

## Psychometric attributes of the Cervantes short-form questionnaire for measuring health-related quality of life in menopausal women



Pluvio J. Coronado<sup>a,\*</sup>, Rafael Sánchez-Borrego<sup>b</sup>, Miguel A. Ruiz<sup>c</sup>, Laura Baquedano<sup>d</sup>, Sonia Sánchez<sup>e</sup>, Cristina Argudo<sup>f</sup>, Mariela Fernández-Abellán<sup>g</sup>, Silvia González<sup>h</sup>, Eva Iglesias<sup>i</sup>, Jackie Calleja<sup>j</sup>, Jesus Presa<sup>k</sup>, Alfonso Duque<sup>l</sup>, Fernando Ruiz<sup>m</sup>, Borja Otero<sup>n</sup>, Javier Rejas<sup>o</sup>

<sup>a</sup> Department of Obstetrics and Gynaecology, Hospital Clínico San Carlos, Madrid, Spain

<sup>b</sup> Clínica DIATROS, Barcelona, Spain

<sup>c</sup> School of Psychology, Universidad Autónoma de Madrid, Spain

<sup>d</sup> Departments of Obstetrics and Gynaecology of Hospital Universitario Miguel Servet, Zaragoza, Spain

<sup>e</sup> Hospital Sagrado Corazón, Barcelona, Spain

<sup>f</sup> Clínica Begoña, Gijón, Asturias, Spain

<sup>g</sup> Clínica Quirón, Málaga, Spain

<sup>h</sup> Gabinete Médico Velazquez, Madrid, Spain

<sup>i</sup> Hospital de Valme, Sevilla, Spain

<sup>j</sup> Clínica Quirón, Madrid, Spain

<sup>k</sup> Hospital Clínico de Jaén, Spain

<sup>l</sup> Clínica Ruber, Madrid, Spain

<sup>m</sup> Hospital Clínico Universitario, Valencia, Spain

<sup>n</sup> Hospital Universitario de Cruces, Baracaldo, Vizcaya, Spain

<sup>o</sup> HEOR Department, Pfizer SLU, Alcobendas, Spain

### ARTICLE INFO

#### Article history:

Received 21 May 2015

Received in revised form

14 September 2015

Accepted 24 October 2015

#### Key words:

Cervantes scale

Abridged

Menopause

Health-related quality of life

### ABSTRACT

**Objective:** To analyse the psychometric properties of the Cervantes scale short-form (SF) in the peri- and post-menopausal periods.

**Methods:** Outpatients women 45–65 years with menstrual problems associated with the climacteric syndrome were analysed. Original and SF versions of the Cervantes scale were administered along with the EuroQol-5D (EQ-5D) and work productivity and activity impairment questionnaire (WPAI) scales. Conceptual model, burden of administration, feasibility, reliability, criteria validity and construct validity were assessed.

**Results:** 317 women [55.7 ± 5.3 years (mean ± standard deviation)] were recruited: 75.4% were post- and 22.3% were peri-menopausal. The Cervantes-SF was completed in 2.5 ± 1.6 min, and 86% answered all items. Cronbach's  $\alpha$  was 0.820, and ranged from 0.510 (Aging) to 0.918 (Vasomotor Symptoms) for individual dimensions. The scale structure matched the structure of the original version,  $\chi^2/(\text{degrees of freedom}) = 3.6$ , Comparative Fit Index = 0.848, Tucker–Lewis Index = 0.850, and root mean square error of approximation = 0.099, although differences were found between sexual activity statuses. Criteria validity was good ( $r = 0.890$ ), concurrent validity was congruent with *a priori* hypothesis using either the EQ-5D or the WPAI scales. The scale discriminated significantly the severity of both vasomotor and genital climacteric associated symptoms.

**Conclusion:** The Cervantes-SF has shown good psychometric properties for measuring Health related quality of life in peri- and post-menopausal women who regularly attended gynaecology clinics in Spain.

© 2015 Elsevier Ireland Ltd. All rights reserved.

## 1. Introduction

Menopause is defined as the permanent cessation of menstruation, confirmed after 12 consecutive months of amenorrhoea. It can be natural, occurring without other pathological or physiolog-

\* Corresponding author at: Servicio de Obstetricia y Ginecología, Hospital Clínico San Carlos, C/Prof. Martín Lagos, s/n 28040 Madrid, Spain.  
E-mail address: [pcoronado@gmail.com](mailto:pcoronado@gmail.com) (P.J. Coronado).

ical causes, or induced by surgery, radiation, or drugs [1]. Natural menopause is the result of the loss of ovarian follicular activity and occurs at about 50–52 years of age with ample variations [1]. During this period, women experience important physical and psychological changes as a consequence of decline in oestrogenic activity [2,3]. It is worth noting that the effects of menopause-related changes on health-related quality of life (HRQoL) are closely related to personal and sociocultural characteristics, which in turn decisively influence how each woman perceives many of these changes [4,5]. HRQoL assessment has become an essential component when studying the effects of menopause on well-being as well as when evaluating the benefits of hormonal treatments or any other therapies used at this stage of a woman's life [6,7].

The Cervantes scale is a self-administered questionnaire designed to measure HRQoL in peri- and post-menopausal women [8]. The original questionnaire consisting of 31 items was reduced to a 16-item short-form version (Cervantes-SF) [9] preserving the original structure with four first-order dimensions and three second-order dimensions: menopause and health (vasomotor symptoms, health, and aging), psychological, sexuality, and couple relations. While first order dimensions are directly assessed by the corresponding indicator-items, a second order dimension aggregates several related first-order dimensions (three in our case) in a common score (see Fig. 1). Along with the shortening of the Cervantes scale, a rescaling of correction method was proposed, converting the original scores (computed by direct addition of points obtained in the corresponding item Likert scales) into a 0–100 scale, where 0 means no impact of menopausal symptoms on woman's HRQoL and 100 implies the strongest possible impact on HRQoL [10]. Although appropriate for use in clinical trials and other health studies [8,9], the 31-item version is time-consuming and thus difficult to use on a routine basis in daily medical practice. Following the usual guidelines for development of multidimensional instruments [10,11], the original version was reduced by identifying those items which could be discarded without deteriorating the original psychometric properties [12]. Nevertheless, psychometric properties were computed by extracting the best candidate items from the extended version, previous reported values could be biased and we ignored what properties the instrument would have when used on a standalone basis [13]. It could be the case that some carriage effect could help to answer questions when they are to be answered in the presence of the discarded items or that an item could gain meaning from the surrounding items. The aim of this study was to assess the psychometric attributes of the conceptual and measurement model and the reliability and validity of the 16-item Cervantes questionnaire (Cervantes-SF) when presented in its definite form.

