
E-mail: [christakism@smh.ca](mailto:christakism@smh.ca)

the impact of VMS<sup>6</sup> demonstrated that 79% of perimenopausal women and 65% of postmenopausal women experienced hot flushes or night sweats in the prior month. Twenty-four percent of women reported having hot flushes every day. In addition, 9% of perimenopausal women and 7% of postmenopausal women reported at least seven moderate to very severe hot flushes and/or night sweats per day.<sup>6</sup>

Results of the Study of Women's Health Across the Nation (SWAN) showed variation in experience of VMS by biological, social, and behavioral factors. SWAN followed 3,302 mature women nationwide from five ethnic groups over 10 years. African American women and women with an elevated body mass index ( $\geq 27 \text{ kg/m}^2$ ) were at an increased risk of suffering VMS.<sup>8</sup> Smoking, low physical activity level, low socioeconomic status, and having a mother who experienced VMS were also associated with an increased risk of hot flushes.<sup>8,9</sup>

VMS have consistently been reported as the leading predictor of health-related quality of life (HRQOL) among perimenopausal women.<sup>10-14</sup> Although menopause stage was not associated with HRQOL in SWAN, VMS were strongly associated with reduced HRQOL.<sup>7</sup> Similar to other studies, this association was strongest in women with more frequent VMS.<sup>7,10</sup> SWAN results also indicated a strong association between VMS and key quality-of-life outcomes including sleep, mood, and cognitive function.<sup>7</sup> In a US population-based study of 2,703 postmenopausal women, participants reported that VMS affected their work (46%), social activities (44%), leisure activities (48%), sleep (82%), mood (69%), concentration (69%), sexual activity (41%), total energy level (63%), and overall quality of life (69%).<sup>11</sup>

VMS also create workplace challenges for perimenopausal women. A 2016 systematic review of the literature on menopause and the workplace identified 75 relevant studies. The review concluded that a constellation of menopausal symptoms, including VMS, sleep disturbance, and mood changes, may be negatively associated with several occupation-related health constructs, including work ability, capacity to work, absence from work, and job performance.<sup>15</sup> Several studies found that compared to other menopause-related symptoms, VMS were the strongest predictor of poor work performance.<sup>14,16-18</sup> Two cross-sectional studies, involving surveys among a total of 5,970 postmenopausal women, found that VMS are more likely to have a negative impact on work when they are more frequent and bothersome.<sup>15</sup> In a survey of 3,267 postmenopausal women selected from the 2010 US National Health and Wellness Survey, Whiteley et al<sup>19</sup> found that women were significantly more likely to report impairment while at work over the last 7 days when suffering from moderate to severe (24%) symptoms than women with mild VMS (4%). In addition, 32% of women with moderate/severe symptoms reported activities of daily living impairment, compared to 6% of those experiencing mild VMS.<sup>19</sup>

### Vasomotor symptoms and health-related quality of life in hormone therapy trials

In 2004, a Cochrane meta-analysis evaluated whether oral hormone therapy is more effective than placebo in the

management of VMS.<sup>20</sup> A systematic search of relevant databases and a hand search of 20 relevant journals and conference proceedings<sup>20</sup> yielded 24 double-blinded, randomized, placebo-controlled clinical trials of at least 3 months duration (Table 1). The meta-analysis included 18 trials; 6 were not suitable for meta-analysis due to inadequate data about the number of women with VMS, the hot flush frequency, or severity.

The Cochrane review is widely regarded as the highest quality of available evidence supporting the use of oral hormone therapy for the treatment of VMS, and is cited by The North American Menopause Society in their 2017 position statement.<sup>21</sup> The meta-analysis concluded that oral estrogen and combined oestrogen/progestogen therapy significantly reduced weekly hot flush frequency compared to placebo.<sup>20</sup> Symptom severity was also significantly reduced.<sup>20</sup>

The purpose of this article is to critically appraise the VMS assessment tools used in the 18 trials in the meta-analysis of the Cochrane review. A distinction will be made between tools that evaluate symptoms, and those that evaluate HRQOL, some also assessing the impact of VMS. The outcome measures include symptom inventory tools, VMS frequency scales, VMS severity scales, composite VMS outcome measures (hot flush weekly weighted score), and quality of life (QOL) scales (General Health Questionnaire [GHQ], Green Climacteric Scale [GCS]).

### Health-related quality of life

In assessing QOL, investigators attempt to understand the overall perceived well-being of an individual including physical, functional, and psychological impairment.<sup>22</sup> HRQOL seeks to assess how the burden of disease affects an individual's QOL and their perceived health potential.<sup>22</sup> The value of patient-reported outcomes, including HRQOL, in evaluating health care practices and interventions is increasingly recognized with mounting evidence supporting the use of patient-reported outcomes, defined as any direct measure of the patient's experience, to enhance treatment adherence and improve outcomes.<sup>23</sup> There also has been an important shift in clinical research to investigation of outcomes that are comprehensive and patient-centered, including QOL measures such as pain, work function, and overall satisfaction.<sup>22</sup>

A description of symptom inventory tools will be followed by discussion of HRQOL measures used to evaluate the impact of hormone treatment in the Cochrane meta-analysis.

## RESULTS

### Symptom inventory tools

Symptom inventory tools, while a basic method of documenting the presence of VMS, improve upon a single question that captures overall VMS. As menopause is associated with a constellation of symptoms including VMS, sleep disturbance, mood lability, genitourinary complaints, and sexual dysfunction, investigators often list these symptoms and ask study participants to indicate which they experienced.

