DOI: 10.7860/JCDR/2026/85707.23945

Original Article

Physiotherapy Section

Development and Psychometric Evaluation of the Menopausal Symptom Bothersomeness Questionnaire: A Cross-sectional Study

PRIYASINGH BHAGIRATHSINGH RANGEY[1], NEHAL SHAH[2]

ABSTRACT

Introduction: Menopausal females experience a diversity of symptoms related to various organ systems, which significantly affect their daily function and Quality of Life (QoL). Bothersomeness depends on each individual's tolerance level; hence, bothersomeness due to menopausal symptoms is subjective and varies with each woman.

Aim: To develop a valid and reliable questionnaire to assess the bothersomeness of menopausal symptoms.

Materials and Methods: A cross-sectional study was conducted among menopausal women from multiple states of India (New Delhi, Haryana, Uttar Pradesh, Gujarat, Rajasthan, Madhya Pradesh, Maharashtra, and Karnataka), India from May 2023 to April 2025. A mixed-methods approach was used to determine the content domain and generate items, and a preliminary questionnaire of 46 items was created. A nine-member expert panel was formed to review and validate the content. Factor analysis was performed to reduce items. Reliability and construct validity were assessed. Content validity was determined using Content Validity Index (CVI), Content Validity Ratio (CVR) and Kappa values.

Results: Among a total of 106 menopausal women, 52 females were peri-menopausal, and 54 were post-menopausal. Mean age of the sample was 51.13±5.067 years. Mean ages of peri-menopausal and post-menopausal women were 47.83±2.185 years and 54.31±5.024 years, respectively. The questionnaire's validity was assessed using multiple methods. Content validity indices (I-CVI: 0.67-1.00; S-CVI/Ave: 0.95) guided item selection, resulting in a draft of 33 questions. Exploratory factor analysis conducted with 106 participants identified a nine-factor structure comprising 33 questions. Internal consistency was acceptable to excellent (Cronbach's α range: 0.635-0.921). Test-retest reliability was moderate to excellent, with Spearman's rho and ICC values exceeding 0.5 in a subsample of 22 participants. Construct validity was established through comparison with the Greene Climacteric Scale and the Menopausal QoL questionnaire in a different sample of 107 participants. Item 24 was removed due to insufficient loading, leading to 32 items in the final questionnaire.

Conclusion: The Menopausal Symptom Bothersomeness Questionnaire (MSBQ), containing 32 items divided into nine domains, was developed to assess the bothersomeness due to menopausal symptoms in menopausal females. The questionnaire has acceptable internal consistency and moderate to excellent test-retest reliability and has established construct and content validity. It is a valid and reliable comprehensive Patient-Reported Outcome Measure (PROM) that can be used in the community setting.

Keywords: Factor analysis, Menopause, Quality of life

INTRODUCTION

Menopause is defined as the permanent cessation of menstruation as a result of loss of ovarian follicular activity [1]. Menopause leads to estrogen-deficiency symptoms such as hot flushes, day and night sweats, vulvovaginal atrophy, disturbed sleep, lethargy, low sexual drive, depressed mood, increased anxiety and brain fog. These symptoms are experienced by the majority of menopausal women, and for quite a significant proportion, they are severe and may last for a long time [2].

The bothersomeness of each symptom may vary from one menopausal female to another. Bothersomeness is defined as the extent to which symptoms interfere with the daily life of an individual [3]. Menopausal symptoms are mild initially and increase later with prolonged amenorrhea and hypoestrogenism [4].

Prevalence and severity of menopausal symptoms depend on confounding factors like lifestyle, psychological status, social status and body composition [5]. Studies have reported a negative impact of menopausal symptoms on QoL [5-8]. Rathnayake N et al., conducted a study to evaluate the prevalence and severity of menopausal symptoms and QoL amongst premenopausal and postmenopausal women. They concluded that both the prevalence and severity of menopausal symptoms and impaired QoL were higher amongst postmenopausal women than premenopausal women [5].

A similar study was conducted by Whiteley J et al., [6]. They concluded that women with menopausal symptoms reported lower levels of health-related QoL, along with higher work impairment and healthcare utilisation. Depression, joint stiffness and anxiety were strongly associated with health outcomes [7].

Another study by Kupcewicz E et al., compared the QoL in somatic, social, mental and environmental dimensions amongst nurses in the full reproductive period with those in the peri-menopausal period and found that significant differences exist in the general QoL in the somatic and environmental domains. Nurses in the reproductive period had a higher level of satisfaction with overall QoL compared to those in the peri-menopausal period. Increased frequency and intensity of menopausal symptoms along with lower QoL were reported by nurses in the 45-55 age group in the somatic, social, mental and environmental domains [8].