## 2. Methods

### 2.1. Design and sample of women

A non-interventional cross-sectional validation study was designed, with a portion of the recruited sample being visited a second time for test-retest reliability assessment (see below). The present work was carried out using a sample of women recruited at random in gynaecology outpatient clinics all over Spain. The women were recruited by a collaborative group, composing of 13 gynaecology specialists from the Spanish Menopause Society (SMS), formed to validate the short form of the Cervantes Scale. Recruitment was competitive until the pre-specified sample size of the study was reached or the entire enrolment period ended, whichever came first. The following inclusion criteria were applied: women, between 45 and 65 years of age, in the peri-menopausal stage due to any menstrual disorder associated with the climac-

teric syndrome or the post-menopausal stage with one or more years of menstrual cessation, being able to understand and answer written questionnaires in Spanish, and having signed the written informed consent. Exclusion criteria were as follows: presenting regular menstruation, suffering from cancer and being under treatment, suffering from a current severe medical condition or having had one that was resolved within the three months prior to recruitment (autoimmune diseases such as lupus, heart disease, severe bronchial disorders, renal or respiratory failure, advanced chronic liver disease, etc.), and receiving psychopharmacologic treatment for a major psychiatric condition.

Sample size was determined following the rule of thumb of four to seven subjects per item in the questionnaire, and also taking into account the minimum size needed to carry out a confirmatory factor analysis, which required at least ten subjects per item of the questionnaire [12]. The minimum sample calculation was increased by 15% to protect the study from data loss. Nevertheless, all subjects enrolled during the enrolment period were included in the analysis.

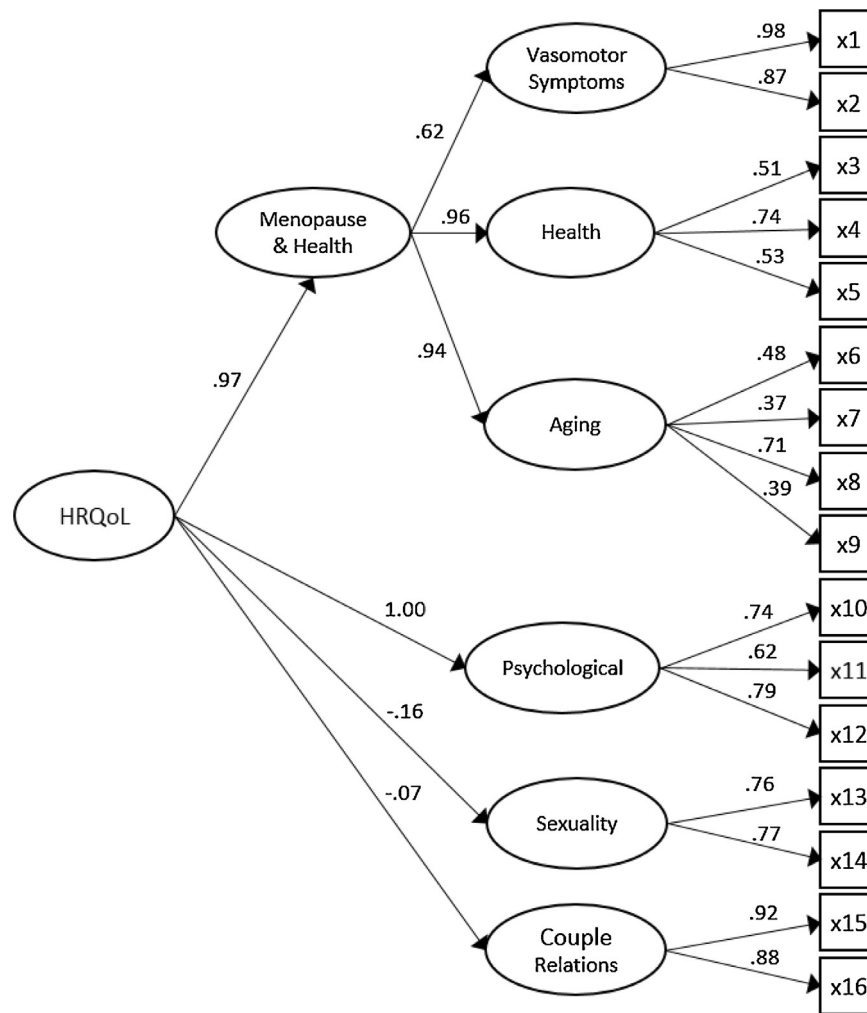
Relevant variables were collected in a single visit after informed consent was obtained. In order to avoid trailing effects when completing the two versions of the Cervantes scale, a random half of the sample completed the original 31-item Cervantes scale first, followed by the EQ-5D, a quality of life visual analogue scale, work productivity and activity impairment questionnaire (see references below), and the 16-item Cervantes scale, while the other half completed the Cervantes short-form first and the original version last.

This study followed the ethical principles of the Declaration of Helsinki for studies involving human subjects and was approved by the Clinical Research Ethical Committee of the Hospital Clínico San Carlos (Madrid).

### 2.2. Instruments

In addition to the 31-item and the 16-item versions of the Cervantes scale, three other scales were used. The instructions to correct the 16-item short form have been published elsewhere [10]. The EuroQoL 5D-3L (EQ-5D) [14] is a generic HRQoL instrument composed of five dimensions (Mobility, Personal Care, Daily Activities, Pain, Anxiety/Depression) assessed by three levels of impact each. It allows the computation of utility values ranging between 0 and 1 associated with each health profile, where 1 implies full health and 0 implies death (in theory it would be possible to obtain negative values associated with health conditions which are considered worse than death). The EQ-5D is accompanied by a visual analogue scale (EQ-VAS) consisting of a thermometer measuring the patient self-assessed current health status, anchored at 0 (worse imaginable health) and 100 (best imaginable health). The work productivity and activity impairment questionnaire: General Health V2.0 (WPAI: GH) [15] is a 6-item well validated instrument measuring impairments in work and activities as a consequence of current health condition. It measures absenteeism and presenteeism, as well as impairments in unpaid activity due to a health problem during the past seven days.

A symptom severity scale was used in measuring the intensity (1 = low intensity, 5 = high intensity) and the daily frequency of six menopause-related symptoms: suffocation, sweating crisis, tremors, and decrease in sexual drive, vaginal dryness, and dyspareunia. This scale was extracted from the 2010 US National Health and Wellness Survey [16]. The intensity of the first three symptoms was used to compute a vasomotor severity index and the intensity of the last two was used to compute a genital symptom severity index, based on the higher score obtained in any of the composing items (absent = 0, mild  $\leq 1$ , moderate  $\leq 2$ , severe  $\geq 3$ ), as previously performed by other researchers [16].