**TABLE 1.** Vasomotor symptom outcome assessment methods used in the 24 trials reviewed in the 2004 Cochrane review of the oral estrogens compared to placebo in the management of vasomotor symptom

	Trial design	Number of participants included in analysis	VMS outcome assessment	Length of follow-up
Baerug, 1998	Parallel, double-blind, multicenter, placebo-control, clinical trial	108	1. Kupperman Index (completed by investigators) 2. HFWWS and number of responders 3. Severity: mild/moderate/severe (with qualifiers re: perspiration) 4. Greene climacteric scale 5. Mean frequency	12 wk
Baumgardner, 1978	Parallel, double-blind, multicenter, placebo-control, randomized, clinical trial	156	1. Daily frequency (reported /week) 2. Severity: none, mild, moderate, marked	24 wk
Bech, 1998	Parallel, double-blind, placebo-control, randomized clinical trial	105	1. Kupperman Index—12-item version (not included in meta-analysis) 2. GHQ-10 neurasthenia 3. GHQ- 11 quality of life scale	12 mo
Blumel, 1994	Parallel, double-blind, placebo-control, single-center, randomized clinical trial	48	1. Kupperman Index	6 mo
Chung, 1996	Cross-over, double-blind, placebo-control, single-center, randomized clinical trial	83	1. Kupperman Index but added qualifiers for mild/mod/severe (analyzed moderate and severe groups together)	12 months (6 months for each phase)
Conard 1995	parallel, double-blind, placebo-control, randomized clinical trial (number of centers not stated)	50	1. Daily frequency (day and night time) 2. Severity: None/mild/medium/intense	3 mo
Coope, 1975	Cross-over, double-blind, single center, randomized clinical trial	30	1. Daily frequency 2. Kupperman Index	6 mo (3 mo for each crossover phase)
Coope, 1981	Crossover, double-blind, single-center, randomized clinical trial	55	1. Daily frequency (reported/week)	14 mo (6 mo for each treatment and an intervening 2-mo “washout” period)
Derman, 1995	Parallel, double-blind, multicenter, placebo-control, randomized clinical trial	70 For VMS frequency, 78 for VMS severity	1. Daily frequency 2. Greene Climacteric Scale	16 wk
Jensen, 1983	Parallel, placebo-control, single-center, randomized clinical trial	87	1. Kupperman Index (not in meta-analysis. Number with hot flashes only)	12 mo
Jensen, 1987	Parallel, double-blind, placebo-control, randomized clinical trial	57	1. Kupperman Index	2 y
Marslew, 1992	Parallel, double blinded, single center, placebo controlled, randomized clinical trial	39	1. Kupperman Index (not in meta-analysis)	24 mo
Notelovitz, 2000	Parallel, double-blind, multicenter, placebo-controlled, randomized clinical trial	280 For moderate/ severe VMS, 324 for HFWWS	1. Daily frequency (reported /week) 2. HFWWS & number of responders 3. Severity: mild/moderate/severe (analyzed mod & severe together)	3 mo
Paterson, 1982	Crossover, double-blind, single-center, placebo-control, randomized clinical trial	20	1. Daily frequency (reported /week) 2. Severity: 0-3 hot flashes and night sweats separate	6 mo (3 mo for each crossover phase)
PEPI 1998	Parallel, double-blind, multicenter, placebo- control trial	846 (Total); 170 for arm 1 and 166 for arm 5 (placebo) 187 For VMS frequency	1. Dichotomous for 3 vasomotor items: present/absent (hot flashes, night sweats and cold sweats)	36 mo
Symons, 2000 Study 1	Parallel, double-blind, multicenter, placebo-controlled, randomized clinical trial	187 For VMS frequency	1. Daily frequency (reported /week)	4 mo
Symons, 2000 Study 2	Parallel, double-blind, multicenter, placebo-controlled, randomized clinical trial	261 For VMS frequency	1. Daily frequency (reported /week)	3 mo
Vikhlyaeva, 1997	Parallel, double-blind, multicenter, placebo-control trial	32 in treatment group, 28 in placebo group	1. Daily frequency	24 wk
Archer, 1992	Parallel, double-blinded, multicenter, placebo-control, randomized clinical trial	<sup>a</sup>	<sup>a</sup>	12 wk
Campbell, 1976	Crossover, double-blind, single-center, randomized clinical trial	<sup>a</sup>	<sup>a</sup>	12 mo
Daidsen, 1974	Double-blind, cross-over, single-center, randomized clinical trial	<sup>a</sup>	<sup>a</sup>	6 mo (3 mo for each cross-over phase)
Dennerstein, 1978	Double-blind, cross-over, single-center, randomized clinical trial	<sup>a</sup>	<sup>a</sup>	12 mo (3 mo for each of the four treatment groups)
Hagen, 1982	Parallel, double-blind, placebo- control, single-center randomized clinical trial	<sup>a</sup>	<sup>a</sup>	24 mo
Martin, 1971	Parallel, double-blind, placebo-control, randomized clinical trial	120	Only loss to follow up data included	3 mo

GHQ, General Health Questionnaire; HFWWS, Hot Flush Weekly Weighted Score; PEPI, Postmenopausal Estrogen/Progestin Interventions Study; VMS, vasomotor syndromes.