Assessing bothersomeness is important, as it can help the healthcare provider understand which symptoms bother the menopausal woman more and need to be prioritised. Existing scales focusing on women's health or menopause evaluate either the frequency or severity of menopausal symptoms or the QoL associated with them. The Menopause Rating Scale [9] and Modified Kupperman Index [10] evaluate the severity and frequency of menopausal symptoms. The Greene climacteric scale evaluates the bothersomeness related to menopausal symptoms; however, it does not include symptoms

related to trophic changes and the urogenital and gastrointestinal domains [11]. Similarly, the Women's Health Questionnaire also includes questions pertaining to bothersomeness, but it does not include items related to memory, trophic changes, weight changes, and urogenital symptoms [12].

Hence, most of the available menopause-related questionnaires or scales assess the prevalence of menopausal symptoms or the QoL related to menopause. This study aimed to develop and validate a questionnaire in the English language assessing bothersomeness due to menopausal symptoms.

MATERIALS AND METHODS

A cross-sectional study was conducted among menopausal women from multiple states of India (New Delhi, Haryana, Uttar Pradesh, Gujarat, Rajasthan, Madhya Pradesh, Maharashtra, and Karnataka), India, from May 2023 to April 2025. Ethics clearance was taken from the Institutional Ethics Committee. Informed consent was obtained from all members of the expert panel and all participants in the psychometric evaluation process. Recruitment was conducted through convenience sampling, using Google Forms circulated via personal contacts and gynaecologists.

Inclusion criteria: Women aged 45-65 years who were peri-menopausal (experiencing menstrual irregularities such as prolonged/heavy bleeding intermixed with episodes of amenorrhoea) or post-menopausal (self-reported cessation of menses for at least one year) were included. Participants were required to be fluent in English and able to provide informed consent. Data collection was conducted online; therefore, internet access was necessary.

Exclusion criteria: Pre-menopausal women (with regular menstrual cycles), women with serious medical or psychiatric conditions that could interfere with questionnaire completion, and those unwilling or unable to provide consent were excluded.

Study Procedure

Data were collected using Google Forms. To ensure eligibility, the link was distributed only to confirmed menopausal women. The form included study details, eligibility criteria, and an informed consent section, and participants proceeded only after providing consent.

It was ensured that the inclusion and exclusion criteria remained uniform across the various phases of data collection. Data were collected in a phased manner - beginning with content validation and continuing with factor analysis, reliability testing and construct validation.

A literature review was conducted using PubMed, CINAHL, Cochrane, and Scopus to identify questionnaires or scales related to menopausal symptoms. It included the Menopause Rating Scale [9,13], the Women's Health Questionnaire [12], the Menopause-Specific QoL Questionnaire [14], etc. Related items from these questionnaires were retrieved. Items related to common menopausal symptoms were retrieved, with overlapping items excluded.

A focus group comprising three gynaecologists and three menopausal women was convened. Gynaecologists with more than 10 years of experience working with menopausal women were included. The mean clinical experience of the gynaecologists was 26.83 years. The menopausal women selected had a mean duration of menopause of 15.3 years, had good educational qualifications, and were willing to share their experiences and problems.

Two group discussions and individual interviews were conducted. During the Focus Group Discussions (FGDs), the primary investigator acted as the moderator.

In the first group discussion, the moderator introduced all the members, explained the purpose of the focus group, and initiated discussion on menopause and menopausal symptoms. Opinions of each member were heard and noted by the moderator. After each member had voiced their opinions, the platform was opened to group discussion on the difficulties faced by menopausal women, the various symptoms they experience, how bothersomeness differs according to each individual's tolerance, and how it may change over the months of menopause. Discussion was also done on the framing of the questionnaire, instructions that should be provided for the easier understanding of the general population and scoring of the items. The focus group discussion was closed after a summary of the group's points.

Individual interviews were conducted with each focus group member after the first focus group discussion, to allow each member another opportunity to express their views independently. None of the members provided any other significant additional information during this.

After the FGDs and individual interviews, the investigators prepared a primary draft of the questionnaire. This primary draft was discussed by the focus group in the second discussion. A few grammatical and linguistic changes were made to this primary draft. 'Memory problems' was changed to 'Difficulty remembering things', 'Breast pain' was changed to 'Breast pain or tenderness', 'Flatulence' was changed to 'Flatulence/Gas' and 'Palpitations' was changed to 'Heart beating quickly or strongly' for easier understanding. The order of some of the items was changed, and a preliminary questionnaire comprising 46 items was prepared after refining and organising them sequentially. Once all focus group members had approved the preliminary draft, it was finalised for the next phase. Data collection and analysis were conducted concurrently. Saturation was considered achieved when no new codes or themes emerged from two consecutive interviews.

After the first round, 13 items were eliminated, and the questionnaire for round two comprised 33 items. In the first round of review, the panel members suggested simplifying the scoring by reducing the Likert scale to four points and changing the score to 'not bothered at all', 'mildly bothered', 'moderately bothered' and 'extremely bothered' as understanding the five-point Likert scale would be difficult for the target population. Literature also suggests that forced-choice scales (even-numbered scales) reduce central tendency bias and are appropriate when a neutral option is undesirable or ambiguous [15,16]. A few words in the questionnaire were changed as suggested by the panel members, and the second draft was prepared. In the second round of review, no changes were suggested; hence, the questionnaire was accepted and finalised.

