



Psychometric Evaluation of the MENQOL Instrument in Women Experiencing Vasomotor Symptoms Associated with Menopause

Neil M. Schultz · Antonia Morga · Emad Siddiqui ·
Stephanie E. Rhoten

Received: September 25, 2023 / Accepted: January 15, 2024 / Published online: February 24, 2024
© The Author(s) 2024, corrected publication [2024]

ABSTRACT

Introduction: Vasomotor symptoms (VMS) associated with menopause can negatively affect health-related quality of life (HRQoL). The Menopause-Specific Quality of Life (MENQOL) questionnaire has been developed to assess QOL specific to menopause. The objective of the current study was to assess the psychometric properties, sensitivity to change, and clinically meaningful within-patient change of the MENQOL using data from the fezolinetant SKYLIGHT 1 and 2 studies in individuals with VMS.

Methods: Individuals aged ≥ 40 to ≤ 65 years with moderate-to-severe VMS (\geq seven hot

flashes/day) were enrolled. In addition to MENQOL, eight patient-reported outcome (PRO) measures were used for the psychometric evaluation. All PRO assessments were completed at weeks 4 and 12 during the treatment period, and most were completed at baseline. Psychometric analyses included factor analysis and reliability, construct validity, and sensitivity to change assessments. The within-patient threshold for a clinically meaningful change in MENQOL was derived.

Results: In total, 1022 individuals were included from SKYLIGHT 1 and 2. Mean MENQOL total score at baseline was 4.30, improving to 3.16 at week 12. The confirmatory factor analysis supported established MENQOL domain structure, including the overall score. The internal consistency of the MENQOL overall and domain scores was supported using Cronbach's alpha and McDonald's omega, and MENQOL construct validity was supported for overall and domain scores. Item-to-item and item-total correlations were generally sufficient, and moderate test-retest reliability was noted. The scales against which construct validity and responsiveness for MENQOL domains were examined were moderately related to the MENQOL domains in general, providing additional support for acceptable measurement properties of MENQOL in this population. A reduction in MENQOL overall score of ≥ 0.9 points was identified as responding to treatment (a clinically important threshold).

The original online version of this article was revised due to update in introduction text.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12325-024-02787-z>.

N. M. Schultz
Medical Affairs, Astellas Pharma, Northbrook, IL,
USA

A. Morga (✉) · E. Siddiqui
Medical Affairs, Astellas Pharma Europe Ltd,
Addlestone KT15 2NX, UK
e-mail: antonia_morga@astellas.com

S. E. Rhoten
Patient Centered Solutions, IQVIA, San Francisco,
CA, USA

Thresholds of 2.0 points for the vasomotor domain and 0.9 for the psychosocial domain were estimated, in addition to distribution-based threshold estimates of 0.8 and 1.2 for the physical and sexual domains, respectively.

Conclusions: The psychometric properties of the MENQOL overall and domain scores support use of this instrument to capture experiences among individuals with moderate-to-severe VMS associated with menopause and assess related endpoints in clinical trials.

Trial Registration: ClinicalTrials.gov identifiers NCT04003155 and NCT04003142.

Keywords: Fezolinetant; Menopause; MENQOL; Psychometric; Vasomotor symptoms

Key Summary Points

Why carry out this study?

Vasomotor symptoms (VMS) associated with menopause can negatively affect health-related quality of life (HRQoL). The Menopause-Specific Quality of Life (MENQOL) questionnaire has been developed to assess QOL specific to menopause.

To support use of MENQOL in studies of novel therapies for VMS, it is important to evaluate its properties and establish a definition of meaningful change (i.e., treatment response).

What was the hypothesis of the study?

The objective of the current study was to assess the psychometric properties, sensitivity to change, and clinically meaningful within-patient change of the MENQOL using data from the SKYLIGHT 1 and 2 studies, featuring fezolinetant, a novel non-hormonal neurokinin 3 receptor antagonist approved for treatment of VMS due to menopause.

What was learned from the study?

Overall, the results provide evidence of acceptable psychometric properties of the MENQOL overall and domain scores using 1-week recall, supporting use of this instrument to capture experiences among people with moderate-to-severe VMS associated with menopause and assess related endpoints in clinical trials.

What has been learned from the study?

These results support the internal consistency reliability, convergent reliability, and structural validity of the MENQOL found in other populations such as breast cancer survivors and participants with diabetes.

INTRODUCTION

Menopause is often accompanied by bothersome vasomotor symptoms (VMS), particularly hot flashes and night sweats [1–3]. Hormone therapy is an effective VMS treatment but is contraindicated in women with a history of breast or endometrial cancer, coronary heart disease, venous thromboembolism, or stroke [4]. For some, therefore, non-hormonal treatment may be preferable to hormone therapy.

Fezolinetant is a novel selective non-hormonal neurokinin 3 receptor antagonist approved by the US Food and Drug Administration (FDA) for treatment of moderate-to-severe VMS due to menopause and by the European Medicines Agency for treatment of moderate-to-severe VMS associated with menopause. In phase 2 placebo-controlled studies in participants aged 40–65 years, fezolinetant significantly reduced the frequency and severity of moderate-to-severe VMS and improved a range of patient-reported outcomes (PROs) [5, 6]. Consequently, two phase 3 studies—SKYLIGHT 1 and 2—investigated efficacy and safety of fezolinetant.

During SKYLIGHT 1 and 2, quality of life (QOL) was assessed using a range of PRO measures, including the Menopause-Specific Quality of Life (MENQOL) questionnaire, which assesses

QOL specific to menopause. Other PROs that were not menopause specific were the Patient-Reported Outcomes Measurement Information System Sleep Disturbance—Short Form 8b (PROMIS SD SF 8b), the Patient Global Impression of Change in Sleep Disturbance (PGI-C SD), and the Patient Global Impression of Severity in Sleep Disturbance (PGI-S SD). While not specific to menopause, to support use of MENQOL in studies of novel therapies for VMS, it is important to evaluate its psychometric properties and sensitivity to change and establish a definition of meaningful change (i.e., treatment response). On the basis of guidance issued by the US FDA [7, 8], we assessed these properties using pooled data from SKYLIGHT 1 and 2 in participants diagnosed with menopause-associated moderate-to-severe VMS.

METHODS

Study Design and Patients

SKYLIGHT 1 (NCT04003155) and 2 (NCT04003142) were phase 3, randomized, placebo-controlled, double-blind studies with identical designs. These studies were conducted in accordance with the Declaration of Helsinki, Good Clinical Practice, and International Council for Harmonisation guidelines. An independent ethics committee or institutional review board reviewed the ethical, scientific, and medical appropriateness of the study at each site before data collection. Individuals aged ≥ 40 to ≤ 65 years with moderate-to-severe VMS (minimum average seven hot flashes/day) who were female at birth were randomized to once-daily fezolinetant 30 mg, fezolinetant 45 mg, or placebo (1:1:1) for 12 weeks. All women had spontaneous amenorrhea for at least 12 consecutive months, spontaneous amenorrhea for at least 6 months with biochemical criteria of menopause (follicle stimulating hormone > 40 IU/L), bilateral oophorectomy at least 6 weeks before the screening visit (with or without hysterectomy), and a BMI of 18–38 kg/m². A 40-week open-label extension followed, in which all participants received active treatment (individuals initially

randomized to placebo were re-randomized to fezolinetant 30 mg or 45 mg, and fezolinetant-treated individuals continued their original dose). Full details of the study designs and inclusion/exclusion criteria have been published [9, 10]. A total of 1022 women were randomized and received ≥ 1 dose of study drug across both studies (placebo, $n = 342$; fezolinetant 30 mg, $n = 339$; fezolinetant 45 mg, $n = 341$). Mean (standard deviation [SD]) age was 54.3 (5.0) years, and the majority of the women were White (828 [81.1%]). Demographic data were largely balanced across groups, although mean (range) time since onset of VMS was slightly longer in the placebo group (81.9 [2–422] months) versus the fezolinetant 30 mg (76.7 [3–370] months) and 45 mg (76.9 [1–396] months) groups.