<sup>a</sup>Relevant outcomes including number of women with vasomotor symptoms, hot flush severity/frequency, losses to follow-up, and side effect outcomes were not suitable for meta-analysis.

The PEPI trial included in the Cochrane review used one such symptom inventory tool. Patients were presented with 25 symptoms, 3 of which were related to VMS, and were asked to report a dichotomous “yes” or “no.”<sup>24</sup> The presence of hot flushes alone, however, does not imply impact on functioning. This measurement of presence or absence of symptoms limited conclusions with respect to HRQOL from this trial. Other studies evaluated VMS frequency and/or severity, presuming these variables to be surrogate measures of HRQOL.

### Vasomotor symptoms frequency measures

Measures of VMS frequency were most commonly used in the reviewed trials. Of the 18 trials in the meta-analysis, 10 studies assessed VMS by recording participants’ daily frequency of hot flushes or night sweats.<sup>25-33</sup> Daily symptom frequency was most often reported as the mean number of VMS per week, allowing an objective means by which the treatment group can be compared to the placebo group.<sup>29-33</sup> Using weekly VMS frequency, however, may misrepresent a woman’s experience by distorting effect sizes in the trials. For example, a dramatic change in a small group of women who began with a high frequency of symptoms may yield a significant mean reduction in symptom frequency related to oral hormone therapy, even when a majority of patients did not experience a change.

A woman’s experience also may be misrepresented by weekly VMS frequency as it does not provide information on whether the recorded symptoms are bothersome to patients. The intervention group may display a significant decrease in mean weekly VMS frequency without patients feeling they benefitted from the treatment. In four trials, weekly VMS frequency was the only outcome measure, and VMS symptom severity was not assessed.

In several studies, women were required to complete a daily symptom diary for an extended period of time. Paterson et al<sup>32</sup> required patients to complete daily symptom diaries for 6 months. This approach can yield valuable, detailed data at minimal risk of recall bias if prospectively completed. However, from a practical standpoint, daily symptom frequency diaries over long periods of time can be burdensome to participants. It also is possible that responses in the Patterson study differed over time simply because of waning interest and attentiveness to completing the diary for the study.

### Symptom severity measures

Four trials included in the meta-analysis used a categorical rating scale to assess the severity of VMS.<sup>26,30,31,34</sup> All four studies also assessed frequency of symptoms, although these variables were analyzed and reported separately. All scales were unidirectional, and used similar adjectives to describe response categories. Conard et al<sup>26</sup> used the terms none, mild, medium, or intense, whereas Baumgardner et al<sup>30</sup> asked participants to rate symptoms as none, mild, moderate, or marked. Notelovitz et al<sup>31</sup> used the scale of mild, moderate, or severe. Baerug et al<sup>34</sup> used the same terms, but included a

description of symptoms to guide participants. Women were asked to select “mild” if they had VMS without perspiration, “moderate” if it was accompanied by perspiration, and “severe” if the hot flush required a change of clothing.<sup>34</sup> These details were likely added in an attempt to further understand the impact of VMS, but conclusions regarding HRQOL cannot be drawn from these studies.

An additional five trials in the Cochrane review used the Kupperman index (KI) as a tool to assess the severity of VMS.<sup>25,34-40</sup> The KI was developed in 1952 as a means of evaluating menopausal symptoms among women treated with estrogen compounds (see KI, Supplemental Digital Content 1, <http://links.lww.com/MENO/A434>).<sup>41</sup> The menopausal index includes 11 items, each of which is assigned a value of 0 (none), 1 (slight), 2 (moderate), or 3 (marked), depending on the patient’s reported symptom severity.<sup>41</sup> Based on expert opinion, 4 of the 11 items were deemed to be of greater importance to patients.<sup>41</sup> The VMS item is multiplied by 4, and insomnia, paresthesia, and nervousness are each multiplied by 2. The weighting resulted in a maximum score of 51. Critics of the KI indicate that the questions are outdated and poorly defined, and that the weighting is arbitrary.<sup>42</sup> More importantly, the KI does not specifically assess how a patient is functioning or how symptoms affect their QOL. Of the five trials that used the KI, three trials included it as the sole method of evaluating VMS.<sup>35,36,40</sup> Two combined the KI with other assessment tools.<sup>25,34</sup>

The KI was not consistently completed in the five trials. In one trial, a physician completed the symptom inventory tool based on assessment of a patient’s symptoms.<sup>34</sup> In the other trials, the survey method was not clearly stated. Completion of the KI by a physician may not reflect a woman’s actual symptoms<sup>43</sup> or it may introduce interviewer bias, as physicians may ask questions in a manner which may influence how a patient responds.<sup>43</sup> Furthermore, women may be less truthful in their responses when responding to an interviewer, as compared to privately completing a questionnaire.<sup>44</sup>

There are also concerns regarding the psychometric properties of the KI. There is limited data to support the validity, reliability, and responsiveness of the tool,<sup>43</sup> but psychometric testing suggests poor construct validity.<sup>45</sup> Bech et al<sup>37</sup> reported that items in the KI have an unacceptably low level of internal consistency. Nonetheless, the KI continues to be used because of its ease of use, concise design and availability in multiple languages. As the earliest measure to be developed it has been used in many well-reported studies, allowing for ease of comparison with the literature.