A panel comprising nine members - experts (three gynaecologists, two physiotherapists working in the field of women's health and one psychologist) and three post-menopausal women was created for the purpose of determining content validity.

Boateng GO et al., recommend using a combination of experts and target population users to determine content validity [16]. The experts were first called by telephone to confirm their participation in the study. Only after they approved were further details provided. Each panel member received a cover letter describing the questionnaire, explaining the need for participation, and providing clear instructions for rating the items, along with a copy of the preliminary questionnaire. A critical appraisal sheet was also provided to rate the items. Each item of the preliminary questionnaire was assessed for relevance, clarity, and essentiality. Relevance and clarity were assessed on a four-point Likert scale.

Recommendations for any other changes were also welcomed by the authors from the panel members for the qualitative assessment of content validity. This encompassed observations regarding the grammar of the text in the questionnaire, the sequence of the items, scoring, and the use of appropriate and correct words. Communication with panel members was made via telephone calls, e-mail and personal contact.

After content validity analysis, the final questionnaire comprised 33 items, scored on a four-point Likert scale, and was named the MSBQ.

Quantitative assessment of content validity involved the determination of CVR, CVI and Kappa values. CVR was determined for the entire scale, whereas CVI was computed for each item of the whole scale. I-CVI values of 0.78 or higher were considered evidence of good content validity [17]. Kappa values above 0.74 are considered excellent, between 0.60 and 0.74 good and 0.40 and 0.59 fair [18,19]. Psychometric evaluation was conducted using factor analysis and reliability testing, including internal consistency, test-retest reliability, and construct validity.

Conventional Guidelines recommend a sample size of ≥300 participants for robust factor analysis, or at least a minimum item-to-response ratio of 1:3 [20,21]. In the current study, 33 items were retained. Based on this recommendation, for the 33 items of the MSBQ, a minimum of 99 participants was required; however, due to practical challenges in recruiting English-fluent menopausal women aged 45-65 years across multiple states in India, data were collected from 106 participants.

Exploratory factor analysis was conducted using principal axis factoring with direct Oblimin rotation. To assess the internal consistency of MSBQ, Cronbach's alpha was computed for each subscale of the questionnaire [22,23].

STATISTICAL ANALYSIS

Exploratory factor analysis was conducted using the Statistical Package for the Social Sciences (SPSS), version 20.0. Principal axis factoring was employed as the extraction method, with Direct Oblimin rotation applied since the factors were expected to be correlated. The number of factors was determined using the Kaiser-Guttman criterion (eigenvalues >1). Reliability was assessed through internal consistency (Cronbach's alpha) and test-retest reliability {Spearman's rho and Intraclass Correlation Coefficient (ICC)}. Validity was examined by evaluating convergent and divergent validity using Spearman's rho correlations.

RESULTS

Thirteen items had CVR values less than 0.78 and hence, were eliminated. The CVR score for the whole instrument was 0.74. I-CVI values ranged from 0.67 to 1.00. Thirty items had an I-CVI value of 1.00, eleven had 0.89, four had 0.78, and one (Item 14) had 0.67. S-CVI/UA was recorded as 0.65, and S-CVI/Ave was recorded as 0.95. The Kappa values calculated based on the available data ranged from 0.34 to 0.996. Average clarity scores ranged from 2.78 to 4. The overall clarity score was 3.805 [Table/Fig-1].