**Fig. 1.** Hierarchical confirmatory factor analysis. Standardised estimates. HRQoL = health-related quality of life. X1–X16: individual scale items. Error terms not represented. Values above arrows represent regression scores (similar to factor loadings).

### 2.3. Statistical analysis

Demographics and descriptive statistics were carried out to describe the sample used to test the measurement attributes of the Cervantes-SF questionnaire [13]. In the first step, item analysis was carried out under classic test theory item analysis assumptions [12,17]. Blank responses, unimodal response distribution, ceiling and floor effect, and skewness were studied. For the Cervantes 16-item short form, the conceptual and measurement model of the shortened version was assumed to be the same as that for the largest version except for describing the target population and its variability (including floor and ceiling rates, dispersion and standard error of measurement) and confirming its structural validity (original dimensionality) that was approached through confirmatory factor analysis (CFA), using listwise treatment of missing values. The original structure proposed by the scale developers was assumed: four dimensions with one of the dimensions divided into three subdimensions (see Fig. 1). Given that goodness-of-fit statistics are very sensitive to large sample size (over 200–400 cases), and trying to prevent over-fitting a single sample, a multigroup strategy was adopted. A subgroup analysis was also carried out separating sexually-active and non-active women, assuming a simpler structure where all dimensions loaded onto a single second-order dimension and using a WLSMV algorithm specifically designed for ordinal variables. The following goodness-of-fit (GOF) criteria were used [12,13]: chi square

( $\chi^2$ )/df < 4; GOF Index (GFI) > 0.90; Adjusted GFI (AGFI) > 0.90; comparative fit index (CFI) > 0.90; Tucker–Lewis index (TLI) > 0.90; and root mean square error of approximation (RMSEA) < 0.05. Because of the sample size, it was expected that  $\chi^2$  would be significant in all cases, while the significance of the difference between nested models,  $\Delta\chi^2$ , was considered informative.

Reliability was assessed by means of internal consistency, which was studied for each dimension and subdimension. Cronbach’s alpha (along with intra-class correlation coefficient with 95% confidence interval) and item–total and adjusted item–total correlations (deleting each particular item from the total score) were computed. An  $\alpha$  value above 0.95 was considered excellent, above 0.90 very good, above 0.80 good, above 0.70 acceptable, and below 0.70 poor [11]. Item–total statistics were computed considering the corresponding dimension total score. Reproducibility was also measured by test–retest correlation coefficient, globally and by dimension, between two administrations of the Cervantes-SF 24 to 48 h apart in a subsample of enrolled patients, without changes in the health status of the patients.

Validity attributes were assessed through various strategies. Criterion validity was estimated by computing the Pearson correlation between the proposed short version and the already-validated extended version, expecting high correlations between matching dimensions for the overall score between versions [13]. Construct concurrent validity included studying the relationship with sim-

ilar scales and/or related outcomes; thus, correlations were also computed with EQ-5D quality of life scores, health state VAS, number of visits due to menopause problems and WPAI:GH indexes. It was hypothesised a priori that Cervantes impairment scores would correlate negatively with HRQoL scores obtained using the EQ-5D and the health state VAS. Conversely, it was also hypothesised that a positive correlation would be found between Cervantes impairment scores and both the vasomotor and genital severity indexes, and also with the number of visits to the doctor due to menopause symptoms. With respect to work productivity impairment due to menopause, it was expected that a small correlation could be observed with the percentage of time lost, since menopause is not a severely disabling condition, but that a higher correlation would be observed with the percentage of work disability and with the percentage of disability in daily living activities. Known group (discriminant) validity was assessed as part of the construct validity by the vasomotor severity index, genital symptom severity index, and number of concomitant comorbidities [13]. Content validity was not tested as it was assumed to remain in the shortened version whether CFA could reproduce the original dimensionality of the scale. Feasibility was tested by determining burden of administration in terms of time devoted to fill in the abridged version of the Cervantes scale, percentage of missing questionnaires, and possible problems with item wording.

All analyses were carried out using IBM SPSS 20.0 software. Confirmatory factor analysis was carried out using IBM AMOS 20.0 module and MPLus 7.1.

### 3. Results

A total sample of 317 women was recruited, while the effective sample size was composed of 308 women (97.2% of recruited women). Mean ( $\pm$ standard deviation) age of the sample was  $55.7 \pm 5.3$  years, with a minimum age of 45 and a maximum of 71 years. Table 1 shows the main demographic and clinical descriptors. Most women (75.4%) were at the post-menopausal stage, 22.3% were at the peri-menopausal stage, 78.7% were sexually active, and the average BMI was  $25.2 \pm 4.2$  kg/m<sup>2</sup>. Over 69% of women were attending menopausal consultation. A total of 25.6% of women were current smokers, with an average consumption of  $10.3 \pm 7.5$  cigarettes/day. On the other hand, 25.2% of women were alcohol consumers with an average of  $1.2 \pm 0.4$  units/day. No other toxic habits were mentioned. The most frequent climacteric symptoms were the following: suffocation (73.8% of women), decrease in sexual drive (71.8%), vaginal dryness (70.2%), and sweating crisis (65.4%). Only 9% of women did not report any symptoms. Climacteric symptom mean intensity was close to the midpoint of the scale (0 = low intensity, 5 = high intensity): vaginal dryness ( $2.8 \pm 1.4$ ), decrease in sexual drive ( $2.8 \pm 1.4$ ), suffocation ( $2.5 \pm 1.4$ ), dyspareunia ( $2.5 \pm 1.4$ ), sweating crisis ( $2.4 \pm 1.4$ ), and tremors ( $1.3 \pm 0.7$ ). Attending to the severity indexes derived from symptoms, 25.6% of women did not present vasomotor symptoms, 24.0% presented mild, 30.8% moderate, and 19.5% severe symptoms. Additionally, 27.3% did not present genital symptoms, 16.6% presented mild, 32.1% moderate, and 24% severe genital symptoms.