MENQOL Questionnaire

The MENQOL questionnaire is a self-reported measure that assesses QOL specific to menopause [11, 12]. Respondents are asked whether they have experienced any of the 29 symptoms within the past week and to rate how bothersome each symptom was on a 7-point Likert scale (0 = not at all bothered; 6 = extremely bothered). The 29 items are combined into four domains: vasomotor (three items), psychosocial (seven items), physical (16 items), and sexual (three items). For analysis, the questionnaire score becomes 1 for no; 2 for yes, not bothered through to 8 for yes, extremely bothered. The score by domain is the mean of the converted item scores forming that domain and ranges from 1 to 8 [12].

Additional PRO Measures

Eight other PRO measures included in the SKYLIGHT 1 and 2 studies were used to evaluate the MENQOL questionnaire: PRO Measurement Information System Sleep-Related Impairment—Short Form 8a (PROMIS SRI SF 8a); PROMIS Sleep Disturbance—Short Form 8b (PROMIS SD SF 8b); VMS episodes captured using an electronic diary; Patient Global Impression of Severity Sleep Disturbance (PGI-S SD); Patient

Global Impression of Change Sleep Disturbance (PGI-C SD); Patient Global Impression of Change Vasomotor Symptoms (PGI-C VMS); EuroQoL 5-dimension 5-level (EQ-5D-5L) questionnaire, including the EQ visual analog scale (VAS); and Work Productivity and Activity Impairment questionnaire specific to Vasomotor Symptoms (WPAI-VMS). Further details on these instruments are included in Supplementary Table S1.

PRO assessments were completed at baseline and weeks 4 and 12 during the treatment period, except for PGI-C assessments, which were completed at weeks 4 and 12 only, as PGI-C measures change from baseline.

Descriptive Analyses

All PRO analyses were performed on the full analysis set (all randomized participants who received ≥ 1 dose of study drug). Completion rates were the number of participants with a completed item entry at each clinic visit divided by the total number in the full analysis set.

Descriptive analysis was performed at item and score level using pooled arms at baseline, week 4, and week 12 for MENQOL items. At each timepoint, the number (%) of participants with completed responses for each item was recorded and descriptive statistics generated for each item and for domain and total scores. Floor ($> 20\%$ of responses for the lowest/least severe option) and ceiling ($> 20\%$ of responses for the highest/most severe option) effects were investigated.

Psychometric Evaluation

Confirmatory factor analysis (CFA) was conducted on MENQOL using a second-order model including four factors, as previously described [13], with three latent factors represented for vasomotor, psychosocial, and sexual domains, with physical domain defined as the observed factor using the physical score (i.e., mean score of 16 domain items). These factors were loaded on the second-order latent factor representing the overall score (main construct). The maximum likelihood method was used, as

this assumes input data are multivariate normal, which may be approximately true for average scores using ordinal data with eight categories. CFA produced several goodness-of-fit measures or fit indices to evaluate the model, including the standardized root mean residual, root mean square error of approximation, and comparative fit index. McDonald's omega was also calculated as a measure of internal consistency and reliability.

Internal consistency was assessed using Cronbach's alpha coefficient (values ≥ 0.70 indicate acceptable reliability [14]); 95% confidence intervals (CI) were calculated for the total score and four MENQOL domains at baseline and week 12. Alpha-if-item-deleted results were calculated.

Test-retest reliability was assessed using a two-way mixed, absolute agreement, single measure intraclass correlation coefficient (ICC) at baseline (test) and week 4 (retest). ICC values of 0.50–0.90 and > 0.90 represent moderate-to-good and excellent reliability, respectively [15].

Spearman correlation coefficients were used to evaluate convergent validity between MENQOL overall and domain scores and other PROs. At least moderate correlations ($r > 0.30$) were expected between overall and domain scores of similar constructs. MENQOL assessed QOL specific to menopause, which was predicted be moderately associated with VMS (frequency and severity), sleep disturbance and related impairments (PGI-S SD, PROMIS SD SF 8b, PROMIS SRI SF 8a scores), overall health (EQ VAS), as well as VMS-related work productivity and activity (WPAI scores). Known-groups validity was evaluated by comparing mean MENQOL overall and domain scores across groups defined using EQ VAS and VMS severity. Analysis of variance with orthogonal planned comparisons was used to test hypotheses that MENQOL scores differ significantly between adjacent quartile groups defined using EQ VAS scores at baseline. The known-groups served as the independent variable, and MENQOL scores were dependent variables in the analysis of variance with $\alpha = 0.05$.

Sensitivity to change was evaluated using Spearman correlations and analysis of covariance. Changes in PRO scores from baseline to

week 12 were correlated with changes in MENQOL scores. Concurrent improvement in PRO measures was expected to result in moderate-to-strong correlations.

Changes in MENQOL overall and domain scores from baseline to week 12 were assessed using mean changes for improved versus non-improved participants and through significance testing using separate analysis of covariance models controlling for baseline with $\alpha = 0.05$. Change from baseline to week 12 was calculated for MENQOL scores among EQ VAS response groups (response = improvement of ≥ 0.5 SD of baseline score) and PGI-C VMS response groups (responders = participants reporting better, much better, or very much better). The dependent variable was change from baseline at week 12, and the model included the responder or improvement factor as fixed.

Thresholds for meaningful within-patient change on MENQOL were estimated using anchor-based approaches together with distribution-based estimates and receiver operating characteristic (ROC) curves. These methods are consistent with guidance from the FDA for determining responder thresholds [7, 8].

Meaningful within-patient change was evaluated using the PGI-C VMS as an anchor (a suitable anchor typically has a correlation of ≥ 0.30 [16]). Spearman correlations between categories of response scores in PGI-C VMS (Supplementary Table S2) and MENQOL summary score response at weeks 4 and 12 were assessed, as were descriptive statistics for raw change. Using the PGI-C VMS anchor, meaningful within-subject change was defined as a PGI-C VMS score of + 2 (moderately better). Mean changes in MENQOL domain and total scores from baseline to weeks 4 and 12 were calculated for the following categories of change: + 3 or + 2 (much better or moderately better), + 1 (a little better), 0 (no change), - 1 (a little worse) and - 2 or - 3 (moderately worse or much worse) (Supplementary Table S2). Mean changes in MENQOL scores from baseline to weeks 4 and 12 were calculated for ≥ 1 -point increase, no change (0 points), and ≥ 1 -point decrease. Sensitivity and specificity were calculated and ROC curves derived

using logistic regression analyses, with anchor change group variables collapsed into two groups: moderate or better improvement and minimal/no improvement or deterioration (Supplementary Table S3). Responder status was the dependent variable in each model, and change from baseline in MENQOL score was the independent variable.

The clinically meaningful threshold was defined by the change value corresponding to the cutoff in the ROC space that minimizes the sum of squares of $(1 - \text{sensitivity})$ and $(1 - \text{specificity})$ and is therefore closest to the top-left corner (1, 0) of the ROC space [17].