Item number	CVR value	I-CVI value	Pc value	Kappa value	Average clarity scores	Interpretation
1	0.78	1.00	0.002	0.996	4	Retained
2	0.33	0.78	0.07	0.64	3.44	Eliminated
3	0.33	1.00	0.002	0.996	3.89	Eliminated
4	0.78	1.00	0.002	0.996	4	Retained
5	0.33	0.89	0.014	0.86	3.67	Eliminated
6	0.55	0.78	0.07	0.64	3.67	Eliminated
7	0.78	1.00	0.002	0.996	3.89	Retained
8	1.00	1.00	0.002	0.996	3.89	Retained
9	0.78	0.78	0.07	0.64	3.89	Retained
10	0.55	1.00	0.002	0.996	3.55	Eliminated
11	0.78	1.00	0.002	0.996	2.78	Retained
12	0.78	1.00	0.002	0.996	4	Retained
13	0.33	1.00	0.002	0.996	3.89	Eliminated
14	0.11	0.67	0.164	0.34	2.89	Eliminated
15	0.55	0.89	0.014	0.86	3.89	Eliminated
16	0.78	1.00	0.002	0.996	4	Retained
17	0.78	1.00	0.002	0.996	3.89	Retained
18	0.78	0.89	0.014	0.86	4	Retained
19	1.00	0.89	0.014	0.86	4	Retained
20	1.00	0.89	0.014	0.86	4	Retained
21	0.78	0.78	0.07	0.64	3.44	Retained
22	1.00	1.00	0.002	0.996	4	Retained
23	1.00	1.00	0.002	0.996	4	Retained
24	0.78	1.00	0.002	0.996	4	Retained
25	0.78	1.00	0.002	0.996	3.89	Retained
26	0.78	1.00	0.002	0.996	3.89	Retained
27	1.00	1.00	0.002	0.996	4	Retained
28	1.00	1.00	0.002	0.996	4	Retained
29	0.78	0.89	0.014	0.86	4	Retained
30	0.78	0.89	0.014	0.86	4	Retained
31	0.78	0.89	0.014	0.86	3.67	Retained
32	0.78	1.00	0.002	0.996	3.78	Retained
33	0.55	1.00	0.002	0.996	3.67	Eliminated
34	0.55	0.89	0.014	0.86	3.67	Eliminated
35	0.78	0.89	0.014	0.86	3.67	Retained
36	0.78	1.00	0.002	0.996	3.89	Retained
37	0.78	1.00	0.002	0.996	3.89	Retained
38	0.78	1.00	0.002	0.996	3.78	Retained
39	0.55	0.89	0.014	0.86	3.55	Eliminated
40	0.55	1.00	0.002	0.996	3.78	Eliminated
41	0.55	1.00	0.002	0.996	3.89	Eliminated
42	0.78	1.00	0.002	0.996	3.89	Retained
43	0.78	1.00	0.002	0.996	3.67	Retained
44	0.78	1.00	0.002	0.996	3.78	Retained
45	1.00	1.00	0.002	0.996	3.78	Retained
46	1.00	1.00	0.002	0.996	3.89	Retained

[Table/Fig-1]: Content validity analysis for each item of the Menopausal Symptom Bothersomeness Questionnaire (MSBQ).
*Elimination of any item was based on a combined decision rule. If any item had values lower than the cut-off values for any of the statistical values (CVR, I-CVI and Kappa), the item was eliminated

For factor analysis, data were collected from a sample of 106 menopausal women. Of these 106, 52 females were peri-menopausal, and 54 were post-menopausal. Mean age of the sample was 51.13±5.067 years. Mean ages of peri-menopausal and post-menopausal women were 47.83±2.185 years and 54.31±5.024 years, respectively. Bartlett's test of sphericity was significant at <0.001, suggesting patterned relationships in the data. Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy value was 0.843, suggesting that the data were suitable for factor analysis. Items were considered representative of a component if they loaded ≥0.50. Principal axis factoring revealed a nine-factor solution (eigenvalues >1) for the 32 items. The extracted factors together explained 68.59% of the cumulative variance. Initial eigenvalues accounted for 76.87% of the variance before extraction [Table/Fig-2].

Exploratory factor analysis using principal axis factoring with oblimin rotation yielded a nine-factor solution. The adequacy of the model was supported by the Reproduced Correlation Matrix, which showed that only 8% of the non redundant residuals had absolute values greater than 0.05, indicating that the factor model provided a good fit to the data. The structure matrix demonstrated that all items loaded ≥0.50 on their respective factors, with two exceptions. Item 24 ("Heart beating quickly or strongly," loading=0.499) was removed due to insufficient loading. Item 15 ("Dryness of vagina," loading=0.464) was retained despite its lower loading, as expert panel discussions and participant feedback emphasised its clinical importance in representing menopausal

Total variance explained							
	Initial Eigen values			Extraction sums of squared loadings			Rotation sums of squared loadings
Factor	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	12.517	37.931	37.931	12.214	37.012	37.012	6.594
2	2.410	7.302	45.234	2.162	6.551	43.563	4.450
3	2.183	6.615	51.849	1.919	5.815	49.379	6.184
4	1.727	5.234	57.083	1.436	4.353	53.731	6.585
5	1.645	4.986	62.069	1.298	3.935	57.666	3.962
6	1.395	4.229	66.297	1.075	3.258	60.923	4.949
7	1.366	4.139	70.436	1.052	3.188	64.112	4.940
8	1.107	3.355	73.792	.782	2.369	66.481	3.759
9	1.016	3.079	76.870	.695	2.106	68.587	1.996

[Table/Fig-2]: Exploratory factor analysis results (principal axis factoring).
*Note: Eigenvalues >1 retained. Initial eigenvalues show total variance before extraction; extraction sums of squared loadings show variance explained by the retained factors; rotation redistributes variance across factors for interpretability but does not change the total variance explained

symptoms. Negative loadings were observed for Items 25-29, 32, and 33, reflecting inverse associations with their respective factor. This pattern is consistent with the oblique rotation method, which allows correlated factors and can produce both positive and negative loadings. Items 30 and 31, however, loaded positively on the same factor [Table/Fig-3].