EQ-5D utility scores ranged from 0.13 to 1, with a mean value of 0.796 (0.177). The distribution of scores was negatively skewed, with 45.9% of scores equal to or greater than 0.8. According to the WPAI:GH scores, although 71.1% of women in the sample reported to be professionally active, 96.8% reported not having lost any work time due to climacteric symptoms, 30.7% reported having experienced some level of work problems due to this health condition, and 42.8% reported some level of interference in daily activities. The WPAI:GH derived indexes gauged a mean work-time loss (absenteeism) of  $1.3 \pm 8.8\%$ , a mean percentage of work disability due

**Table 1**  
Demographic and clinical data of the analysed women ( $n = 308$ ).

Age (years): mean (SD)	55.7 (5.3)
BMI (kg/m <sup>2</sup> ): mean (SD)	25.2 (4.2)
Menopausal stage (%)	
Post-menopausal	232 (75.4)
Peri-menopausal	69 (22.3)
Sexually active: (%)	235 (78.7)
Attending menopause consultation: (%)	213 (69.0)
Climacteric symptoms: (%)	
Suffocation	228 (73.8)
Decrease in sexual drive	222 (71.8)
Vaginal dryness	217 (70.2)
Sweating crisis	202 (65.4)
Dyspareunia	150 (48.5)
Tremors	92 (29.8)
Irritability	12 (3.9)
Dysthymia	8 (2.6)
Insomnia	7 (2.3)
Asthenia	6 (1.9)
Changes in mood	6 (1.9)
Joint pain	6 (1.9)
Nervousness	5 (1.6)
None	27 (9)
Educational Level (%)	
None	0 (0)
Primary	56 (18)
Secondary	108 (35)
Professional degree	1 (0.3)
Higher	142 (17)
Working status (%)	
Active	216 (70)
Disabled	4 (1)
Unemployed	16 (5)
Retired	11 (4)
Homemaker	57 (19)
Civil status (%)	
Single	25 (8)
Married	221 (72)
Divorced	42 (14)
Widowed	18 (6)
Other	2 (1)
Smoker (%)	78 (25.6)
Cigarettes (units/day): mean (SD)	10.3 (5.50)
Alcohol consumption (%)	77 (25.2)
Units/day: mean (SD)	1.15 (0.49)

SD: standard deviation; BMI: body mass index. Some patients did not response to all questions in the case record form. Climacteric symptoms may sum more than 100%.

to menopause problems (presenteeism) of  $7.0 \pm 13.9\%$ , an overall work disability of  $8.2 \pm 16.3\%$ , and a mean overall disability in daily activities due to menopause symptoms of  $13.0 \pm 20.7\%$ .

#### 3.1. Reliability and item analyses

None of the items in the 16-item scale had more than 10% missing responses. Items with higher non-response rates were 13 (5.5%), 14 (4.2%), 15 (8.5%), and 16 (7.2%), the first two of which correspond to the sexuality dimension and the last two to the couple relationship dimension. After filtering out the women who reported not having a current partner (14%), the non-response rate was reduced to 0.4%, and after filtering out the women who reported not being sexually active, the non-response rate dropped to 1.7%. Most items exhibited a negatively skewed distribution of responses, except the last four items which were worded in inverted scale. Six items accumulated more than 50% of the responses in the lower category, with a maximum of 65% for item 8, suggesting the presence of floor effect.

Response time for the 16-item scale was  $2.6 \pm 2.3$  min, while it rose to  $5.0 \pm 2.1$  min for the 31-item scale. This time gain difference was statistically significant ( $t_{274} = 19.9$ ,  $p < 0.001$ ).

**Table 2**

Cervantes short-form scale reliability estimates: internal consistency (Cronbach's alpha) and test–retest correlation, % missing values, % floor effect, and % ceiling effect.

Dimension	# Items	Cronbach $\alpha$	95% CI	Test-retest ICC	SEM	Missing (%)	Floor (%)	Ceiling (%)
Menopause & health	9	.775	.734–.812	.844	8.82	4.9	1.6	0
Vasomotor symptoms	–2	.918	.897–.935	.798	8.95	0.6	26.4	4.9
Health	–3	.596	.510–.669	.638	12.89	2.6	15.3	0
Aging	–4	.51	.413–.595	.853	12.56	2.6	10.1	0
Psychological	3	.732	.675–.780	.716	11.93	1.6	22.1	1
Sexuality	2	.738	.670–.793	.925	14.69	7.5	6	9.4
Couple relations	2	.897	.870–.919	.888	9.6	8.8	34	8.4
Total	16	.82	.786–.850	.868	6.7	14	0	0

ICC: intra-class correlation coefficient, CI: confidence interval. SEM: standard error of measurement. Floor and ceiling effect means % of items with, respectively, lower or higher response in the items of the dimension.

Internal consistency reliability for the overall score was  $\alpha = 0.820$  (95% confidence interval: 0.786–0.850), and test-retest stability was  $r = 0.868$  (Table 2). Internal consistency varied in the first-order dimensions between 0.897 (couple relations) and 0.732 (psychological), and stability between 0.925 (sexuality) and 0.716 (Psychological).

### 3.2. Conceptual and measurement model

Confirmatory factor analysis solution for the hierarchical structure proposed originally obtained GOF statistics of  $\chi^2 = 360.45$  ( $p < 0.001$ ),  $\chi^2/df = 3.6$ , GFI = 0.869, AGFI = 0.824, CFI = 0.848, TLI = 0.850, and RMSEA = 0.099 (Fig. 1). Nevertheless, some of the parameter estimates were low, particularly the factor loadings for the sexuality ( $\lambda = -0.16$ ) and couple relations ( $\lambda = -0.07$ ) dimensions (Fig. 1). The solution reached would suggest that these two dimensions lack the power to measure HRQoL in our sample, which was not consistent with other concurrent results in our sample.

Given previous evidence, a separate estimation was carried out with sexually-active and non-active women, this time assuming a simpler structure where all dimensions loaded onto a second-order HRQoL dimension and with the WLSMV estimation method, which is specific for categorical (ordinal) variables. Results for the sexually-active group gave GOF indexes of CFI = 0.913 and TLI = 0.893. All estimated parameters were statistically significant ( $p < 0.001$ ) and were in the direction expected by theory (Fig. 2a). The Psychological dimension exhibited the highest second-order factor loading ( $\gamma_{41} = 0.997$ ), followed by the Health dimension ( $\gamma_{21} = 0.966$ ). As expected by theory, the sexual ( $\gamma_{51} = -0.437$ ) and couple relations ( $\gamma_{61} = -0.419$ ) dimensions obtained negative loadings, since they are worded in inverted scale. Fig. 2b shows the confirmatory structure of the scale in non-sexually-active women, which corresponded to the one expected *a priori*, with moderate to higher standardised estimates and all of them being statistically significant. The non-sexually-active group had a relatively small sample size ( $n = 64$ ) and also obtained a good fit (CFI = 0.917, TLI = 0.900). The dimension with the highest second-order factor loading was aging ( $\gamma_{31} = 0.994$ ), followed by health ( $\gamma_{21} = 0.992$ ). In this group, the second-order factor loadings for the Sexual ( $\gamma_{51} = 0.395$ ) and couple relations ( $\gamma_{61} = 0.630$ ) dimensions were positive.