For the supportive distribution-based methods, statistical parameters included effect size (Cohen's d), SD, and standard error of measurement (defined as $SD \cdot \sqrt{[1 - r]}$, where r is the internal consistency of the instrument). Effect sizes were defined as small (0.2), medium (0.5), or large (0.8) [18].

Statistical Analyses

Statistical comparisons were made using two-sided tests at the $\alpha = 0.05$ level. For point estimates, 95% CIs were used. All analyses were performed using SAS version 9.3 or higher (SAS Institute, Cary, North Carolina, USA). See Fig. S1 in the electronic supplementary material for details of the CFA second-order model code.

RESULTS

Completion Rates

Of the expected 1022 participants in SKYLIGHT 1 and 2, 99.5–99.7% had baseline data, with completion rates remaining high at 91.2% at weeks 4 and 84.6% at week 12. Compliance rates were $> 94\%$ at all timepoints.

Mean MENQOL Scores

The mean MENQOL total score at baseline was 4.30, improving to 3.25 at week 4 and 3.16 at week 12 (Fig. 1). All item mean scores demonstrated the greatest decrease from baseline to

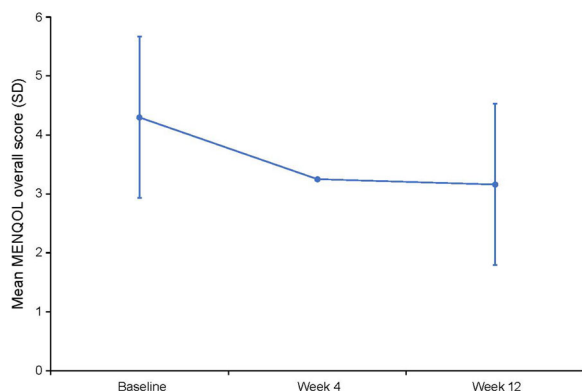


Fig. 1 MENQOL total score over time. MENQOL total score over time. *MENQOL* Menopause Quality of Life questionnaire, *SD* standard deviation

week 4, with smaller improvements between weeks 4 and 12 (Supplementary Table S4). The greatest symptom bother at baseline was reported for items in the vasomotor domain (mean range 6.41–6.58) and the individual item “difficulty sleeping” (mean 5.60). While large ceiling effects were observed for all other items at baseline (mean range 2.13–4.87), all improved at weeks 4 and 12 (Supplementary Table S4). For domain scores, baseline vasomotor score was the highest (mean 6.52), improving to 4.74 at week 4 and 4.36 at week 12. Lowest mean scores were observed for the psychosocial, physical, and sexual domains, ranging from 3.37 to 3.69 at baseline, 2.52–2.88 at week 4, and 2.51–2.91 at week 12.

Psychometric Evaluation

Excellent fit was observed for the second-order model (Table 1; Supplementary Fig. S2). Factor loadings were consistently high for each domain, ranging from 0.75 to 0.82 for vasomotor, 0.64–0.74 for psychosocial, and 0.67–0.88 for the sexual domain. Latent domain factor loadings and physical domain scores were also highly related to the general factor (range 0.43–0.93).

Overall, MENQOL showed a high degree of internal consistency (Cronbach’s alpha: baseline, 0.93; week 12, 0.94; Table 2). Internal consistency for the domain scores was also high at both timepoints (baseline, 0.83–0.90;

week 12, 0.84–0.91). Furthermore, the coefficients for MENQOL scores when each item was individually deleted further supported inter-item consistency (MENQOL overall score—item omitted: baseline, 0.93–0.93; week 12, 0.71–0.91; Table 2).

Correlations between items within the same domain at baseline were generally sufficient without suggesting redundancy (defined as $r = 0.40$ – 0.90), particularly for the vasomotor ($r = 0.76$ – 0.80), psychosocial ($r = 0.42$ – 0.67), and sexual domains ($r = 0.67$ – 0.84 ; Supplementary Table S5). However, some weaker correlations were observed for physical domain items 14–17 (difficulty sleeping, aches in back of neck or head, decrease in physical strength, and decrease in stamina), items 21–22 (increased facial hair and changes in appear/texture/tone of skin), and items 24–26 (low backache, frequent urination, and involuntary urination when laugh/cough), which ranged from 0.23 to 0.40. Excluding these weak correlations, the physical item correlations ranged from 0.40 to 0.88.

Moderate-to-strong item-total correlations were observed at baseline ($r = 0.44$ – 0.73) and week 12 ($r = 0.40$ – 0.73) for MENQOL overall score, with no redundant items (Table 3). Furthermore, strong correlations were observed between items within each of the domain scores, notably the relationship between items 1–3 and the vasomotor score at baseline ($r = 0.72$ – 0.73) and week 12 ($r = 0.81$ – 0.82 ; Table 3).

Moderate test–retest reliability was observed for MENQOL overall score when PGI-C VMS was used to define stable participants (ICC 0.71) as well as the domain scores (ICC 0.61–0.71) (Table 4). When EQ VAS was used to define stable participants, test–retest reliability was moderate for the physical and sexual domain scores (ICC 0.53 and 0.65, respectively) but lower for other domain scores (ICC 0.21–0.47; Table 4).

MENQOL overall scores at baseline moderately correlated ($r > 0.30$) with PROMIS SRI SF 8a, PROMIS SD SF 8b, PGI-S SD, EQ VAS, and WPAI activity impairment, presenteeism, and overall work productivity loss scores (Table 5). Weak correlations were observed for frequency

Table 1 Confirmatory factor analysis of MENQOL overall score at baseline (second-order model)

Hypothesized model structure	Statistic ^a	MENQOL overall score (N = 1017)
All items constitute one factor	RMSEA (90% CI)	0.06 (0.05, 0.06)
	SRMR	0.04
	CFI	0.96
	McDonald’s omega	0.82
Factor loadings	Domain	General
Vasomotor domain factor		0.43
Item 1: hot flushes or flashes	0.75	
Item 2: night sweats	0.82	
Item 3: sweating	0.80	
Psychosocial domain factor		0.84
Item 4: dissatisfaction with personal life	0.67	
Item 5: feeling anxious or nervous	0.68	
Item 6: poor memory	0.64	
Item 7: accomplishing less than used to	0.73	
Item 8: feeling depressed, down, or blue	0.74	
Item 9: being impatient with other people	0.67	
Item 10: feelings of wanting to be alone	0.66	
Physical domain factor		0.93
Sexual domain factor		0.56
Item 27: decrease in sexual desire	0.88	
Item 28: vaginal dryness	0.66	
Item 29: avoiding intimacy	0.82	

CFI comparative fit index, CI confidence interval, MENQOL Menopause-Specific Quality of Life questionnaire, RMSEA root mean square error of approximation, SRMR standardized root mean residual

^aRules for assessing the model fit: SRMR: ≤ 0.08, RMSEA ≤ 0.06; CFI ≥ 0.95

and severity of VMS ($r = 0.09$ and 0.18 , respectively) and for WPAI absenteeism ($r = 0.20$). For the vasomotor domain score, correlations were moderate with PROMIS SD SF 8b, PGI-S SD, and WPAI scores except absenteeism (absolute $r = 0.34$ – 0.43). Correlations between vasomotor domain score and frequency and severity of VMS ($r = 0.21$ and 0.23 , respectively) were lower, as were correlations with PROMIS SRI SF