Structure matrix									
	Factor								
	1	2	3	4	5	6	7	8	9
Item 6	**0.783**								
Item 2	**0.745**								
Item 5	**0.737**								
Item 3	**0.695**								
Item 7	**0.676**								
Item 4	**0.639**								
Item 1	**0.611**								
Item 17		**0.916**							
Item 18		**0.904**							
Item 16		**0.782**							
Item 15		**0.464**							
Item 21			**0.911**						
Item 22			**0.873**						
Item 23			**0.872**						
Item 28				**-0.885**					
Item 27				**-0.783**					
Item 26				**-0.738**					
Item 25				**-0.718**					
Item 29				**-0.677**					
Item 13					**0.748**				
Item 11					**0.707**				
Item 14					**0.687**				
Item 12					**0.569**				
Item 19						**0.867**			
Item 20						**0.832**			
Item 8							**0.820**		
Item 9							**0.731**		
Item 10							**0.681**		
Item 24							**0.499**		
Item 33								**-0.820**	
Item 32								**-0.758**	
Item 31									**0.667**
Item 30									**0.624**

Extraction Method: Principal Axis Factoring; Rotation Method: Oblimin with Kaiser Normalisation.

[Table/Fig-3]: Nine-factor solution - factor loadings and structure matrix.
Item numbers reflect the original questionnaire sequence. Items eliminated during content validity are not included in this table

The final draft of the MSBQ comprised nine factors and 32 items [Table/Fig-4]. Each factor represents a distinct symptom domain, with items grouped according to their strongest loadings in the exploratory factor analysis.

Factors	Variables
Sleep	Items 32 and 33
Sexual	Items 16, 17 and 18
Gastrointestinal	Items 21, 22 and 23
Psychological	Items 25, 26, 27, 28 and 29
Physical	Items 1, 2, 3, 4, 5, 6 and 7
Vasomotor	Items 19 and 20
Urogenital	Items 11, 12, 13 and 14,15
Trophic changes	Items 8, 9 and 10
Cognitive	Items 30 and 31

[Table/Fig-4]: Menopausal Symptom Bothersomeness Questionnaire (MSBQ).

Internal consistency of the MSBQ was evaluated in a sample of 45 women (24 peri-menopausal, 21 post-menopausal; mean age 52.6±5.91 years). Cronbach's alpha values across the nine factors ranged from 0.635 to 0.921, indicating acceptable to excellent reliability. The sexual (α=0.921), sleep (α=0.840), vasomotor (α=0.850), and cognitive (α=0.818) domains demonstrated strong internal consistency. The psychological (α=0.798), trophic changes (α=0.785), and gastrointestinal (α=0.758) domains showed good reliability. The physical (α=0.675) and urogenital (α=0.635) domains were somewhat lower but still within the acceptable range [Table/Fig-5].

Factors	Number of items	Cronbach's alpha
Sleep	2	0.840
Sexual	3	0.921
Gastrointestinal	3	0.758
Psychological	6	0.798
Physical	7	0.675
Vasomotor	2	0.850
Urogenital	5	0.635
Trophic changes	3	0.785
Cognitive	2	0.818

[Table/Fig-5]: Cronbach's alpha for evaluation of internal consistency.

Test-retest reliability was assessed in 22 women (7 peri-menopausal and 15 post-menopausal; mean age 54.91±5.38 years) at baseline and after one week. ICC and Spearman's rho values across the nine factors ranged from 0.730 to 0.936, indicating good to excellent stability over time. The physical (ICC=0.925, rho=0.871), gastrointestinal (ICC=0.935, rho=0.918), and urogenital (ICC=0.889, rho=0.936) domains demonstrated particularly strong reproducibility. The psychological (ICC=0.745, rho=0.747) and sleep (ICC=0.730, rho=0.641) domains showed moderate but acceptable reliability [Table/Fig-6].

The analysis of 107 females (mean age 51.16±4.67 years) showed that the MSBQ demonstrated strong convergent validity with related domains of Greene Menopausal Scale (GMS) and Menopause-Specific Quality of Life Questionnaire (MENQOL). Spearman's rho values ranged from 0.752 to 0.923, indicating consistently high correlations. The strongest association was observed for urogenital symptoms (ρ=0.923, MENQOL), while other domains such as vasomotor (ρ=0.864), sexual (ρ=0.855), psychological (ρ=0.853), and sleep (ρ=0.849) also showed robust convergence [Table/Fig-7].

As shown in [Table/Fig-8,9], MSBQ components demonstrated moderate correlation with other components of the questionnaire (r=0.3-0.5), whereas most components of MSBQ exhibited strong correlation with non corresponding components of GMS or MENQOL (r>0.6).

Factors	Spearman's rho	Intra-class correlation coefficient
Physical	0.871	0.925
Trophic changes	0.891	0.889
Urogenital	0.936	0.889
Sexual	0.860	0.891
Vasomotor	0.834	0.762
Gastrointestinal	0.918	0.935
Psychological	0.747	0.745
Cognitive	0.820	0.817
Sleep	0.641	0.730

[Table/Fig-6]: Test–retest reliability of the Menopausal Symptom Bothersomeness Questionnaire (MSBQ) {Spearman's rho and Intraclass Correlation Coefficient (ICC)}
Note: ICC values ≥0.70 were considered acceptable indicators of reliability.

