

Climacteric symptoms and quality of life: validity of women's health questionnaire

Sintomas climatéricos e qualidade de vida: validação do questionário da saúde da mulher

Carlos Rodrigues da Silva Filho^a, Edmundo Chad Baracat^b, Lucieni de Oliveira Conterno^a, Mauro Abi Haidar^b and Marcos Bosi Ferraz^c

^aDepartamento de Medicina. Faculdade de Medicina de Marília. Marília, SP, Brasil. ^bUniversidade Federal de São Paulo (Unifesp). São Paulo, SP, Brasil. ^cDepartamento de Medicina. Unifesp. São Paulo, SP, Brasil

Keywords

Climacteric. Quality of life. Women's health. Validity. Questionnaires.

Abstract

Objective

To evaluate the reliability and validity of the Portuguese version of the Women's Health Questionnaire.

Methods

In order to evaluate the Women's Health Questionnaire (WHQ), an analytical cross-sectional study was carried out at the women's menopause outpatient clinic of a university hospital in São Paulo, Brazil. There were studied 87 women in perimenopause or menopause, defined as experiencing at least one year's absence of menstrual flow. The following variables were collected: demographic data, clinical variables (Kupperman index and correlate numeric scale) and quality of life indexes (SF-36 and utility).

Results

The WHQ proved to be a questionnaire easily translated into Portuguese and well-adjusted to Brazilian women. The internal consistency of the overall WHQ was excellent (Cronbach alpha =0.83; 95% CI: 0.71-0.91). Test-retest reliability was also excellent (intraclass correlation coefficient [ICC]=0.92; 95% IC: 0.86-0.96) and had good absolute agreement (0.84; 95% CI: 0.71-0.92). A satisfactory clinical validity was observed. The construct validity was corroborated by clear associations with others scales. A good index of responsiveness after the intervention was reached.

Conclusions

The Portuguese version of the WHQ is of easy and fast administration and understanding. Its measuring properties were related, allowing its use in the evaluation of Brazilian climacteric women's quality of life for various purposes.

Descritores

Climatério. Qualidade de vida. Saúde da mulher. Validade. Questionários.

Resumo

Objetivo

Validar para o português o instrumento de avaliação de qualidade de vida no climatério, Women's Health Questionnaire.

Métodos

Para avaliação do Women's Health Questionnaire, foi realizado estudo transversal analítico no ambulatório de assistência ao climatério de um hospital universitário do Município de São Paulo. Foram estudadas 87 mulheres na peri-menopausa ou menopausa, definida como ao menos um ano sem apresentar fluxo menstrual, e analisadas as seguintes variáveis: demográficas, índices clínicos (índice

Correspondence to:

Carlos Rodrigues da Silva Filho
Faculdade Estadual de Medicina de Marília
Av. Monte Carmelo, 800 Bairro Fragata
17519-030 Marília, SP, Brasil
E-mail: silvacr@famema.br or
conterno.rodrigues@flash.tv.br

Financial support by INCLEN Trust (International Clinical Epidemiology Network - Grant n. 1307/1999) Philadelphia, PA, USA.

Received on 19/5/2003. Reviewed on 16/8/2004. Approved on 9/11/2004.

menopausal de Kuppermann, e escala numérica correlata) índices de qualidade de vida (SF-36, e utility).

Resultados

A consistência interna do WHQ traduzido foi muito boa (Coeficiente Alfa de Cronbach =0,83 IC 95%: 0,71-0,91), assim como a correlação intra-classe (teste-reteste =0,92; IC 95%: 0,86-0,96), e boa concordância absoluta (0,84; IC 95%: 0,71-0,92). Sua validade de construto foi corroborada pela boa associação com outras escalas. A validade clínica foi considerada satisfatória e um bom índice de sensibilidade após intervenção foi alcançado.

Conclusões

A versão para o português do Women's Health Questionnaire é de fácil e rápida aplicação e compreensão. Suas propriedades de medida foram avaliadas e provadas podendo ser utilizada para a avaliação da qualidade de vida das mulheres brasileiras no climatério, para vários objetivos.

INTRODUCTION

Due to declining fertility rates and increasing life expectancy – a global yet heterogeneous trend –, the number of women reaching menopause has been steadily on the rise.