8a, EQ VAS, and WPAI absenteeism (absolute $r = 0.02$ – 0.30). Similar patterns of moderate correlations were observed for the psychosocial and physical domain scores between the same measures (absolute $r = 0.35$ – 0.46 and 0.32 – 0.48 , respectively), except for notably higher correlations for PROMIS SRI SF 8a and EQ VAS (absolute $r = 0.42$ – 0.61). Similar to overall scores, weak correlations were found for VMS

Table 2 Internal consistency reliability analysis of MENQOL overall and domain scores

Item	Baseline		Week 12	
	Alpha item omitted	Cronbach's alpha ^a (95% CI)	Alpha item omitted	Cronbach's alpha ^a (95% CI)
MENQOL overall score		0.93 (0.93; 0.94)		0.94 (0.93; 0.94)
Item 1: hot flushes or flashes	0.93		0.94	
Item 2: night sweats	0.93		0.94	
Item 3: sweating	0.93		0.94	
Item 4: dissatisfaction with personal life	0.93		0.94	
Item 5: feeling anxious or nervous	0.93		0.94	
Item 6: poor memory	0.93		0.94	
Item 7: accomplishing less than used to	0.93		0.93	
Item 8: feeling depressed, down, or blue	0.93		0.94	
Item 9: being impatient with other people	0.93		0.94	
Item 10: feelings of wanting to be alone	0.93		0.94	
Item 11: passing gas or gas pains	0.93		0.94	
Item 12: aching in muscles and joints	0.93		0.94	
Item 13: feeling tired or worn out	0.93		0.93	
Item 14: difficulty sleeping	0.93		0.94	
Item 15: aches in back of neck or head	0.93		0.94	
Item 16: decrease in physical strength	0.93		0.93	
Item 17: decrease in stamina	0.93		0.93	
Item 18: lack of energy	0.93		0.93	
Item 19: dry skin	0.93		0.94	
Item 20: weight gain	0.93		0.94	
Item 21: increased facial hair	0.93		0.94	
Item 22: changes in appear/texture/tone of skin	0.93		0.94	
Item 23: feeling bloated	0.93		0.94	
Item 24: low backache	0.93		0.94	
Item 25: frequent urination	0.93		0.94	

Table 2 continued

Item	Baseline		Week 12	
	Alpha item omitted	Cronbach's alpha ^a (95% CI)	Alpha item omitted	Cronbach's alpha ^a (95% CI)
Item 26: involuntary urination when laugh/cough	0.93		0.94	
Item 27: decrease in sexual desire	0.93		0.94	
Item 28: vaginal dryness	0.93		0.94	
Item 29: avoiding intimacy	0.93		0.94	
Vasomotor domain score		0.83 (0.81; 0.85)		0.90 (0.88; 0.91)
Item 1: hot flushes or flashes	0.79		0.84	
Item 2: night sweats	0.75		0.85	
Item 3: sweating	0.76		0.86	
Psychosocial domain score		0.86 (0.85; 0.87)		0.88 (0.87; 0.89)
Item 4: dissatisfaction with personal life	0.84		0.87	
Item 5: feeling anxious or nervous	0.84		0.86	
Item 6: poor memory	0.85		0.87	
Item 7: accomplishing less than used to	0.84		0.87	
Item 8: feeling depressed, down, or blue	0.83		0.85	
Item 9: being impatient with other people	0.84		0.86	
Item 10: feelings of wanting to be alone	0.84		0.86	
Physical domain score		0.90 (0.90; 0.91)		0.91 (0.90; 0.92)
Item 11: passing gas or gas pains	0.90		0.91	
Item 12: aching in muscles and joints	0.90		0.90	
Item 13: feeling tired or worn out	0.90		0.90	
Item 14: difficulty sleeping	0.90		0.91	
Item 15: aches in back of neck or head	0.90		0.90	
Item 16: decrease in physical strength	0.90		0.90	
Item 17: decrease in stamina	0.89		0.90	
Item 18: lack of energy	0.90		0.90	

Table 2 continued

Item	Baseline		Week 12	
	Alpha item omitted	Cronbach's alpha ^a (95% CI)	Alpha item omitted	Cronbach's alpha ^a (95% CI)
Item 19: dry skin	0.90		0.91	
Item 20: weight gain	0.90		0.91	
Item 21: increased facial hair	0.90		0.91	
Item 22: changes in appear/texture/tone of skin	0.90		0.91	
Item 23: feeling bloated	0.90		0.90	
Item 24: low backache	0.90		0.91	
Item 25: frequent urination	0.90		0.91	
Item 26: involuntary urination when laugh/cough	0.90		0.91	
Sexual domain score		0.83 (0.81; 0.85)		0.84 (0.82; 0.86)
Item 27: decrease in sexual desire	0.70		0.72	
Item 28: vaginal dryness	0.84		0.89	
Item 29: avoiding intimacy	0.74		0.71	

CI confidence interval, *MENQOL* Menopause-Specific Quality of Life questionnaire

^aCronbach's alpha is calculated between each item and the total score omitting the item

frequency and severity and WPAI absenteeism with psychosocial and physical scores (absolute $r = 0.04$ – 0.19 and 0.04 – 0.19 , respectively). The *MENQOL* sexual domain score had low correlations with all PRO measures (absolute $r = 0.04$ – 0.26) (Table 5).

In the known-groups validity analysis, there were significant differences in *MENQOL* overall and domain scores across EQ VAS quartiles and VMS severity groups at baseline (all $p \leq 0.012$), except for vasomotor score using EQ VAS quartiles and sexual scores using VMS severity (Table 6). *MENQOL* overall scores were significantly different between adjacent EQ VAS quartile and VMS severity groups (all $p \leq 0.013$).

Moderate correlations ($r > 0.30$) were observed for the *MENQOL* overall change score ($r = 0.37$ – 0.50) and vasomotor domain change

score ($r = 0.38$ – 0.60) with PGI-S SD, PGI-C SD, PGI-C VMS, PROMIS SRI SF 8a, PROMIS SD SF 8b, and frequency of VMS (Table 7). Moderate correlations were also observed for change from baseline in *MENQOL* psychosocial domain score with PGI-S SD, PGI-C VMS, PROMIS SRI SF 8a, and PROMIS SD SF 8b ($r = 0.30$ – 0.46) and change from baseline in physical score with PGI-S, PROMIS SRI SF 8a, PROMIS SD SF 8b, and EQ VAS (absolute $r = 0.32$ – 0.49 ; Table 7).

Moderate correlations were found between the anchor and *MENQOL* overall score ($r = 0.47$ – 0.48) and vasomotor domain score ($r = 0.60$ – 0.60 ; Supplementary Table S6). Anchor support was demonstrated for the psychosocial score, although the correlations were lower ($r = 0.30$ – 0.30). Changes in physical and sexual scores were weakly correlated with PGI-C

Table 3 Item-total correlation analysis of MENQOL overall and domain scores at baseline and week 12