Chung et al., used the KI, but added further descriptors to the survey. Respondents were asked to treat 0 as none, 1 as an “occasional experience without causing distress”, 2 as a “daily occurrence which caused some distress” and 3 as “severe, occurring more often than once each day and interfering with the woman’s life”.<sup>40</sup> Although this approach helps to quantify how debilitating VMS symptoms are perceived by women, investigators combined categories 2 & 3 for analysis. The approach of combining categories to detect

change was used in several other trials. For example, Notevitz et al<sup>31</sup> also combined the moderate and severe groups.<sup>31</sup> While pooling the data reduces the amount of information available for the most debilitated (severe) patients, it often is required where there are insufficient numbers in each category to allow for meaningful analysis, as is likely to be the case for the most severe distress category.

Given concerns regarding separate evaluation of frequency and severity in the previously mentioned studies and with the KI, a composite score was developed,<sup>34</sup> the Hot Flush Weekly Weighted Score (HFWS). This measure was developed with that idea that it better captures a woman's symptoms, although it also does not evaluate the impact of symptoms.

### Hot flush weekly weighted score

The Hot Flush Weekly Weighted Score (HFWS) is a composite score reported on a weekly basis which captures both the frequency and severity of VMS experienced by a woman. With the HFWS, mild hot flushes are multiplied by a factor of 1, moderate by a factor of 3 and severe hot flushes by a factor of 4.<sup>34</sup> This approach attempts to eliminate some problems noted above about pooling or analyzing frequency and severity data separately. For example, women with severe symptoms which occur with low frequency are represented with this method. Although the composite outcome allows for analysis of both severity and frequency, it does not address how symptoms affect women's function or their QOL.

Two trials in the Cochrane review utilized the HFWS over a 3 month period.<sup>31,34</sup> Both studies defined the dependent variable as the number of women with a greater than 90% reduction in their symptoms.<sup>31,34</sup> This number was then compared between the treatment and placebo group at 3 months. A 90% reduction in symptoms was selected to yield a dichotomous outcome, but it did not assess the magnitude of the mean change in symptoms. The physiology of VMS is often represented along a continuous scale and response to treatment may not be best represented by a binary variable. Additionally, patients whose scores worsened over the study were not assessed with this variable. The mean change in symptoms, both positive and negative, would have provided additional information about the impact of treatment.

The previously described measures, used in studies of hormone therapy, evaluate VMS occurrence. The measures do not, however, assess the impact of VMS on HRQOL, or HRQOL more broadly, as experienced by study participants. Some studies attempted to do so, and the tools used in these studies are described below. The measures varied in the studies, with few using the same instruments except the KI to evaluate the impact of treatment.

### The general health questionnaire

One study in the Cochrane review used multiple measures of mental health, including a generic QOL subscale, in addition to the KI.<sup>37</sup> Bech et al<sup>37</sup> asked participants to complete two subscales of the 60-item GHQ, as well as a customized symptom rating scale for depression and anxiety,

based on the Beck Depression Inventory.<sup>37</sup> The outcome from this study described in the meta-analysis was the GHQ QOL 11-item subscale.<sup>46</sup> The subscale has been translated into several languages and validated in multiple populations, although it has not been previously used in relation to menopausal symptom relief.<sup>46</sup> It was used in the trial to assess an indirect association between VMS and HRQOL. The trial results showed a weak correlation of the KI with the GHQ-11 subscale, and no difference in the GHQ between the treatment and placebo group.<sup>37</sup> No other available studies of VMS used this subscale.

Studies that use a generic QOL questionnaire in conjunction with VMS inventories may provide important information about the impact of an intervention on QOL and VMS frequency/severity. This approach did not confirm a link between the intervention and QOL due to improvement in VMS, but it was used in only one study. The GCS, described below, is menopause specific.

### Green climacteric scale

The GCS is a 21-item self-administered scale (see GCS, Supplemental Digital Content 2, <http://links.lww.com/MENO/A435>) that measures the presence of menopause symptoms and their impact on QOL.<sup>47</sup> The questionnaire is separated into four subscales: vasomotor, psychological, somatic, and sexual dysfunction. Respondents are asked to identify how bothered they are by symptoms with response categories: "not at all," "a little," "quite a bit," and "extremely."<sup>47</sup> Each item is assigned a value from 0 to 3, with total scores ranging from 0 to 63.

Greene<sup>47</sup> demonstrated content validity by constructing this questionnaire after reviewing seven factorial studies, which used a mathematical algorithm to delineate important menopausal symptoms. Factorial studies allow for the identification of clustered symptoms and delineate symptoms that are an essential part of a syndrome.<sup>47</sup> The careful selection of items and subscales is an advantage over other previous tools. Compared to the KI, the GCS also uses contemporary terms.<sup>42</sup> It is a more direct assessment of HRQOL, as participants are asked to describe the severity of their symptoms and their impact.