### 3.3. Criterion and construct validity

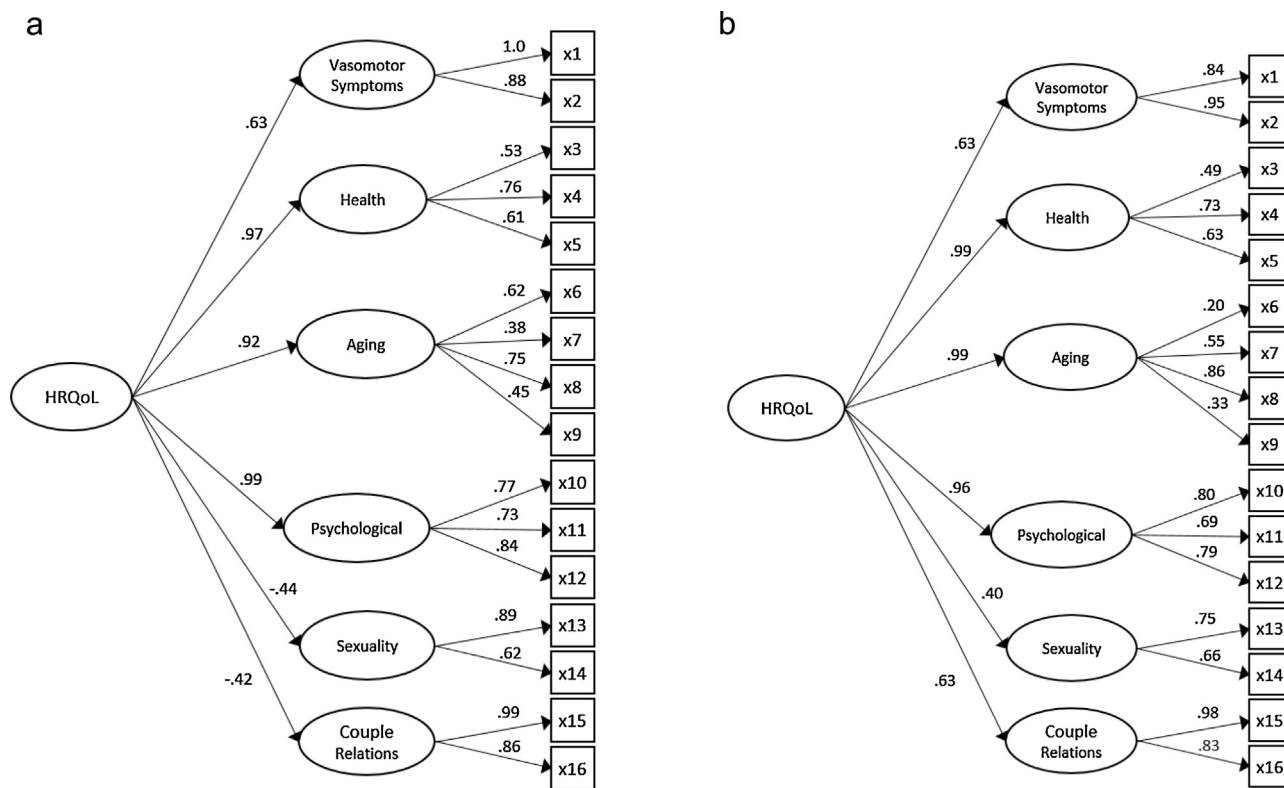
Correlations between matching dimensions from the 16-item short version and the 31-item original version were all close to 0.8 or higher and significant ( $p < 0.001$ ). The correlation between overall scores for both versions was high ( $r = 0.890$ ). The individual dimension with the highest correlation was Couple Relations ( $r = 0.864$ ), while the one with the lowest correlation was Health ( $r = 0.791$ , see Table 3).

The 16-item overall score negatively correlated with the EQ-5D utility score ( $r = -0.515$ ), suggesting that lower quality of life health states assessed by the Cervantes scale are related to lower utility values, and vice versa (Table 3). By dimension, the one exhibiting the highest correlation with the utility score was the Psychological dimension ( $r = 0.568$ ). The sexuality dimension obtained a low but significant correlation with utility scores ( $r = -0.173$ ,  $p = 0.003$ ) and the couple relations dimension attained a correlation close to null ( $r = -0.068$ ,  $p = 0.256$ ). The 16-item overall score also negatively correlated with the health state VAS ( $r = -0.453$ ), and the individual dimensions exhibited a similar pattern of correlations as mentioned above. Once more the couple relations dimension attained a correlation close to null ( $r = -0.083$ ,  $p = 0.170$ ).

Indexes computed from the symptom severity scales also correlated with the Cervantes individual dimensions. The vasomotor severity index correlated with the vasomotor dimension ( $r = 0.592$ ) and the genital symptom severity index correlated with the aging dimension ( $r = 0.360$ ,  $p < 0.001$ ) but not with the sexuality ( $r = 0.096$ ) or couple relations ( $r = 0.061$ ) dimensions. Nevertheless, when the sexually-active subsample was studied, the genital symptom severity index did correlate significantly with the sexuality dimension ( $r = 0.238$ ,  $p < 0.001$ ). The number of visits to the doctor due to menopause symptoms attained the highest correlation with the dimensions of menopause and health ( $r = 0.265$ ,  $p < 0.001$ ), followed by the aging dimension ( $r = 0.244$ ,  $p < 0.001$ ). Medical visits were not significantly correlated with either the sexuality or couple relations dimensions of the Cervantes-SF.

Attending to indexes derived from the WPAI:GH scale (Table 3), in the subsample of labour-active women, there was no significant correlation between the Cervantes-SF overall score and the percentage of work lost due to menopause problems (absenteeism,  $r = 0.027$ ,  $p = 0.692$ ), or with any of the Cervantes dimensions. A significant correlation was observed between the percentage of work impact (presenteeism) due to menopause problems and overall Cervantes scores ( $r = 0.317$ ,  $p < 0.001$ ). Furthermore, a significant correlation was found between the percentage of overall work load disability and the overall score ( $r = 0.278$ ,  $p < 0.001$ ). Taking into account all women in the sample (labour-active or not), a significant correlation was observed between the percentage of impairment in daily living activities and the Cervantes score ( $r = 0.454$ ,  $p < 0.001$ ).