This brings on the need to understand in what health these women reach this age, so that they can be provided, as a group or individually, with adequate care. Recently, there has been an increased interest in monitoring patient's self-perception and their response to therapy.⁸

Unlike biomedical testing, qualitative health aspects are better mirrored in quality of life assessment. These scales focus on subjective symptoms as perceived by patients and allow to understanding how they affect well-being and daily affairs, reaching beyond traditional clinical indexes that provide relevant additional information, routinely overlooked in habitual procedures but of equal or superior importance as compared to routinely applied clinical, biochemical or physiological indexes.¹

Additionally, for better understanding this issue, general quality of life aspects, or those aspects specific to some diseases or situations, can be developed by preferably leaving binary measurement scales aside and favoring broader methods. This is due to the subjectivity of a specific case, allowing for sufficient responsiveness in detecting symptomatic alterations following an intervention.

Their results comprise a set of symptoms and signs, which indicate the same condition or symptom complex, manifested differently in each individual.²

In this study, it was applied a previously developed questionnaire, which has been used in another

context under a structured methodology, which on first inspection seemed adequate due to its range of coverage, and the use of terms and expressions, which seemed easily understandable to the study subjects, considering the simplicity of use (face validity).

This scale, known as Women's Health Questionnaire (WHQ), was developed in English and is pertinent to women's health, since it seeks to assess not only menopause-related complaints, but also global transformations in women's lives, which may affect their quality of life.⁷

The WHQ scale is well accepted internationally. Many translations were done according to international methodological recommendations (French, Swedish, Afrikaans, Bulgarian, Danish, Dutch, Belgium Dutch, Australian and Canadian English, German, Italian, Spanish, and other language cultures).^{8,10,15}

In order for these instruments to be used in different cultures, it is necessary for them to possess well-documented psychometric characteristics, i.e., reliability, validity and responsiveness to alterations.

There are further elements involved in these cross-cultural translation problems, such as the use of language and potential communication problems.⁴ Translation equivalency problems are frequently observed. Different types of equivalency problems are involved: vocabulary, language, grammar and syntax, in addition to procedures.¹²

The purpose of this study was to provide psychometric documentation details of the translation into Portuguese of the Women's Health Questionnaire.

METHODS

The WHQ was developed in England in 1986, and is characterized specifically for avoiding an empha-

sis on clearly post-menopausal symptoms, permitting an overall assessment of other changes, which occur in women in this phase of their lives, which may affect its quality.

In addition the WHQ is the first quality of life measure to be included in the International Health-Related Quality of Life Outcomes Database (IQOD).⁸

It comprises 36 symptoms and signs, rated on a 4-point scale, including: somatic symptoms, depressed mood, cognitive difficulties, anxiety and fear, sexual functioning, vasomotor symptoms, sleep problems, menstrual problems, and self-perceived attraction.⁷ It provides individual dimensions and overall scores. The higher the score, the more pronounced the suffering and dysfunction.

The WHQ has been validated in Portuguese,* following internationally accepted methodology:⁵

- 1) Translation. Two translators, with good proficiency in both languages, did the first translation, aware of the concept but not of its objectives.
- 2) Back translation. Two different translators, unfamiliar with the original questionnaire, translated it back into English. The result was compared to the original questionnaire and the discrepancies were identified. At this point, a panel composed of two geriatricians, two gynecologists, two English professors, and three other physicians from different parts of the country met and discussed the discrepancies, until they came to a consensus, which was then put into writing.
- 3) The cultural equivalence evaluation. After adding an option, "does not apply," to the possible answers, nine women, were summoned to the climacteric outpatient clinic. They all came from different parts of the country, in order to, following completion of the questionnaire, detect any part of the translation which caused comprehension difficulties and identify experiences, which were not a part of their daily life. These items were then rewritten. The result of this process was adopted as the final WHQ.¹³

Assessment of measuring characteristics

- 1) Reliability assessment. The Portuguese version of the WHQ had both its internal consistency and intraclass reliability tested by means of three interviews. It was applied at the climacteric outpatient department of an university consecutively to a group of 45 patients with no history of hormonal replacement therapy (HRT) in the previous six months. Two interviewers conducted individual examinations on the first day (interviewer 1 and 2). Within a 10-day period

following the first interview, interviewer 1 applied the questionnaire again to the patients.