The questionnaire is available in several languages and has been validated for use in evaluating menopausal symptom relief in a variety of populations.<sup>48-51</sup> The psychometric properties of the GCS have been widely studied. In addition to appropriate content validity, confirmatory factor analysis has demonstrated construct validity.<sup>52,53</sup> Cronbach alpha coefficients were calculated for each subscale to assess internal consistency with coefficients for the symptom domain considered acceptable (Cronbach  $\alpha > 0.7$ ). Test-retest reliability coefficients for all subscales have been evaluated as good, with the vasomotor scale reported as having a coefficient of 0.83.<sup>47</sup> Normative values for populations have been established in both low- and high-resource countries.<sup>49,50</sup> The psychometric quality of responsiveness has also been demonstrated, because the GCS can be used to

determine clinically important changes resulting from an intervention. The GCS was most recently used in a study comparing gabapentin, estrogen therapy, and placebo in the treatment of VMS.<sup>54</sup>

Two studies from the meta-analysis used the GCS.<sup>27,34</sup> Both trials were randomized parallel, double-blind, multicenter, placebo-controlled trials that evaluated the effects of oral estradiol combined with a progestin. Baerug et al evaluated two doses of estradiol in combination with norethisterone acetate (NETA) in 119 women.<sup>34</sup> Significant improvements were reported in all subscales from baseline in both treatment arms compared to the placebo.<sup>34</sup> Although the placebo group only saw a 21% reduction in the GCS VMS subscale over 12 weeks, the group receiving 1 mg estradiol/0.5 mg NETA ( $n=38$ ) reported an 88% reduction.<sup>34</sup> Women receiving 1 mg estradiol/0.25 mg NETA ( $n=40$ ) had an even greater reduction in the VMS subscale of 91%.<sup>34</sup> Derman et al reported outcomes on 78 patients, 40 of whom were in the treatment arm<sup>27</sup> based on the vasomotor, psychological, and somatic subscales of the GCS. Investigators did not assess the sexual dysfunction subscale. They also altered the severity score for the response categories to a scale from 0 to 12, an approach that has not been validated in the literature. The hormone therapy group experienced significant improvements in all subscales, whereas the placebo group showed no improvement.<sup>27</sup> Both studies showed an improvement in the GCS in the hormone therapy group compared to placebo, but both included a small sample of participants. Combined, only 115 women were in the placebo group and 118 women received treatment.<sup>27,34</sup>

## DISCUSSION

HRQOL is a concept that captures multiple domains, including physical, mental, emotional, and social functioning.<sup>55</sup> Assessments of HRQOL for menopausal women should evaluate somatic symptoms (VMS, urogenital concerns), psychological symptoms (depression, anxiety, irritability) and the impact of menopause on their function (productivity and employment) and social life.<sup>55</sup> Furthermore, HRQOL allows for assessment of complex concepts, including a patient's sexuality and self-image. By measuring the impact of an intervention on a woman's HRQOL, investigators provide meaningful information to clinicians and prospective patients beyond what can be learned from assessments of somatic symptoms alone.

Significant interest has emerged in understanding the impact of hormone therapy on HRQOL since the Cochrane review was reported. Most recent trials have used a variety of generic QOL measures. In two large, randomized trials evaluating oral hormone therapy, HRQOL was a secondary outcome. The Women's Health Initiative and the Heart and the Estrogen/progestin Replacement Study used the RAND 36-Item Health Survey.<sup>56,57</sup> Neither showed significant improvement in QOL.<sup>56,57</sup> However, in both trials the majority of participants were asymptomatic, as the primary study objective was to determine the impact of hormone therapy on chronic disease.<sup>10</sup> The lack of symptoms also is likely due to

the older ages of women in both trials (mean age above 65 years well after the onset of menopause), and that all women were postmenopausal. Also the QOL tools has not been validated in menopausal women for this purpose.<sup>10</sup>

The Women's International Study of Long Duration Oestrogen after Menopause of 3,721 women also found no difference in QOL at 1 year using EuroQOL, another generic QOL tool.<sup>58</sup> However, investigators concluded in 2008 that hormone therapy improved HRQOL in the domains of VMS, sleep, and sexual functioning at 1 year on the Women's Health Questionnaire (WHQ), a menopause-specific QOL tool.<sup>58</sup> The WHQ contains 36 total items evaluated on a 4-point scale in the 9 domains of depressed mood, somatic symptoms, anxiety/fears, VMS, sleep problems, sexual behavior, menstrual symptoms, memory/concentration, and attractiveness.<sup>59</sup> The WHQ is available in multiple languages and has been validated for use in healthy, perimenopausal women.<sup>59</sup>

Only one study has examined the reliability of the WHQ using the Portuguese translation.<sup>60</sup> This version was shown to have good test-retest reliability and internal consistency (Cronbach alpha  $>0.8$  for each domain).<sup>60</sup> Responsiveness of the questionnaire has not been demonstrated.<sup>43</sup> Although the WHQ is no longer in frequent use, it demonstrated the potential importance of a menopause-specific QOL life scale in assessing hormone therapy.

Recently, the Menopause-Specific Quality of Life Questionnaire (MENQOL) has been more widely used to assess the impact of VMS on HRQOL (see MENQOL, Supplemental Digital Content 3, <http://links.lww.com/MENO/A436>).<sup>61-64</sup> Developed in 1995, the MENQOL is a self-administered instrument consisting of 29 items.<sup>65</sup> Each item asks women to indicate whether they experienced a particular symptom in the past month. If yes, the participants rate how much they were bothered by the symptom on a scale from 0 to 6.<sup>65</sup> The instrument covers the following domains: physical, vasomotor, psychosocial and sexual symptoms, and global QOL.<sup>65</sup> The MENQOL was validated in North American women 2 to 7 years postmenopause as a means of assessing how menopausal symptoms affect HRQOL.<sup>65</sup> The vasomotor domain has been found to have good internal consistency and test-retest reliability.<sup>65</sup> MENQOL has also been shown to be valid and reliable in breast cancer survivors when evaluating HRQOL at the menopausal transition.<sup>66,67</sup>

Compared to the GCS and the WHQ, the psychometric properties of the MENQOL have been more widely studied and confirmed. Although the MENQOL has eight more items than the GCS, the survey is still considered simple to complete and only requires a modest amount of time. As a result, it is unlikely that women would feel burdened when completing this survey, even when repeated over the course of a study to assess their response to treatment.