MSBQ	GMS/MENQOL	Spearman's rho
Physical	Physical (GMS)	0.752
Psychological	Psychological (GMS)	0.853
Sexual	Loss of interest in sex (GMS)	0.855
Vasomotor	Vasomotor (GMS)	0.864
Sleep	Difficulty in sleeping (GMS)	0.849
Cognitive	Difficulty in concentrating (GMS)	0.770
Trophic changes	Trophic changes (MENQOL)	0.836
Gastrointestinal	Gastrointestinal (MENQOL)	0.815
Sleep	Difficulty sleeping (MENQOL)	0.849
Cognitive	Experiencing poor memory (MENQOL)	0.770
Urogenital	Urogenital (MENQOL)	0.923

[Table/Fig-7]: Convergent validity of MSBQ with related components of GMS and MENQOL.

Factor	Factor	Spearman's rho
Physical	Gastrointestinal	0.450
Trophic changes	Urogenital	0.435
Urogenital	Vasomotor	0.479
Sexual	Gastrointestinal	0.342
Vasomotor	Gastrointestinal	0.423
Gastrointestinal	Psychological	0.474
Psychological	Vasomotor	0.459
Cognitive	Vasomotor	0.381
Sleep	Gastrointestinal	0.361

[Table/Fig-8]: Divergent validity between different components of MSBQ.

MSBQ Factor	GMS/MENQOL component	Spearman's rho
Sexual	Vasomotor (GMS)	0.421
Psychological	Physical (GMS)	0.578
Cognitive	Physical (GMS)	0.545
Physical	Trophic changes (MENQOL)	0.572
Trophic changes	Gastrointestinal (MENQOL)	0.380
Urogenital	Trophic changes (MENQOL)	0.542
Vasomotor	Trophic changes (MENQOL)	0.539
Gastrointestinal	Physical (MENQOL)	0.458
Sleep	Urogenital (MENQOL)	0.283

[Table/Fig-9]: Divergent validity of MSBQ with non corresponding components of GMS and MENQOL.

DISCUSSION

The MSBQ is a unique PROM that addresses an important gap in menopausal health assessment. Existing scales primarily evaluate symptom prevalence or severity, whereas the MSBQ focuses on

bothersomeness, thereby enabling healthcare providers to identify which symptoms most interfere with daily life and require prioritisation in management.

The development process was rigorous, combining literature review, FGDs, and expert panel validation. Both item-level and scale-level content validity indices were reported, which is an uncommon practice. While the universal agreement method (S-CVI/UA=0.65) suggested moderate validity, the average method (S-CVI/Ave=0.95) indicated excellent validity. Reporting both indices provides a balanced perspective, as S-CVI/UA may underestimate validity in larger panels, whereas S-CVI/Ave may overestimate it [24]. The inclusion of nine experts aligns with recommended practices [23], and the final MSBQ retained only items with an I-CVI ≥0.78, consistent with accepted thresholds for excellent content validity [17,23].

Exploratory factor analysis produced a nine-factor solution with 32 items, supported by strong sampling adequacy (KMO=0.843) and a significant Bartlett's test. Although larger samples are often recommended, Comrey AL and Lee HB classify samples in the range of 100-200 as "fair" [25], and an item-to-response ratio of 1:3 has also been suggested [25]. With 33 items, a minimum of 99 participants was required, and the 106 responses collected satisfied this criterion. Recent methodological reviews also emphasise that adequacy depends more on communalities, factor loadings, and sampling indices than on absolute sample size, supporting the validity of the present sample [26]. The decision to retain the item on vaginal dryness, despite its lower loading, was based on clinical relevance, given its high prevalence among Indian menopausal women (57-77%) [27,28].

Elimination of the items was based on a combined decision rule. If any item had values lower than the cut-off values for any of the statistical values (CVR, I-CVI and Kappa), the item was eliminated. Item 3 had a low CVR value despite having a high I-CVI value, and was eliminated. Item 9 had good CVR, I-CVI and Kappa values and hence was retained. This highlights the importance of balancing statistical criteria with clinical utility in questionnaire development.

The decision to reduce the Likert scale from five points to four points was based on both expert panel feedback and methodological guidance. Lynn MR (1986) advised that a minimum of three experts is sufficient for content validation, while more than ten is unnecessary, and recommended the use of a 4-point ordinal scale to avoid a neutral midpoint. Item ratings are typically conducted on a 4-point scale. Although 3- or 5-point scales may be considered, Lynn MR advocated for a 4-point scale to minimise central tendency bias and ensure clearer judgments of item relevance. This rationale aligned with our panel's recommendation, and therefore the MSBQ adopted a 4-point Likert scale ('not bothered at all' to 'extremely bothered' [29].

During exploratory factor analysis, some items demonstrated negative loadings. This outcome is consistent with the use of oblique rotation, which permits correlated factors and can yield both positive and negative coefficients. Rather than discarding these items solely on statistical grounds, we retained them based on their clinical relevance and strong endorsement from the expert panel. This decision underscores the importance of balancing psychometric rigour with clinical judgment when developing PROM.