Item	Overall		Vasomotor		Psychosocial		Physical		Sexual	
	Baseline	Week 12	Baseline	Week 12	Baseline	Week 12	Baseline	Week 12	Baseline	Week 12
Item 1: hot flushes or flashes	0.48	0.59	0.72	0.82	0.26	0.28	0.32	0.33	0.20	0.24
Item 2: night sweats	0.55	0.60	0.73	0.82	0.33	0.31	0.39	0.32	0.24	0.26
Item 3: sweating	0.53	0.60	0.73	0.81	0.33	0.32	0.39	0.31	0.23	0.27
Item 4: dissatisfaction with personal life	0.63	0.60	0.40	0.37	0.65	0.67	0.55	0.53	0.33	0.32
Item 5: feeling anxious or nervous	0.55	0.59	0.20	0.26	0.67	0.70	0.52	0.57	0.28	0.33
Item 6: poor memory	0.59	0.58	0.28	0.23	0.59	0.60	0.58	0.62	0.32	0.34
Item 7: accomplishing less than used to	0.67	0.64	0.36	0.31	0.68	0.65	0.66	0.65	0.34	0.37
Item 8: feeling depressed, down, or blue	0.62	0.62	0.21	0.22	0.73	0.77	0.59	0.61	0.33	0.36
Item 9: being impatient with other people	0.58	0.59	0.21	0.24	0.63	0.66	0.60	0.59	0.30	0.36
Item 10: feelings of wanting to be alone	0.58	0.64	0.21	0.24	0.64	0.68	0.56	0.62	0.31	0.43
Item 11: passing gas or gas pains	0.45	0.40	0.17	0.14	0.41	0.38	0.54	0.52	0.24	0.24
Item 12: aching in muscles and joints	0.54	0.52	0.28	0.20	0.47	0.48	0.65	0.67	0.26	0.29
Item 13: feeling tired or worn out	0.68	0.70	0.37	0.34	0.67	0.71	0.71	0.77	0.33	0.38
Item 14: difficulty sleeping	0.58	0.61	0.48	0.58	0.51	0.53	0.55	0.51	0.28	0.25
Item 15: aches in back of neck or head	0.55	0.53	0.28	0.23	0.51	0.48	0.64	0.67	0.28	0.30
Item 16: decrease in physical strength	0.68	0.67	0.26	0.26	0.67	0.65	0.74	0.75	0.39	0.44
Item 17: decrease in stamina	0.73	0.69	0.36	0.24	0.67	0.66	0.77	0.78	0.42	0.45
Item 18: lack of energy	0.70	0.70	0.40	0.28	0.68	0.70	0.76	0.77	0.37	0.44
Item 19: dry skin	0.56	0.56	0.23	0.20	0.49	0.51	0.60	0.63	0.36	0.40
Item 20: weight gain	0.52	0.48	0.30	0.19	0.40	0.41	0.61	0.58	0.30	0.31
Item 21: increased facial hair	0.47	0.42	0.21	0.11	0.32	0.32	0.54	0.47	0.34	0.37
Item 22: changes in appear/texture/tone of skin	0.56	0.52	0.24	0.17	0.41	0.45	0.58	0.61	0.43	0.39
Item 23: feeling bloated	0.61	0.56	0.32	0.22	0.48	0.47	0.69	0.68	0.39	0.39

Table 3 continued

Item	Overall		Vasomotor		Psychosocial		Physical		Sexual	
	Baseline	Week 12	Baseline	Week 12	Baseline	Week 12	Baseline	Week 12	Baseline	Week 12
Item 24: low backache	0.54	0.49	0.23	0.16	0.46	0.44	0.65	0.64	0.32	0.30
Item 25: frequent urination	0.53	0.48	0.27	0.20	0.46	0.40	0.56	0.56	0.32	0.35
Item 26: involuntary urination when laugh/cough	0.44	0.45	0.14	0.16	0.37	0.37	0.46	0.52	0.31	0.34
Item 27: decrease in sexual desire	0.71	0.73	0.27	0.31	0.42	0.47	0.50	0.53	0.75	0.75
Item 28: vaginal dryness	0.55	0.55	0.18	0.23	0.29	0.32	0.38	0.38	0.64	0.60
Item 29: avoiding intimacy	0.68	0.70	0.24	0.28	0.41	0.43	0.49	0.50	0.73	0.78

MENQOL Menopause-Specific Quality of Life Questionnaire

For the domain items, polyserial correlation coefficients are calculated between each item and the domain score omitting the item. For the remaining items, polyserial correlation coefficients are calculated between each item and the domain score

Table 4 Test–retest reliability analysis of *MENQOL* overall and domain scores among stable patients from baseline to week 4 in *PGI-C VMS* and *EQ VAS* scores

	<i>n</i>	<i>ICC</i> (95% <i>CI</i>)
<i>PGI-C VMS</i> no change		
<i>MENQOL</i> overall score	193	0.71 (0.63, 0.78)
<i>MENQOL</i> vasomotor score	193	0.61 (0.51, 0.69)
<i>MENQOL</i> psychosocial score	193	0.71 (0.63, 0.77)
<i>MENQOL</i> physical score	193	0.70 (0.61, 0.78)
<i>MENQOL</i> sexual score	193	0.71 (0.63, 0.77)
<i>EQ VAS</i> no change		
<i>MENQOL</i> overall score	522	0.40 (0.06, 0.62)
<i>MENQOL</i> vasomotor score	522	0.21 (– 0.01, 0.39)
<i>MENQOL</i> psychosocial score	522	0.47 (0.28, 0.60)
<i>MENQOL</i> physical score	522	0.53 (0.28, 0.68)
<i>MENQOL</i> sexual score	522	0.65 (0.54, 0.73)

CI confidence interval, *EQ VAS* Euro-QoL visual analog scale, *ICC* intraclass correlation coefficient, *MENQOL* Menopause-Specific Quality of Life questionnaire, *PGI-C VMS* Patient Global Impression of Change in Vasomotor Symptoms

VMS ($r = 0.28–0.29$ and $0.18–0.18$, respectively).

For results with sufficient anchor correlations (area under the curve [*AUC*] > 0.70), *AUCs* for the *MENQOL* overall and domain scores ranged from 0.73 to 0.81 (Supplementary Table S7). Psychosocial score *AUCs* were, however, below the recommended value of 0.70 (week 4, 0.65; week 12, 0.66). Thresholds for overall score were – 1.08 and – 0.91 at weeks 4 and 12 and for the vasomotor domain were – 2.00 at both timepoints. The threshold for the psychosocial score was – 0.71 at both timepoints, although *AUCs* were low. The physical and sexual domains showed low anchor correlations and *AUCs* (Supplementary Table S7).

For the *MENQOL* overall score (score range 1–8), anchor-based mean change estimates at

Table 5 Convergent validity: correlations between MENQOL overall and domain scores and assessments of related constructs at baseline

Measure	n	MENQOL score ^a				
		Overall	Vasomotor	Psychosocial	Physical	Sexual
Frequency of VMS	1017	0.09	0.21	0.04	0.04	0.04
Severity of VMS	1017	0.18	0.23	0.19	0.15	0.07
PROMIS SRI SF 8a ^b	1017	0.60	0.30	0.61	0.58	0.26
PROMIS SD SF 8b ^b	1017	0.43	0.38	0.37	0.44	0.20
PGI-S SD	1017	0.49	0.43	0.46	0.48	0.21
EQ VAS	1016	− 0.34	− 0.02	− 0.42	− 0.46	− 0.14
WPAI activity impairment	1016	0.44	0.41	0.38	0.38	0.22
WPAI absenteeism	622	0.20	0.09	0.19	0.19	0.15
WPAI presenteeism	621	0.39	0.34	0.35	0.32	0.23
WPAI overall work productivity loss	621	0.40	0.34	0.35	0.32	0.24

EQ VAS Euro-QoL visual analog scale, *MENQOL*, Menopause-Specific Quality of Life, *PGI-S SD* Patient Global Impression of Severity Sleep Disturbance, *PROMIS SD SF 8b* Patient-Reported Outcomes Measurement Information System Sleep Disturbance—Short Form 8b, *PROMIS SRI SF 8a* Patient-Reported Outcomes Measurement Information System Sleep-Related Impairment—Short Form 8a, *VMS* vasomotor symptoms, *WPAI* Work Productivity and Activity Impairment

^aSpearman rank correlation coefficients are presented

^bFor the PROMIS SRI SF 8a and PROMIS SD SF 8b instruments, only the total score is used for the correlation

weeks 4 and 12 were − 1.05 and − 1.19, respectively, using the PGI-C VMS anchor (median − 0.98 and − 1.19; Table 8). ROC analyses provided support for the lower end of this range, even when only using the largest ROC estimates; the smallest threshold arising from the ROC analyses was − 0.91 points. Therefore, a threshold of 0.90 points based on the lowest anchor-based median and ROC estimates was selected for MENQOL overall score.