## CONCLUSIONS

VMS have consistently been reported as the leading predictor of HRQOL and poor work performance among menopausal women, and the strongest indication for hormone

therapy. This review discussed the measures used in 18 trials included in the Cochrane publication evaluating oral hormone therapy for the treatment of VMS. Only two trials used the GCS to more directly evaluate the impact of hormone therapy on VMS and HRQOL. Both studies showed a significant improvement in HRQOL in the combined oral estrogen and progestin treatment arm compared to the placebo group but the sample size for the two studies combined was small.

The GCS appears to be a valid, reliable tool in measuring the impact of menopausal symptoms on QOL in menopausal women but is seldom used. Additional HRQOL tools to assess menopausal symptoms have become available, including the MENQOL, a validated patient-reported outcome tool in the menopause literature. Very recently, Simon et al<sup>64</sup> used MENQOL to evaluate the impact of hormone therapy on HRQOL as a secondary outcome in menopausal women. This trial may illustrate an important shift in outcome assessment.

In conclusion, although the Cochrane meta-analysis demonstrated that hormone therapy significantly decreases the frequency and severity of VMS, there is little evidence on how hormone therapy affects HRQOL. Significant heterogeneity in data collection and reporting limit conclusions from these studies about HRQOL. Future studies would benefit from use of validated, menopause-specific, patient-reported outcome measurement tools to directly assess how the treatment of VMS affects HRQOL.