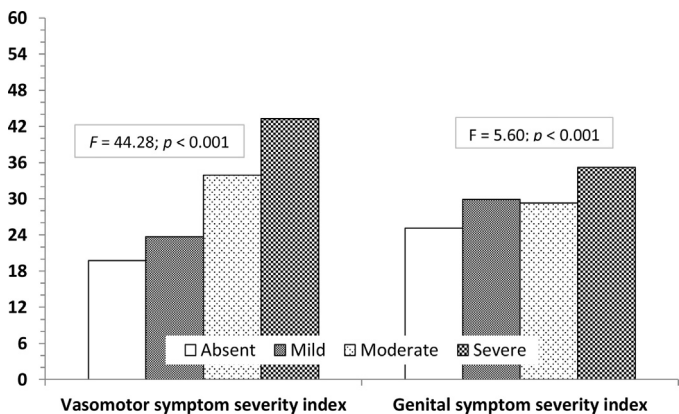
Comparing the mean overall Cervantes scores between the groups defined by vasomotor symptom severity (absent, mild, moderate, severe), statistically significant differences were found among groups ( $F_{3,303} = 44.3$ ,  $p < 0.001$ , Fig. 3). The absent- and mild-symptom groups were not significantly different ( $p = 0.259$ ), while both attained significantly lower scores than the other severity groups ( $p < 0.001$ ), which differed from each other ( $p < 0.001$ ) (Fig. 3). Mean overall score differences were also found among the genital-symptom severity groups ( $F_{3,303} = 5.6$ ,  $p = 0.001$ ). The absent-symptom group differed from the severe-symptom group ( $p < 0.001$ ), and the difference between the mild- and severe-severity groups was close to significance ( $p = 0.063$ ). For



**Fig. 2.** (a) Second-order confirmatory factor analysis: sexually-active women. Standardised estimates. HRQoL = health-related quality of life. X1–X16: individual scale items. Error terms not represented. Values above arrows represent regression scores (similar to factor loadings). (b) Second-order confirmatory factor analysis: non-sexually-active women. Standardised estimates. HRQoL = health-related quality of life. X1–X16: individual scale items. Error terms not represented. Values above arrows represent regression scores (similar to factor loadings).

**Table 3**  
Criterion and construct validity of the 16-item Cervantes scale with the original 31-item version and other concurrent measures.

Dimension	Cervantes 31-item	EQ-5D	Health state VAS	Vasomotor severity	Genital symptoms	Medical visits	Absenteeism <sup>a</sup>	Presenteeism <sup>a</sup>	Total work disability <sup>a</sup>	Total daily activities disability
Health & menopause	.864**	-.507**	-.434**	.533**	.237**	.265**	-.016	.409**	.330**	.550**
Vasomotor	.881**	-.338**	-.319**	.592**	.115**	.210**	.025	.361**	.309**	.377**
Health	.791**	-.438**	-.380**	.308**	.158**	.190**	-.064	.279**	.201**	.380**
Aging	.822**	-.498**	-.372**	.269**	.360**	.244**	-.023	.327**	.257**	.487**
Psychological	.842**	-.568**	-.426**	.281**	.245**	.198**	.045	.402**	.360**	.476**
Sexuality	.807**	-.173**	-.200**	.061	.096	.058	.111	-.073	-.027	.117 <sup>†</sup>
Couple relations	.864**	-.068	-.083	.050	-.061	-.048	.052	-.024	-.003	-.013
Overall	.890**	-.515**	-.453**	.426**	.213**	.215**	.027	.317**	.278**	.454**



**Fig. 3.** Mean overall Cervantes-SF score by vasomotor and genital symptoms severity levels (validity of known groups). Cervantes short-form score (Y-axis) ranging from 0 (no symptom impact on quality of life) to 100 (highest impact).

genital-symptom severity (absent, mild, moderate, severe), statistically significant differences were found among groups ( $F_{3,303} = 5.6$ ,  $p < 0.001$ , Fig. 3), particularly due to the high score of 35.2 in the severe group (Fig. 3).

Mean overall Cervantes score differed depending on the number of comorbidities suffered by women ( $F_{4,301} = 4.2$ ,  $p = 0.003$ ), with an overall pattern of increasing deterioration in quality of life, although only the group with no comorbidities differed statistically from the group with three comorbidities present ( $p = 0.009$ ). The scores were 26.2, 31.5, 30.6, 38.4, and 36.0, respectively for 0, 1, 2, 3, and 4 comorbidities.

**4. Discussion**

This study assessed the psychometric attributes of the conceptual and measurement model, as well as the reliability and validity of the 16-item Cervantes-SF quality of life scale in peri- and post-menopausal women. The sample obtained was gathered from a wide range of clinical settings and by a relatively high number

of clinicians, and could be considered representative of peri- and post-menopausal women in Spain. The prevalence of menopause-related symptoms was high, especially for those symptoms listed in the clinical interviews [16], but their reported intensity was always close to the scale mean value. One in five women presented severe vasomotor symptoms and one in four presented severe genital symptoms. Presenteeism reached 7% of time worked, overall work load disability surpassed 8%, and impact on daily activities was close to 13%.

The Cervantes-SF demonstrated good feasibility with little burden of administration: very few women left responses blank and completion time was just over two and a half minutes, which is significantly less time-consuming than the original 31-item version. The two dimensions relating to sexuality and couple relations could be less pertinent to women who are not sexually active, as might be expected. Responses to items tended to accumulate in the lower impairment values, suggesting that the level of impairment was rather low for this sample. In fact, the mean value for the overall score was 30 on a 0–100 scale. Scale reliability statistics were considered good and not very different from the value of the original version. Scores demonstrated to be stable between presentations. All of these suggest that we were able to measure with precision the level of impairment due to climacteric symptoms in our patients. Reliability for the isolated dimensions was not as good, particularly for the health and aging subdimensions, which only reached mild reliability. This last result was expected, since the short-form version was designed as a global and faster alternative to the 31-item version, which should be preferred by researchers and clinicians when an accurate individual dimension diagnosis is pursued. It should be noted that score correlations between the short form and the extended 31-item form were very high. Thus, criterion validity was highly sustained [11,13].