Cronbach's alpha ranged from 0.635 to 0.921, indicating acceptable to excellent internal consistency across domains. The physical (α=0.675) and urogenital (α=0.635) subscales yielded lower alpha values; while these fall below the conventional 0.70 threshold [30], these domains were retained due to their clinical relevance and theoretical importance.

Future studies should aim to refine item composition to improve reliability. Test-retest reliability ranged from moderate to excellent, confirming stability. Construct validity was established through strong convergent correlations with the Greene Climacteric Scale (GCS) and MENQOL, while divergent validity correlations were appropriately weaker. Compared with the GCS, which explained 48.6% of variance in an Indian sample [31], the MSBQ accounted for 68.59%, suggesting a more comprehensive factor structure. Internal consistency of the MSBQ was comparable to that of the GCS (Cronbach's alpha 0.91) [30], further supporting its reliability. Importantly, the MSBQ complements MENQOL by focusing on bothersomeness rather than QoL alone, thereby filling a unique gap in menopausal assessment.

Limitation(s)

The study had a few limitations. Though the sample size used for factor analysis and reliability testing is justified, it is relatively small. Future studies may be conducted with a larger sample size. The questionnaire was developed in the English language, whereas the population residing in India primarily speaks Hindi and other regional languages. Studies on the translation of the MSBQ into regional languages should be done. The study included only exploratory factor analysis. Studies focusing on confirmatory factor analysis should also be conducted. Future studies can also focus on the determination of responsiveness to change over time.

CONCLUSION(S)

The MSBQ is a PROM developed in English to assess the bothersomeness of menopausal symptoms. It was designed using a mixed-methods approach and demonstrated strong content validity, good internal consistency, and stable test-retest reliability. Construct validity was also established, confirming that the MSBQ captures relevant aspects of menopausal symptomatology. The MSBQ provides healthcare providers with a practical tool to identify the most distressing symptoms and guide individualised management strategies, thereby supporting improved QoL for menopausal women.

Acknowledgement

The authors would like to thank all the members in the review panel for their support, advice and time. The authors extend their gratitude to Dr. Megha Sheth, Dr. Atul Munshi, Dr. Anjana Chauhan and Dr. Janaki Desai for their help and guidance.

REFERENCES

[1] Menopause Terminology [Internet]. International Menopause Society. [cited 2023 Aug 3]. Available from: https://www.imsociety.org/education/menopause-terminology/.
[2] Talaulikar V. Menopause transition: Physiology and symptoms. Best Pract Res Clin Obstet Gynaecol. 2022;81:03-07.
[3] Melby MK, Sievert LL, Anderson D, Obermeyer CM. Overview of methods used in cross-cultural comparisons of menopausal symptoms and their determinants: Guidelines for Strengthening the Reporting of Menopause and Aging (STROMA) studies. Maturitas. 2011;70(2):99-109.
[4] Santoro N, Roeca C, Peters BA, Neal-Perry G. The menopause transition: Signs, symptoms, and management options. J Clin Endocrinol Metab. 2021;106(1):1-15.
[5] Rathnayake N, Lenora J, Alwis G, Lekamwasam S. Prevalence and severity of menopausal symptoms and the quality of life in middle-aged women: A study from Sri Lanka. Nurs Res Pract. 2019;2019:2081507.
[6] Whiteley J, DiBonaventura M daCosta, Wagner JS, Alvir J, Shah S. The impact of menopausal symptoms on quality of life, productivity, and economic outcomes. J Womens Health. 2013;22(11):983-90.
[7] Szadowska-Szlachetka Z, Stasiak E, Leziak A, Irzmańska-Hudziak A, Łuczyk M, Stanisławek A, et al. Intensity of menopausal symptoms and quality of life in climacteric women. Prz Menopauzalny (Menopause Review / Przegląd Menopauzalny). 2020;18(4):217-21.
[8] Kupcewicz E, Drężek–Skrzeszewska M, Roszuk W, Jóźwik M. Relationship between frequency and intensity of menopausal symptoms and quality of life among polish nurses. Risk Manag Healthc Policy. 2021;14:97-107.
[9] Vadugu S, Lanke V. The Menopause Rating Scale (MRS) in Indian women. J Evid Based Med Healthc. 2015;2(12):01-03.
[10] Tao M, Shao H, Li C, Teng Y. Correlation between the modified Kupperman Index and the Menopause Rating Scale in Chinese women. Patient Prefer Adherence. 2013;7:223-29.
[11] Greene JG. Constructing a standard climacteric scale. Maturitas. 2008;61(1):78-84.