## REFERENCES

1. The North American Menopause Society. Menopause Glossary. Available at: <http://www.menopause.org/for-women/menopause-glossary>. Accessed March 5, 2018.
2. Freeman EW, Sammel MD, Sanders RJ. Risk of long-term hot flashes after natural menopause: evidence from the Penn Ovarian Aging Study cohort. *Menopause* 2014;21:924-932.
3. McKinlay SM, Brambilla DJ, Posner JG. The normal menopause transition. *Am J Hum Biol* 1992;4:37-46.
4. US Census Bureau. Projections of the Population by Selected Age Groups and Sex for the United States: 2015-2060. 2012;64. Available at: [www.census.gov/population/projections/data/national/2012/summarytables.html](http://www.census.gov/population/projections/data/national/2012/summarytables.html). Accessed March 5, 2018.
5. US Census Bureau. Projections of the Population by Sex and Age for the United States: 2015 to 2060 2014; National Population Projections Tables. Available at: <https://www.census.gov/data/tables/2014/demo/popproj/2014-summary-tables.html>. Published 2014. Accessed April 10, 2018.
6. Williams RE, Kalilani L, DiBenedetti DB, et al. Frequency and severity of vasomotor symptoms among peri- and postmenopausal women in the United States. *Climacteric* 2008;11:32-43.
7. Thurston RC, Joffe H. Vasomotor symptoms and menopause: findings from the study of women's health across the nation. *Obstet Gynecol Clin North Am* 2011;38:489-501.
8. Gold EB, Sternfeld B, Kelsey JL, Brown C, Mouton C, Reame N. Relation of demographic and lifestyle factors to symptoms in a multi-racial /ethnic population of women 40-55 years of age. *Am J Epidemiol* 2000;152:463-473.
9. Utian WH. Psychosocial and socioeconomic burden of vasomotor symptoms in menopause: a comprehensive review. *Health Qual Life Outcomes* 2005;3:47.
10. Schneider HPG, Birkhäuser M. Quality of life in climacteric women. *Climacteric* 2017;20:187-194.
11. Williams RE, Levine KB, Kalilani L, Lewis J, Clark RV. Menopause-specific questionnaire assessment in US population-based study shows negative impact on health-related quality of life. *Maturitas* 2009;62:153-159.
12. Avis NE, Colvin A, Bromberger JT, et al. Change in health-related quality of life over the menopausal transition in a multiethnic cohort of middle-aged women: study of women's health across the nation. *Menopause* 2009;16:860-869.
13. Guthrie JR, Dennerstein L, Taffe JR, Leher P, Burger HG. The menopausal transition: a 9-year prospective population-based study. The Melbourne Women's Midlife Health Project. *Climacteric* 2004; 7:375-389.
14. Whiteley J, DiBonaventura Md, Wagner JS, Alvir J, Shah S. The impact of menopausal symptoms on quality of life, productivity, and economic outcomes. *J Womens Health (Larchmt)* 2013;22:983-990.
15. Jack G, Riach K, Bariola E, Pitts M, Schapper J, Sarrel P. Menopause in the workplace: what employers should be doing. *Maturitas* 2016;85:88-95.
16. Kleinman NL, Rohrbacker NJ, Bushmakin AG, Whiteley J, Lynch WD, Shah SN. Direct and indirect costs of women diagnosed with menopause symptoms. *J Occup Environ Med* 2013;55:465-470.
17. High RV, Marcellino PA. Menopausal women and the work environment. *Soc Behav Personal* 1994;22:347-354.
18. Griffiths A, MacLennan SJ, Hassard J. Menopause and work: an electronic survey of employees' attitudes in the UK. *Maturitas* 2013; 76:155-159.
19. Whiteley J, Wagner J-S, Bushmakin A, Kopenhafer L, Dibonaventura M, Racketa J. Impact of the severity of vasomotor symptoms on health status, resource use, and productivity. *Menopause* 2013;20:518-524.
20. MacLennan AH, Broadbent JL, Lester S, Moore V. Oral oestrogen and combined oestrogen/progestogen therapy versus placebo for hot flushes. *Cochrane Database Syst Rev* 2004;CD002978.
21. The 2017 hormone therapy position statement of The North American Menopause Society. *Menopause* 2017;24:728-753.
22. Anker SD, Agewall S, Borggrefe M, et al. The importance of patient-reported outcomes: a call for their comprehensive integration in cardiovascular clinical trials. *Eur Heart J* 2014;35:2001-2009.
23. Deshpande PR, Rajan S, Sudeepthi BL, Abdul Nazir CP. Patient-reported outcomes: a new era in clinical research. *Perspect Clin Res* 2011;2:137-144.
24. Greendale GA, Reboussin BA, Hogan P, et al. Symptom relief and side effects of postmenopausal hormones: results from the postmenopausal estrogen/progestin interventions trial. *Obstet Gynecol* 1998;92:982-988.
25. Coope J. Double blind cross-over study of estrogen replacement therapy. *The Management of the Menopause and Post-Menopausal Years*. Lancaster: MTP; 1975.
26. Conard J, Basdevant A, Thomas JL, et al. Cardiovascular risk factors and combined estrogen-progestin replacement therapy: a placebo-controlled study with norgestrel acetate and estradiol. *Fertil Steril* 1995;64:957-962.
27. Derman RJ, Dawood MY, Stone S. Quality of life during sequential hormone replacement therapy—a placebo-controlled study. *Int J Fertil Menopausal Stud* 1995;40:73-78.
28. Vikhlyaeva E, Zaidieva Y, Lobova T, Shishkina AV, Larsen S. Trisequens in perimenopausal women with climacteric syndrome: a randomised double-blind trial. *Acta Obstet Gynecol Scand* 1997;76:58.
29. Coope J. Is oestrogen therapy effective in the treatment of menopausal depression? *J R Coll Gen Pract* 1981;31:134.
30. Baumgardner SB, Condrea H, Daane TA, et al. Replacement estrogen therapy for menopausal vasomotor flushes. Comparison of quiniestrol and conjugated estrogens. *Obstet Gynecol* 1978;51:445-452.
31. Notelovitz M, Lenihan JP, McDermott M, Kerber IJ, Nanavati N, Arce J. Initial 17beta-estradiol dose for treating vasomotor symptoms. *Obstet Gynecol* 2000;95:726-731.
32. Paterson MEL. A randomised, double-blind, cross-over study into the effect of sequential mestranol and norethisterone on climacteric symptoms and biochemical parameters. *Maturitas* 1982;4:83-94.
33. Symons J, Kempfert N, Speroff L. Vaginal bleeding in postmenopausal women taking low-dose norethindrone acetate and ethinyl estradiol combinations. *Obstet Gynecol* 2000;96:366-372.
34. Baerug U, Winge T, Nordland G, et al. Do combinations of 1 mg estradiol and low doses of NETA effectively control menopausal symptoms? *Climacteric* 1998;1:219-228.
35. Jensen PB, Jensen J, Riis BJ, Rødbro P, Strøm V, Christiansen C. Climacteric symptoms after oral and percutaneous hormone replacement therapy. *Maturitas* 1987;9:207-215.
36. Blumel JE, Roncagliolo ME, Gramegna G, et al. Double-blind method of the effect of menopause symptoms, lipid profile, and endometrial