The structure of responses to the questionnaire fitted the configuration proposed by theory, but it was also found that the dimensions of sexuality and couple relations behaved in a different manner depending on whether the women were sexually active or not. While both dimensions behaved as expected for sexually-active women, in the non-active group the dimensions loaded inversely to the expected sign. This result could be explained by several facts: non-sexually-active women presented a very high level of impairment in these two dimensions (higher than those for the sexually-active women and higher than the impairment in other dimensions); being sexually active was statistically related to having a stable partner; and sexually-active women were, on average, three and a half years younger. Whatever the explanation might be, we should recommend interpreting with care scores in these two dimensions when a patient is not sexually active [17]. Group divergent results for the two dimensions related to sexual activity explain why estimates for the overall sample are so small in value.

On the other hand, scores obtained with the Cervantes-SF correlated significantly with other quality of life or severity scores, which *a priori* were hypothesised to be related to quality of life in peri- or post-menopausal women [18]. The correlation with health-related utility scores was meaningful, as was that with the health state scores. Vasomotor symptom severity correlated with the corresponding dimensions and, to a lesser extent, genital symptom severity also correlated with the matching dimensions when sexually-active women were studied. The overall score also correlated with the number of visits to the clinician due to climacteric problems, as was hypothesised. As expected, Cervantes-SF scores did not correlate with absenteeism due to menopause symptoms, while they did correlate with presenteeism estimates and impact on daily activities, suggesting that the impact of menopause in daily living could be worryingly disabling. Age did not correlate with

the overall score but it did correlate with aging, sexuality and couple relations dimension scores, this result being in line with the above-mentioned findings for non-sexually-active women. Construct validity of the Cervantes-SF was also confirmed by means of validity of known groups. The scale was able to significantly differentiate the severity of vasomotor symptoms as well as the genital symptoms, but in particular the observed linear relationship between severity of vasomotor symptoms and quality of life was highly significant: the higher the severity of vasomotor symptoms, the higher the impact on the patient's quality of life. In addition, a similar pattern was seen with the number of concomitant comorbidities, with the impact on quality of life being greater in subjects with more comorbidity. This ability of the Cervantes-SF is important as previous published studies have reported the association of poor quality of life in menopausal women with several comorbidities [19–23].

It is important to bear in mind that the original Cervantes scale was developed to also include the assessment of the impact of menopause on couple relations [8], based on previous findings derived from the development of the MENCAV scale [24], where the relevance of couple relations for the Spanish society was identified. At the time of development, inclusion of the sexuality and couple relations dimensions was deemed to be a central aspect differentiating the Cervantes scale from other existing instruments [25–28]. We still contend that these two dimensions are very important for women in this particular age group [17], and rightly so since other disease-specific instruments measuring the impact of induced menopause on quality of life also include sexuality as a relevant dimension, for example those in oncology (see EORTC QLQ BR-23 [29] and FACT-B [30,31]).

The initial purpose of abridging the original Cervantes scale was to reduce the instrument length as far as possible without losing any of the original concepts/dimensions contained in the original extended version and hence preserving its psychometric attributes of measurement. The 16-item version has shown to keep all the original properties of the extended version, and it is proving to be more feasible in a clinical setting. However, our study is not free of possible limitations. First of all, we assumed that our sample was representative of the Spanish peri- and post-menopausal women. However, the rate of unemployed women found in our study was lower than the current rate in Spain at the time of conducting the study. Nevertheless, we believe that this should not have affected the psychometric attributes of the measurement scale. Another limitation is that the instrument has shown a somewhat different behaviour when assessing non-sexually-active women, as could be expected. In this group, the sexuality and couple relations dimension scores might have reflected acceptance of the sexual role instead of interference in HRQoL.

Despite possible limitations, we may conclude that the 16-item Cervantes short-form scale is well understood by women, and it is a pertinent instrument to be used for the assessment of the degree of impact of climacteric symptoms on health-related quality of life. The scale has been shown to be reliable and to reproduce the structure of the theoretical model proposed for the extended version. It has also been shown to correlate as expected with other convergent measures, validating its meaning and purpose. As a whole, the good psychometric properties obtained ensure its use in clinical settings, enabling a fast measurement of the impact of climacteric symptoms on the patient's daily living.

## Funding

This work was promoted and sponsored by the Spanish Menopause Society (SMS). Statistical analysis was carried out by Miguel Ruiz from the Universidad Autónoma de Madrid, and was

funded by a restricted grant from Pfizer, S.L.U. No editorial support was provided except for review of the English manuscript by NOVA Traductors i Intèrprets, S.L., which was funded by Pfizer, S.L.U.

### Authors's contributions

All authors had complete access to the data, participated in the analysis and/or interpretation of results, and drafted the manuscript. PJC, RSB, SP, and JR were responsible for the design of the study. Data analysis was performed by MAR.

### Conflict of interest

The AEEM collaborative group for the study of psychometric validation of the Cervantes Short-Form was constituted by P. Coronado, R. Sánchez-Borrego, A. Duque, M. Fernández-Abellán, S. Sánchez, E. Iglesias, L. Baquedano, S. González, B. Otero, J. Presa, F. Ruiz, J. Calleja, and C. Argudo. Javier Rejas is an employee of Pfizer, S.L.U. Miguel A. Ruiz is a professor of statistics and methodology at the Universidad Autónoma de Madrid and received a grant from Pfizer, S.L.U. for the statistical analysis. All other authors declare that they have no conflicts of interests.

### Ethical approval

This study followed the ethical principles of the Declaration of Helsinki for studies involving human subjects and was approved by the Clinical Research Ethical Committee of the Hospital Clínico San Carlos (Madrid).

### Acknowledgments

The authors wish to thank the Spanish Menopause Society (SMS) for its support and guidance.

### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.maturitas.2015.10.013>.