[12] Hunter M. The Women's Health Questionnaire (WHQ): The development, standardization and application of a measure of mid-aged women's emotional and physical health. Qual Life Res. 2012;9:733-38. Doi: 10.1023/A:1008973822876.
[13] Heinemann LA, Potthoff P, Schneider HP. International versions of the Menopause Rating Scale (MRS). Health Qual Life Outcomes. 2003;1(1):28.
[14] Hilditch JR, Lewis J, Peter A, van Maris B, Ross A, Franssen E, et al. A menopause-specific quality of life questionnaire: Development and psychometric properties. Maturitas. 1996;24(6):161-75.
[15] DeVellis R. Scale development: Theory and applications. 4th ed. Thousand Oaks: SAGE Publications, Inc; 2017.
[16] Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: A primer. Front Public Health. 2018;6:149.
[17] Kishore K, Jaswal V, Kulkarni V, De D. Practical guidelines to develop and evaluate a questionnaire. Indian Dermatol Online J. 2021;12(2):266-75.
[18] Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Res Nurs Health. 2007;30(4):459-67.
[19] Zamanzadeh V, Ghahramanian A, Rassouli M, Abbaszadeh A, Alavi-Majd H, Nikanfar AR. Design and implementation content validity study: Development of an instrument for measuring patient-centered communication. J Caring Sci. 2015;4(2):165-78.
[20] Bujang MA, Omar ED, Foo DHP, Hon YK. Sample size determination for conducting a pilot study to assess reliability of a questionnaire. Restor Dent Endod. 2024;49(1):e3.
[21] Gunawan J, Marzilli C, Aungsuroch Y. Establishing appropriate sample size for developing and validating a questionnaire in nursing research. Belitung Nurs J. 2021;7(5):356-60.
[22] Comrey AL, Lee HB. A first course in factor analysis. 2nd ed. New York: Psychology Press; 2013. 442 p.
[23] Polit DF, Beck CT. The content validity index: Are you sure you know what's being reported? critique and recommendations. Res Nurs Health. 2006;29(5):489-97.
[24] Rodrigues IB, Adachi JD, Beattie KA, MacDermid JC. Development and validation of a new tool to measure the facilitators, barriers and preferences to exercise in people with osteoporosis. BMC Musculoskelet Disord. 2017;18(1):540. Doi: 10.1186/s12891-017-1914-5.
[25] White M. Sample size in quantitative instrument validation studies: A systematic review of articles published in Scopus, 2021. Heliyon. 2022;8(12):e12223. Doi: 10.1016/j.heliyon.2022.e12223. PMID: 36568672; PMCID: PMC9768294.
[26] Goretzko D, Pham TTH, Bühner M. Exploratory factor analysis: Current use, methodological developments and recommendations for good practice. Current Psychology. 2021;40:3510-21. Doi: 10.1007/s12144-019-00300-2.
[27] Thomas T, Kamath N, Kumar A, D'Silva F, Shetty PK. Prevalence and severity of menopausal symptoms among women – a community-based cross-sectional study. Online J Health Allied Sci. 2021;20(1):2.
[28] Ulhe SC, Acharya N, Vats A, Singh A. Study of vulvovaginal atrophy and genitourinary syndrome of menopause and its impact on the quality of life of postmenopausal women in central India. Cureus. 2024;16(2):e54802.
[29] Lynn MR. Determination and quantification of content validity. Nurs Res. 1986;35(6):382-85. PMID: 3640358.
[30] 30. Taber KS. The use of Cronbach's alpha when developing and reporting research instruments in science education. Res Sci Educ. 2018;48:1273-96. https://doi.org/10.1007/s11165-016-9602-2.
[31] Chattha R, Kulkarni R, Nagarathna R, Nagendra HR. Factor analysis of Greene's Climacteric Scale for Indian women. Maturitas. 2008;59(1):22-27.

PARTICULARS OF CONTRIBUTORS:
1. PhD Scholar, Department of Physiotherapy, Gujarat University, Ahmedabad, India; Assistant Professor, LJ Institute of Physiotherapy, LJ University, Ahmedabad, India.
2. Principal, SBB College of Physiotherapy, Ahmedabad, India.

NAME, ADDRESS, E-MAIL ID OF THE CORRESPONDING AUTHOR:
Priyasingh Bhagirathsingh Rangey,
C-1004, Satyamev Vista, Near Gota Cross Roads,
Gota, Ahmedabad-382481, Gujarat, India.
E-mail: Priya_singh9192@yahoo.in

PLAGIARISM CHECKING METHODS: [Jain H et al.]
- Plagiarism X-checker: Dec 21, 2025
- Manual Googling: Apr 11, 2026
- iThenticate Software: Apr 14, 2026 (1%)

ETYMOLOGY: Author Origin

EMENDATIONS: 8

AUTHOR DECLARATION:
- Financial or Other Competing Interests: None
- Was Ethics Committee Approval obtained for this study? Yes
- Was informed consent obtained from the subjects involved in the study? Yes
- For any images presented appropriate consent has been obtained from the subjects. NA

Date of Submission: **Dec 04, 2025**
Date of Peer Review: **Dec 27, 2025**
Date of Acceptance: **Apr 16, 2026**
Date of Publishing: **Aug 01, 2026**