Vasomotor score anchor-based mean change estimates at weeks 4 and 12 were larger than for overall score (− 1.91 and − 2.02, respectively; median − 2; score range 1–8; Table 8). Again, ROC analyses provided support for the lower end of this range, with thresholds estimated at − 2.0 points. The same assessment for vasomotor score led to a threshold of 2.0 points, above the distribution-based estimate (0.5 SD at baseline) and the largest estimate for the “no change” anchor category (0.75). Thresholds for psychosocial score at weeks 4 and 12 were −

0.83 and − 0.87, respectively, for the anchor-based mean change estimates (median − 0.43 and − 0.71; score range 1–8); the ROC estimate was − 0.71, resulting in a threshold of 0.9 for the psychosocial score, in line with distribution-based and “no change” estimates (0.5 SD at baseline) as well as participants reporting “no change” (− 0.45).

DISCUSSION

This analysis was designed to evaluate the psychometric properties, sensitivity to change, and clinically meaningful within-patient change of the MENQOL questionnaire in individuals experiencing moderate-to-severe VMS related to menopause who were treated with fezolinetant. MENQOL completion rates were high, ranging from 99.5% at baseline to 84.6% at week 12. Overall, the results provide evidence of acceptable psychometric properties of the MENQOL

Table 6 Known-groups validity: analysis of mean MENQOL overall and domain scores by EQ VAS quartile groups at baseline

MENQOL score	EQ VAS categories	n	LS mean (SE)	95% CI	p value ^a	Contrast p value ^b
Overall	Quartile 1	265	4.95 (0.08)	(4.79; 5.10)	< 0.0001	
	Quartile 2	240	4.36 (0.08)	(4.20; 4.53)		< 0.0001
	Quartile 3	265	4.07 (0.08)	(3.92; 4.23)		0.013
	Quartile 4	246	3.77 (0.08)	(3.61; 3.94)		0.0095
Vasomotor domain	Quartile 1	265	6.57 (0.09)	(6.39; 6.75)	0.70	
	Quartile 2	240	6.44 (0.10)	(6.25; 6.63)		0.31
	Quartile 3	265	6.57 (0.09)	(6.39; 6.75)		0.31
	Quartile 4	246	6.50 (0.10)	(6.31; 6.68)		0.56
Psychosocial domain	Quartile 1	265	4.37 (0.10)	(4.17; 4.57)	< 0.0001	
	Quartile 2	240	3.58 (0.11)	(3.37; 3.79)		< 0.0001
	Quartile 3	265	2.98 (0.10)	(2.78; 3.18)		< 0.0001
	Quartile 4	246	2.48 (0.11)	(2.27; 2.69)		0.0008
Physical domain	Quartile 1	265	4.65 (0.09)	(4.48; 4.83)	< 0.0001	
	Quartile 2	240	3.90 (0.09)	(3.72; 4.09)		< 0.0001
	Quartile 3	265	3.38 (0.09)	(3.20; 3.55)		< 0.0001
	Quartile 4	246	2.80 (0.09)	(2.62; 2.98)		< 0.0001
Sexual domain	Quartile 1	265	4.19 (0.15)	(3.91; 4.48)	< 0.0001	
	Quartile 2	240	3.53 (0.15)	(3.23; 3.83)		0.0018
	Quartile 3	265	3.37 (0.15)	(3.08; 3.65)		0.45
	Quartile 4	246	3.32 (0.15)	(3.02; 3.62)		0.83

CI confidence interval, EQ VAS EuroQoL visual analog scale, LS least squares, MENQOL Menopause-Specific Quality of Life questionnaire, SE standard error

^ap value is based on ANOVA comparisons with alpha = 0.05 level

^bContrast p value comparisons examined adjacent quartile groups (i.e., quartiles 1 vs 2, 2 vs 3, and 3 vs 4)

Table 7 Sensitivity to change: correlations between MENQOL overall and domain scores and PRO variables change from baseline to week 12

Change from baseline to week 12 in MENQOL score	Change from baseline to week 12						
	PGI-S SD	PGI-C SD ^a	PGI-C VMS ^a	PROMIS SRI SF 8a ^b	PROMIS SD SF 8b ^b	EQ VAS	Frequency of VMS
Overall	0.47	0.41	0.47	0.50	0.47	– 0.27	0.37
Vasomotor	0.45	0.48	0.60	0.38	0.46	– 0.14	0.48
Psychosocial	0.35	0.28	0.30	0.46	0.38	– 0.28	0.22
Physical	0.40	0.29	0.28	0.49	0.43	– 0.32	0.21
Sexual	0.20	0.18	0.18	0.23	0.16	– 0.16	0.14

EQ VAS EuroQoL visual analog scale, *MENQOL* Menopause-Specific Quality of Life, *PGI-C SD* Patient Global Impression of Change Sleep Disturbance, *PGI-C VMS* Patient Global Impression of Change Vasomotor Symptoms, *PGI-S SD* Patient Global Impression of Severity Sleep Disturbance, *PROMIS SD SF 8b* Patient-Reported Outcomes Measurement Information System Sleep Disturbance—Short Form 8b, *PROMIS SRI SF 8a*, Patient-Reported Outcomes Measurement Information System Sleep-Related Impairment—Short Form 8a, *VMS* vasomotor symptoms

^aFor the PGI-C SD and PGI-C VMS instruments, response at follow-up is used

^bFor the PROMIS SRI SF 8a and PROMIS SRI SF 8a instruments, change in total score is used

overall and domain scores using 1-week recall, supporting use of this instrument to capture experiences among people with moderate-to-severe VMS associated with menopause and assess related endpoints in clinical trials.

At baseline, greatest symptom bother was reported for vasomotor domain items (rated moderately to highly bothersome), and symptom improvement was seen at weeks 4 and 12. Baseline scores for psychosocial, physical, and sexual domains were low to moderate at baseline relative to vasomotor scores. Correlations between items were generally moderate ($r > 0.4$), with some weak correlations between physical domain items. Correlations between each item and the overall and domain scores (omitting that item) were also moderate to high, with no redundant items. CFA provided additional support for the established MENQOL domain structure, including overall score. The second-order model demonstrated acceptable fit and generally strong relationships between the items, domains, and overall score.

Internal consistency of the MENQOL overall and domain scores were supported using Cronbach’s alpha and McDonald’s omega, and MENQOL construct validity was supported for

overall and domain scores. Overall and domain scores differentiated well between groups defined by EQ VAS severity at baseline except for vasomotor scores. Additionally, MENQOL scores between VMS severity groups were significantly different at baseline and week 12 except for sexual scores at baseline. Adequate convergent validity was generally demonstrated by moderate correlations between overall and scale scores and PRO measures of related constructs, although some weak correlations were observed. Longitudinal analysis using two timepoints provided support for sensitivity to change (baseline to week 12) and test–retest reliability (baseline to week 4; ICC 0.50–0.90) except for sexual domain scores. Test–retest reliability was lower using EQ VAS than PGI-C VMS, possibly because the EQ VAS is a general health measure, while the PGI-C VMS is symptom specific and more related to MENQOL overall and domain scores.