- 2) Validity assessment. Results were correlated to four other testing instruments, equally administered to 67 patients at the same outpatient clinic, namely: Kupperman menopausal index,⁹ a numeric scale comprising the four most meaningful items of the former index, the overall life quality assessment questionnaire SF-36,² and a utility, obtained from the construction of three different scenarios depicting different degrees of menopausal symptoms with higher grades indicating more moderate conditions, in which the patients sought to summarize their health.

The authors' hypothesis predicted that a higher total score on the dimensions of the WHQ would correspond to a higher score on Kupperman menopausal index and on results of the numeric scale (NS), and lower scores in the areas of the SF-36 and utility.

- 3) Responsiveness. To test this, score variations in the Portuguese version of the WHQ were correlated with variations in other clinical parameters in 20 menopausal patients, not included in the previous group, to whom hormonal replacement therapy had been prescribed by their physicians (estrogen or estrogen-progesterone) in an unsupervised and continuous manner, orally or transcutaneously, during an average 89.7 day period, with pre- and post-therapy assessments by the same interviewer.

In the cultural equivalence assessment phase, for convenience purposes, the questionnaire was applied to and discussed with a group of nine patients. In the assessment of the measuring properties, the questionnaire was applied to two other groups of patients. In the first, 67 women were tested consecutively to test for validity and reliability (45 patients; 22 did not return for the second interview in the study period). The second was a separate group of 20 patients, in which responsiveness was tested.

For validation purposes, the dimension of the sample was calculated through the Systematized Program for Sample Size Calculation in Research Design,¹¹ which uses a formula proposed by Walter et al.¹⁴ It was adopted as the interest parameter the interclass correlation coefficient, a 0.05 error type 1 and 0.20 error type 2.

Adopting 0.80 as an adequate intraclass correlation coefficient and the above described errors, in a ± 15 confidence interval amplitude, the necessary sample would be at least 40 patients.

Descriptive statistics were used to characterize

Table 1 - Socio-demographic features of studied women.

| | |
|--------------------------------------|-------------|
| Number of subjects | 87 (100%) |
| Age | |
| Mean (SD) | 54.8 (9.0) |
| Minimum - maximum | 31-73 |
| Skin color | |
| White | 52 (60%) |
| Black | 19 (22%) |
| Yellow | 1 (1%) |
| Mulatto | 15 (17%) |
| Marital status | |
| Single | 9 (10%) |
| Married | 46 (53%) |
| Widowed | 11 (13%) |
| Divorced | 13 (15%) |
| Others | 8 (9%) |
| Occupation | |
| Homemaker | 43 (49%) |
| Professional | 30 (35%) |
| Retired | 14 (16%) |
| Socio-economic status | |
| A | 0 (0%) |
| B | 14 (16%) |
| C | 32 (37%) |
| D | 31 (36%) |
| E | 10 (11%) |
| Number of gestations | |
| 0 | 11 (13%) |
| 1 | 7 (8%) |
| 2 | 15 (17%) |
| 3 or + | 54 (62%) |
| Length of time of amenorrhea (years) | |
| Mean (SD) | 8.2 (6.9) |
| Minimum - maximum | <1 up to 27 |

SD: Standard deviation

patient demographics, and interclass correlation coefficients for the evaluation of test-retest and absolute agreement reliability. Cronbach alpha coefficient was used to assess the questionnaire's internal consistency and Pearson's correlation coefficient was used for validity. For the responsiveness assessment, the following ratio was used: change magnitude measured by the score of each one of the WHQ dimensions/variability of those changes among different kinds of people (average change/standard error).

RESULTS

No question was considered non-applicable by more than one in the group of nine women who participated in the cultural equivalence assessment.

However, there was a doubt concerning a question about menstrual bleeding in menopausal women, which was then clarified with the explanation that both perimenopausal women and those under hormone replacement therapy were included and, therefore, would experience bleeding.