### References

- [1] WHO Scientific Group, Research on the menopause in the 1990 WHO Technical Report Series, vol. 866, World Health Organization, Geneva, 1999, pp. 1–3.
- [2] A. Pines, Vasomotor symptoms and cardiovascular disease risk, *Climacteric* 14 (2011) 535–536.
- [3] H. Busch, A.S. Barth-Olofsson, S. Rosenhagen, A. Collins, Menopausal transition and psychological development, *Menopause* 10 (2003) 179–187.
- [4] N.E. Avis, R. Stellato, S. Crawford, J. Bromberger, P. Ganz, V. Cain, et al., Is there a menopausal syndrome? Menopausal status and symptoms across racial/ethnic groups, *Soc. Sci. Med.* 52 (2001) 345–356.
- [5] C. Castelo-Branco, S. Palacios, J. Ferrer-Barriendos, et al., Understanding how personality factors may influence quality of life: development and validation of the Cervantes Personality Scale, Cervantes Study Group, *Menopause* 15 (2008) 914–918.
- [6] J.E. Blümel, C. Castelo-Branco, N. Kerrigan, M.J. Cancelo, B. Blümel, J. Haya, et al., Influences of hormone replacement therapy on postmenopausal women's health perceptions, *Menopause* 10 (2003) 235–240.
- [7] I.B. Meeuwssen, M.M. Samson, S.A. Duursma, H.J. Verhaar, The influence of tibolone on quality of life in postmenopausal women, *Maturitas* 41 (2002) 35–43.
- [8] S. Palacios, J. Ferrer-Barriendos, J.J. Parrilla, C. Castelo-Branco, M. Manubens, X. Alberich, et al., Calidad de vida relacionada con la salud en la mujer española durante la perimenopausia y posmenopausia. Desarrollo y validación de la Escala Cervantes, *Med. Clin. (Barc.)* 122 (2004) 205–211.
- [9] P. Coronado, R. Sánchez Borrego, S. Palacios, M.A. Ruiz, J. Rejas, Structural validity of a 16-item abridged version of the Cervantes health-related quality of life scale for menopause: the Cervantes Short-Form scale, *Menopause* 22 (2015) 325–336, <http://dx.doi.org/10.1097/gme.0000000000000321>.
- [10] O. Chassany, P. Sagnier, P. Marquis, S. Fullerton, N. Aaronson, for the European Regulatory Issues on Quality of Life Assessment Group: patient-reported outcomes: the example of health-related quality of life—a European guidance document for the improved integration of health-related quality of life assessment in the drug regulatory process, *Drug Inf. J.* 36 (2002) 209–238.
- [11] J.C. Nunnally, I.H. Bernstein, *Psychometric Theory*, 3rd ed., McGraw-Hill, New York, 1994.
- [12] N. Aaronson, J. Alonso, A. Burnam, Assessing health status and quality-of-life instruments: attributes and review criteria, *Qual. Life Res.* 11 (2002) 193–205.
- [13] P.M. Fayers, D. Machin, *Quality of life: The assessment, analysis and interpretation of patient-reported outcomes*, second ed., John Wiley and Sons, Ltd., Chichester, UK, 2007:1–10.
- [14] The EuroQol Group, EuroQol—a new facility for the measurement of health-related quality of life, *Health Policy* 16 (1990) 199–208.
- [15] A.S. Zbrozek, E.M. Dukes, The validity and reproducibility of a work productivity and activity impairment instrument, *Pharmacoeconomics* 4 (1993) 353–365.
- [16] J. Whiteley, J.S. Wagner, A. Bushmakim, L. Kopenhafer, M. DiBonaventura, J. Racketta, Impact of the severity of vasomotor symptoms on health status, resource use and productivity, *Menopause* 20 (2013) 518–524, <http://dx.doi.org/10.1097/GME.0b013e31827d38a5>.
- [17] M.K. Connor, N.N. Maserejian, L. De Rogatis, C.M. Meston, E.P. Gerstenberger, R.C. Rosen, Sexual desire, distress, and associated factors in premenopausal women: preliminary findings from the hypoactive sexual desire disorder registry for women, *J. Sex Marital Ther.* 37 (2011) 176–189, <http://dx.doi.org/10.1080/0092623X.2011.560533>.
- [18] S. Palacios, R.S. Borrego, A. Forteza, The importance of preventive health care in post-menopausal women, *Maturitas* 52 (Suppl. 1) (2005) S53–S60.
- [19] C. Cooper, F. Jakob, C. Chinn, et al., Fracture incidence and changes in quality of life in women with an inadequate clinical outcome from osteoporosis therapy: the Observational Study of Severe Osteoporosis (OSSO), *Osteoporos. Int.* 19 (2008) 493–501.
- [20] S. Palacios, J.L. Neyro, S. Fernández de Cabo, J. Chaves, J. Rejas, Impact of osteoporosis and bone fracture on health-related quality of life in postmenopausal women, *Climacteric* 17 (2014) 60–70, <http://dx.doi.org/10.3109/13697137.2013.808182>.
- [21] J. Ferrer, J.L. Neyro, A. Estevez, Identification of risk factors for prevention and early diagnosis of a-symptomatic post-menopausal women, *Maturitas* 52 (Suppl. 1) (2005) S7–S22.
- [22] J. Whiteley, M. DiBonaventura, J.S. Wagner, J. Alvir, S. Shah, The impact of menopausal symptoms on quality of life, productivity, and economic outcomes, *J. Womens Health (Larchmt.)* 22 (2013) 983–990, <http://dx.doi.org/10.1089/jwh.2012.3719>.
- [23] B.J. Bankowski, L.M. Gallicchio, M.K. Whiteman, L.M. Lewis, H.A. Zacur, J.A. Flaws, The association between menopausal symptoms and quality of life in midlife women, *Fertil. Steril.* 86 (2006) 1006–1008.
- [24] J. Buendía Bermejo, R. Rodríguez Segarra, N. Yubero Bascuñana, V. Martínez Vizcaíno, Diseño y validación de un cuestionario en castellano para medir la calidad de vida en mujeres posmenopáusicas: el cuestionario MENCAV, *Aten. Primaria* 27 (2001) 94–100.
- [25] J.R. Hilditch, J. Lewis, A. Peter, B. Van Maris, A. Ross, E. Franssen, et al., A menopause-specific quality of life questionnaire: development and psychometric properties, *Maturitas* 24 (1996) 161–175.
- [26] H.P.G. Schneider, L.A.J. Heinemann, H.P. Rosemeier, P. Potthoff, H.M. Behre, The Menopause Rating Scale (MRS): comparison with the Kupperman index and quality-of-life scale SF-36, *Climacteric* 3 (2000) 50–58.
- [27] J.G. Greene, *Guide to the Greene Climacteric Scale*, University of Glasgow, Glasgow, 1991.
- [28] W.H. Utian, J.W. Janata, S.A. Kingsberg, M. Schluchter, J.C. Hamilton, The Utian Quality of Life (UQOL) Scale: development and validation of an instrument to quantify quality of life through and beyond menopause, *Menopause* 9 (2002) 402–410.
- [29] N.K. Aaronson, S. Ahmedzai, B. Bergman, et al., The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology, *J. Natl. Cancer Inst.* 85 (1993) 365–376.
- [30] D.F. Cella, D.S. Tulskey, G. Gray, The functional assessment of cancer therapy (FACT) scale: development and validation of the general measure, *J. Clin. Oncol.* 11 (1993) 570–579.
- [31] M.J. Brady, D.F. Cella, F. Mo, Reliability and validity of the functional assessment of cancer therapy-breast (FACT-B) quality of life instrument, *J. Clin. Oncol.* 15 (1997) 974–986.