Thresholds for defining clinically important responses were estimated on the basis of within-subject change on each scale score. These data were triangulated considering the range of the scale, sufficient anchor correlations, the smallest improvement exceeding the 0.5 SD at

Table 8 Mean change in MENQOL overall and domain scores within PGI-C VMS categories of change from baseline to weeks 4 and 12

Score/anchor	<i>n</i>	Mean change	Median change	95% CI
MENQOL overall score				
PGI-C VMS week 4				
Much or moderately better (collapsed)	461	− 1.58	− 1.45	(− 1.70, − 1.46)
Improvement (collapsed)	708	− 1.31	− 1.20	(− 1.41, − 1.21)
Much better	307	− 1.85	− 1.71	(− 2.00, − 1.70)
Moderately better*	154	− 1.05	− 0.98	(− 1.24, − 0.86)
A little better	247	− 0.80	− 0.78	(− 0.94, − 0.66)
No change	193	− 0.27	− 0.27	(− 0.43, − 0.11)
PGI-C VMS week 12				
Much or moderately better (collapsed)	488	− 1.61	− 1.52	(− 1.73, − 1.49)
Improvement (collapsed)	696	− 1.31	− 1.19	(− 1.41, − 1.20)
Much better	327	− 1.82	− 1.68	(− 1.96, − 1.68)
Moderately better*	161	− 1.19	− 1.19	(− 1.38, − 0.99)
A little better	208	− 0.59	− 0.62	(− 0.77, − 0.41)
No change	129	− 0.31	− 0.26	(− 0.51, − 0.10)
MENQOL vasomotor score				
PGI-C VMS week 4				
Much or moderately better (collapsed)	461	− 2.86	− 2.67	(− 3.04, − 2.68)
Improvement (collapsed)	708	− 2.31	− 2.00	(− 2.46, − 2.16)
Much better	307	− 3.34	− 3.33	(− 3.56, − 3.12)
Moderately better*	154	− 1.91	− 2.00	(− 2.16, − 1.65)
A little better	247	− 1.28	− 1.33	(− 1.48, − 1.09)
No change	193	− 0.25	0.00	(− 0.45, − 0.04)
PGI-C VMS week 12				
Much or moderately better (collapsed)	488	− 3.10	− 3.00	(− 3.28, − 2.93)
Improvement (collapsed)	696	− 2.56	− 2.33	(− 2.71, − 2.40)
Much better	327	− 3.63	− 3.67	(− 3.84, − 3.42)
Moderately better*	161	− 2.02	− 2.00	(− 2.26, − 1.79)
A little better	208	− 1.28	− 1.33	(− 1.52, − 1.03)
No change	129	− 0.48	− 0.33	(− 0.75, − 0.21)

Table 8 continued

Score/anchor	<i>n</i>	Mean change	Median change	95% CI
MENQOL psychosocial score				
PGI-C VMS week 4				
Much or moderately better (collapsed)	461	− 1.25	− 1.14	(− 1.40, − 1.09)
Improvement (collapsed)	708	− 1.03	− 0.86	(− 1.15, − 0.90)
Much better	307	− 1.45	− 1.29	(− 1.65, − 1.26)
Moderately better*	154	− 0.83	− 0.71	(− 1.06, − 0.61)
A little better	247	− 0.61	− 0.43	(− 0.80, − 0.42)
No change	193	− 0.22	0.00	(− 0.43, 0.00)
PGI-C VMS week 12				
Much or moderately better (collapsed)	488	− 1.21	− 0.86	(− 1.36, − 1.05)
Improvement (collapsed)	696	− 0.95	− 0.71	(− 1.08, − 0.82)
Much better	327	− 1.37	− 1.14	(− 1.56, − 1.19)
Moderately better*	161	− 0.87	− 0.43	(− 1.14, − 0.59)
A little better	208	− 0.36	− 0.14	(− 0.56, − 0.15)
No change	129	− 0.17	0.00	(− 0.45, 0.11)

CI confidence interval, *MENQOL* Menopause-Specific Quality of Life questionnaire, *PGI-C VMS* Patient Global Impression of Change Vasomotor Symptoms

*Categories were used for the triangulation of meaningful change thresholds

baseline, and lower 95% CI estimates for participants experiencing “no change” on the anchors. On the basis of FDA guidance, more consideration was allotted to anchor-based estimates [7, 8]. The analysis supports a MENQOL overall score reduction of ≥ 0.9 points as responding to treatment (a clinically important threshold). Thresholds of 2.0 points for the vasomotor domain and 0.9 for the psychosocial domain were proposed, in addition to distribution-based threshold estimates of 0.8 and 1.2 for the physical and sexual domains, respectively.

A few prior publications have assessed clinically/minimally important difference in postmenopausal women with moderate/severe VMS, but they anchored with different PRO measures compared to our study and involved different treatments. In two hormone therapy studies, weekly VMS frequency [19] or severity [20] was anchored to generic CGI (not specific

to VMS) outcomes. Other publications that have reported responder thresholds in moderate-to-severe VMS include those reporting VMS frequency and severity anchored to the Menopause Symptoms Treatment Satisfaction Questionnaire in women treated with desvenlafaxine [21]; VMS frequency anchored to CGI and the MENQOL questionnaire in women treated with hormone therapy [22]; and VMS frequency anchored to the Hot Flash Related Daily Interference scale/Hot Flash Interference scale in women treated with escitalopram [23].

The current results support the good psychometric properties of the MENQOL (internal consistency reliability, convergent reliability, and structural validity) found in other populations such as breast cancer survivors [24, 25] and participants with diabetes [26].

Limitations are primarily due to challenges associated with evaluating MENQOL sexual

domains. Anchors for meaningful change were not optimal for MENQOL physical and sexual domains. However, as noted by Bushmakina et al. [13], an individual may experience some physical symptoms of menopause (such as difficulty sleeping) but not others (such as drying skin). Correlations of anchors with overall and vasomotor domain scores were between 0.47 and 0.60 and were lower for the psychosocial domain (0.30–0.30) because of different relationships between VMS and the domains. More precise estimates of meaningful change may be offered by targeted global impressions of change, focused on specific concepts of interest for the physical and sexual domains. The scales against which construct validity and responsiveness for MENQOL domains was examined were moderately related to MENQOL domains in general, providing additional support for acceptable measurement properties of MENQOL in this population.

CONCLUSION

These analyses confirm the measurement properties of the MENQOL questionnaire. Additionally, within-person clinically important response thresholds have been established using appropriate anchors and distribution-based methods. Overall, these results suggest MENQOL is fit for purpose to evaluate appropriate endpoints in trials investigating moderate-to-severe VMS associated with menopause.

ACKNOWLEDGEMENTS

Medical Writing Assistance Medical writing support was provided by Dan Booth and Sue Cooper from Envision Pharma, Inc., and funded by the study sponsor.

Author Contributions. Neil M. Schultz and Stephanie E. Rhoten designed and executed the analysis. All authors developed and interpreted the data for the study, drafted the manuscript or revised it critically for important intellectual content, and provided final approval of the

published manuscript. We thank the participants of the SKYLIGHT 1 and SKYLIGHT 2 studies.

Funding. This study was funded by Astellas Pharma Inc. who also funded the journal's Rapid Service and Open Access Fees.