- thickness of continuous therapy with estradiol valerate and medroxyprogesterone acetate [in Spanish]. *Rev Chil Obstet Ginecol* 1994;59:354-360.
37. Bech P, Munk-Jensen N, Obel EB, Ulrich LG, Eiken P, Nielsen SP. Combined versus sequential hormonal replacement therapy: a double-blind, placebo-controlled study on quality of life-related outcome measures. *Psychother Psychosom* 1998;67:259-265.
  38. Marslew U, Overgaard K, Riis BJ, Christiansen C. Two new combinations of estrogen and progestogen for prevention of postmenopausal bone loss: long-term effects on bone, calcium and lipid metabolism, climacteric symptoms, and bleeding. *Obstet Gynecol* 1992;79:202-210.
  39. Jensen J, Christiansen C. Dose-response and withdrawal effects on climacteric symptoms after hormonal replacement therapy. A placebo-controlled therapeutic trial. *Maturitas* 1983;5:125-133.
  40. Chung TK, Yip SK, Lam P, Chang AM, Haines CJ. A randomized, double-blind, placebo-controlled, crossover study on the effect of oral oestradiol on acute menopausal symptoms. *Maturitas* 1996;25:115-123.
  41. Kupperman HS, Blatt MHG, Wiesbader H, Filler W. Comparative clinical evaluation of estrogenic preparations by the menopausal and amenorrhoeal indices. *J Clin Endocrinol Metab* 1953;13:688-703.
  42. Alder E. The Blatt-Kupperman menopausal index: a critique. *Maturitas* 1998;29:19-24.
  43. Schneider HPG, MacLennan AH, Feeny D. Assessment of health-related quality of life in menopause and aging. *Climacteric* 2008;11:93-107.
  44. Dayan Y, Paine CS, Johnson AJ. Responding to sensitive questions in surveys: a comparison with results from Online panels, face to face and self-completion interviews. [2014-03-03]. Available at: [http://www.ipsos-mori.com/Assets/Docs/Publications/Ops\\_RMC\\_Responding\\_Sensitive\\_Questions\\_08\\_03\\_10.pdf](http://www.ipsos-mori.com/Assets/Docs/Publications/Ops_RMC_Responding_Sensitive_Questions_08_03_10.pdf). Accessed March 5, 2018.
  45. Cogo-Moreira H, Jackowski AP, De Araújo TRE, Vieira LHL, Haidar MA, Silva I. Poor-fit indexes of Kupperman index: a psychometric analysis based on item response theory. *Menopause* 2015;22:444-447.
  46. Hankins M. The reliability of the twelve-item general health questionnaire (GHQ-12) under realistic assumptions. *BMC Public Health* 2008;8:355.
  47. Greene JG. Constructing a standard climacteric scale. *Maturitas* 1998;29:25-31.
  48. Vasconcelos-Raposo J, Coelho E, Fernandes HM, Rodrigues C, Moreira H, Teixeira C. Factor structure and normative data of the Greene Climacteric Scale among postmenopausal Portuguese women. *Maturitas* 2012;72:256-262.
  49. Chattha R, Kulkarni R, Nagarathna R, Nagendra HR. Factor analysis of Greene's Climacteric Scale for Indian women. *Maturitas* 2008;59:22-27.
  50. Travers C, O'Neill SM, King R, Battistutta D, Khoo SK. Greene Climacteric Scale: norms in an Australian population in relation to age and menopausal status. *Climacteric* 2005;8:56-62.
  51. Barentsen R, Van De Weijer PHM, Van Gend S, Foekema H. Climacteric symptoms in a representative Dutch population sample as measured with the Greene Climacteric Scale. *Maturitas* 2001;38:123-128.
  52. Chen RQ, Davis SR, Wong CM, Lam TH. Validity and cultural equivalence of the standard greene climacteric scale in Hong Kong. *Menopause* 2010;17:630-635.
  53. Committee SA. Assessing Health status and quality-of-life instruments and review criteria. *Qual Life Res* 2002;11:193-215.
  54. Reddy SY, Warner H, Guttuso T, et al. Gabapentin, estrogen, and placebo for treating hot flushes: a randomized controlled trial. *Obstet Gynecol* 2006;108:41-48.
  55. Utian WH, Janata JW, Kingsberg SA, Schluchter M, Hamilton JC. The Utian Quality of Life (UQOL) scale: development and validation of an instrument to quantify quality of life through and beyond menopause. *Menopause* 2002;9:402-410.
  56. Hlatky MA, Boothroyd D, Vittinghoff E, Sharp P, Whooley MA. Quality-of-life and depressive symptoms in postmenopausal women after receiving hormone therapy: results from the Heart and Estrogen/Progestin Replacement Study (HERS) trial. *JAMA* 2002;287:591-597.
  57. Hays J, Ockene JK, Brunner RL, et al. Effects of estrogen plus progestin on health-related quality of life. *N Engl J Med* 2003;348:1839-1854.
  58. Welton AJ, Vickers MR, Kim J, et al. Health related quality of life after combined hormone replacement therapy: randomised controlled trial. *BMJ* 2008;337:a1190.
  59. Hunter MS. The Women's Health Questionnaire (WHQ): frequently asked questions (FAQ). *Health Qual Life Outcomes* 2003;1:41.
  60. Silva Filho CRDa, Baracat EC, Conterno LDO, Haidar MA, Ferraz MB. Climacteric symptoms and quality of life: validity of women's health questionnaire. *Rev Saude Publica* 2005;39:333-339.
  61. Setty P, Redekal L, Warren MP. Vaginal estrogen use and effects on quality of life and urogenital morbidity in postmenopausal women after publication of the Women's Health Initiative in New York City. *Menopause* 2016;23:7-10.
  62. Pinkerton JAV, Bushmakin AG, Komm BS, Abraham L. Relationship between changes in vulvar-vaginal atrophy and changes in sexual functioning. *Maturitas* 2017;100:57-63.
  63. Lethaby A, Ayeleke RO, Roberts H. Local oestrogen for vaginal atrophy in postmenopausal women. *Cochrane Database Syst Rev* (8):2016;CD001500.
  64. Simon JA, Kaunitz AM, Kroll R, Graham S, Bernick B, Mirkin S. Oral 17 $\beta$ -estradiol/progesterone (TX-001HR) and quality of life in postmenopausal women with vasomotor symptoms. *Menopause* 2019;26:506-512.
  65. Hilditch JR, Lewis J, Peter A, et al. A menopause-specific quality of life questionnaire: development and psychometric properties. *Maturitas* 1996;24:161-175.
  66. Radtke JV, Terhorst L, Cohen SM. The menopause-specific quality of life questionnaire: psychometric evaluation among breast cancer survivors. *Menopause* 2011;18:289-295.
  67. Kulasingam S, Moineddin R, Lewis JE, Tierney MC. The validity of the Menopause Specific Quality of Life Questionnaire in older women. *Maturitas* 2008;60:239-243.