Data Availability. Researchers may request access to anonymized participant level data, trial level data, and protocols from Astellas-sponsored clinical trials at www.clinicalstudydatarequest.com. For the Astellas criteria on data sharing see <https://clinicalstudydatarequest.com/Study-Sponsors/Study-Sponsors-Astellas.aspx>. The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of Interest. All authors have received support for medical writing. Neil M. Schultz, Antonia Morga, and Emad Siddiqui are employees of Astellas Pharma Inc. Stephanie E. Rhoten received consulting fees from IQVIA.

Ethical Approval. This analysis used data that were acquired from the SKYLIGHT 1 and 2 studies. These studies were conducted in accordance with the Declaration of Helsinki, Good Clinical Practice, and International Council for Harmonisation guidelines. An independent ethics committee or institutional review board reviewed the ethical, scientific, and medical appropriateness of the study at each site before data collection. The independent ethics committees or institutional review boards that were involved in reviewing the SKYLIGHT 1 and 2 studies were Advarra Headquarters (USA); Advarra Canada (Canada); CEI, Hospital 12 de Octubre, Centro Actividades Ambulatoria (Spain); Egészségügyi Tudományos Tanács Klinikai Farmakológiai Etikai Bizottság (ETT-KFEB; Hungary); Ethics Committee of Communal Nonprofit Enterprise Maternity Hospital #4 of Zaporizhzhya City Council (Ukraine); Eticka komise FN Brno (Czechia); Ethics Committee for Clinical Research, Pauls Stradins Clinical University Hospital Development Society

(Latvia); Komisja Bioetyczna, przy Okregowej Radzie Lekarskiej, Wielkopolskiej Izby Lekarskiej w Poznaniu (Poland); and South Central–Hampshire B Research Ethics Committee; HSC A, Office for Research Ethics Committees Northern Ireland; and East Midlands–Nottingham 2 Research Ethics Committee (UK). Signed informed consent forms were obtained from the participants before any study-related procedures were performed.

Open Access. This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

1. Thurston RC, Joffe H. Vasomotor symptoms and menopause: findings from the Study of Women's Health across the Nation. *Obstet Gynecol Clin N Am.* 2011;38(3):489–501.
2. Gold EB, Colvin A, Avis N, et al. Longitudinal analysis of the association between vasomotor symptoms and race/ethnicity across the menopausal transition: study of women's health across the nation. *Am J Public Health.* 2006;96(7):1226–35.
3. Nappi RE, Kroll R, Siddiqui E, et al. Global cross-sectional survey of women with vasomotor symptoms associated with menopause: prevalence and quality of life burden. *Menopause.* 2021;28(8):875–82.
4. The NAMS 2017 Hormone Therapy Position Statement Advisory Panel. The 2017 hormone therapy position statement of The North American Menopause Society. *Menopause.* 2017;24(7):728–53.
5. Fraser GL, Lederman S, Waldbaum A, et al. A phase 2b, randomized, placebo-controlled, double-blind, dose-ranging study of the neurokinin 3 receptor antagonist fezolinetant for vasomotor symptoms associated with menopause. *Menopause.* 2020;27(4):382–92.
6. Depypere H, Timmerman D, Donders G, et al. Treatment of menopausal vasomotor symptoms with fezolinetant, a neurokinin 3 receptor antagonist: a phase 2a trial. *J Clin Endocrinol Metab.* 2019;104(12):5893–905.
7. US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), Center for Devices and Radiological Health (CDRH). Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims. 2009. <https://www.fda.gov/media/77832/download>. Accessed 5 Feb 2024.
8. US Food & Drug Administration. Patient-focused drug development guidance public workshop. Methods to identify what is important to patients & select, develop or modify fit-for-purpose clinical outcomes assessments. 2018. <https://www.fda.gov/media/116277/download>. Accessed 5 Feb 2024.
9. Lederman S, Ottery FD, Cano A, et al. Fezolinetant for treatment of moderate-to-severe vasomotor symptoms associated with menopause (SKYLIGHT 1): a phase 3 randomised controlled study. *Lancet.* 2023;401(10382):1091–102.
10. Johnson KA, Martin N, Nappi RE, et al. Efficacy and safety of fezolinetant in moderate to severe vasomotor symptoms associated with menopause: a phase 3 RCT. *J Clin Endocrinol Metab.* 2023;108(8):1981–97.
11. Lewis JE, Hilditch JR, Wong CJ. Further psychometric property development of the Menopause-Specific Quality of Life questionnaire and development of a modified version MENQOL-Intervention questionnaire. *Maturitas.* 2005;50(3):209–21.
12. Hilditch JR, Lewis J, Peter A, et al. A menopause-specific quality of life questionnaire: development and psychometric properties. *Maturitas.* 1996;24(3):161–75.
13. Bushmakina AG, Abraham L, Pinkerton JV, Cappelletti JC, Mirkin S. Evaluation of the measurement model and clinically important differences for menopause-specific quality of life associated with

- bazedoxifene/conjugated estrogens. *Menopause*. 2014;21(8):815–22.
14. Fayers PM, Machin D. *Quality of life: the assessment, analysis and interpretation of patient-reported outcomes*. 2nd ed. Chichester: Wiley; 2007.
 15. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155–63.
 16. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol*. 2008;61(2):102–9.
 17. Froud R, Abel G. Using ROC curves to choose minimally important change thresholds when sensitivity and specificity are valued equally: the forgotten lesson of Pythagoras. Theoretical considerations and an example application of change in health status. *PLoS ONE*. 2014;9(12):e114468.
 18. Wyrwich KW, Norquist JM, Lenderking WR, Acaster S, the Industry Advisory Committee of International Society for Quality of Life Research (ISO-QOL). Methods for interpreting change over time in patient-reported outcome measures. *Qual Life Res*. 2013;22(3):475–83.
 19. Gerlinger C, Gude K, Hiemeyer F, Schmelter T, Schafers M. An empirically validated responder definition for the reduction of moderate to severe hot flushes in postmenopausal women. *Menopause*. 2012;19(7):799–803.
 20. Constantine GD, Simon JA, Kaunitz AM, et al. TX-001HR is associated with a clinically meaningful effect on severity of moderate to severe vasomotor symptoms in the REPLENISH trial. *Menopause*. 2020;27(11):1236–41.
 21. Wyrwich KW, Spratt DI, Gass M, Yu H, Bobula JD. Identifying meaningful differences in vasomotor symptoms among menopausal women. *Menopause*. 2008;15(4 Pt 1):698–705.
 22. Constantine GD, Revicki DA, Kagan R, et al. Evaluation of clinical meaningfulness of estrogen plus progesterone oral capsule (TX-001HR) on moderate to severe vasomotor symptoms. *Menopause*. 2019;26(5):513–9.
 23. Carpenter JS, Bakoyannis G, Otte JL, et al. Validity, cut-points, and minimally important differences for two hot flash-related daily interference scales. *Menopause*. 2017;24(8):877–85.
 24. Radtke JV, Terhorst L, Cohen SM. The Menopause-Specific Quality of Life Questionnaire: psychometric evaluation among breast cancer survivors. *Menopause*. 2011;18(3):289–95.
 25. Phligbua W, Smith EML, Barton DL. Psychometric properties of the Menopause Specific Quality of Life questionnaire among Thai women with a history of breast cancer. *Eur J Oncol Nurs*. 2018;36:75–81.
 26. Hasan SS, Ahmadi K, Santiago R, Ahmed SI. The validity of the Menopause-specific Quality of Life questionnaire in women with type 2 diabetes. *Climacteric*. 2014;17(4):456–64.