Question 6, for instance, "I get palpitations or a sensation of butterflies in my stomach or chest" was altered since the sensation of butterflies in one's stomach or chest is not an expression or metaphor in the Portuguese language or culture, being thus, adapted to "beatings or palpitations" in the stomach or chest.

Table 1 shows the sociodemographic data of women through which the assessments of the measuring properties were carried out.

The reliability of a measuring tool refers to the measurement procedure as it is repeated, and to its own homogeneous results. It is the condition of yielding either the same, or very similar results, when submitted to retests.

Table 2 shows the results of test-retest and absolute agreement reliability regarding each item of the Portuguese version of the WHQ.

Table 3 shows the internal consistency reliability (Cronbach alpha coefficient) between the various dimensions of the Portuguese version of WHQ.

Similar to reliability, which assesses measurement consistency, the validity estimates if a quality of life scale measures what it intends to. However, whereas reliability can easily be determined with very few indicators, validity is almost always a continuous process and requires comparisons with other questionnaires used to measure the same phenomenon.

The overall and specific dimension scores of the Portuguese WHQ version were correlated with the clinical parameters adopted, the Kupperman meno-

Table 2 - Intra-class correlation coefficient of each subscale for the Portuguese WHQ version, intra-observer (test-retest) and inter-observer (absolute agreement).

| Dimensions | Correlation coefficient | |
|------------------------|-------------------------|--------------------|
| | Test-retest | Absolute agreement |
| Somatic symptoms | 0.89 | 0.85 |
| Depressed mood | 0.92 | 0.87 |
| Cognitive difficulties | 0.86 | 0.68 |
| Anxiety/ fear | 0.85 | 0.85 |
| Sexual functioning | 0.74 | 0.79 |
| Vasomotor symptoms | 0.91 | 0.72 |
| Sleep problems | 0.90 | 0.35 |
| Menstrual problems | 0.69 | 0.18 |
| Attraction | 0.74 | 0.75 |
| Total (95% CI) | 0.92 (0.86-0.96) | 0.84 (0.71-0.92) |

Table 3 - Internal consistency reliability (Cronbach alpha coefficient) among the various domains of the Portuguese WHQ version.

| | Somatic symptoms | Depressed mood | Cognitive difficulties | Anxiety | Sexual problems | Vasomotor symptoms | Sleep problems | Menstrual problems | Attraction |
|------------------------|------------------|----------------|------------------------|---------|-----------------|--------------------|----------------|--------------------|------------|
| Somatic symptoms | 1 | | | | | | | | |
| Depressed mood | 0.51 | 1 | | | | | | | |
| Cognitive difficulties | 0.50 | 0.43 | 1 | | | | | | |
| Anxiety | 0.73 | 0.50 | 0.71 | 1 | | | | | |
| Sexual functioning | 0.11 | 0.45 | 0.23 | 0.12 | 1 | | | | |
| Vasomotor symptoms | 0.52 | 0.31 | 0.25 | 0.53 | 0.19 | 1 | | | |
| Sleep problems | 0.57 | 0.40 | 0.37 | 0.58 | 0.14 | 0.60 | 1 | | |
| Menstrual problems | 0.58 | 0.29 | 0.37 | 0.53 | -0.07 | 0.16 | 0.29 | 1 | |
| Attraction | 0.42 | 0.46 | 0.23 | 0.32 | -0.02 | 0.20 | 0.19 | 0.30 | 1 |

pausal index and the numeric scale, including the four most relevant items. It was observed a strong agreement between them except for the sexual functioning item, which did not show a significant correlation with any of the items considered according to the numeric scale, as shown in Table 4.

When both the overall and individual dimension scores of the Portuguese WHQ version were correlated with other quality of life parameter measurements, such as the SF-36, statistically significant correlations between all dimensions were found. Exceptions were found for sexual functioning, which only had a statistically significant correlation with the vitality and mental health dimensions; vasomotor symptoms only had a statistically significant correlation with the pain factor; and the attraction factor did not show a statistically significant correlation with the functional capacity dimension of the SF-36.

The utility only failed to have a statistically significant correlation with the vasomotor symptoms and menstrual problems dimensions. In this case, the expected correlations are negative, given that the higher the score on the Portuguese WHQ version (the more symptomatic the patient), the lower the SF-36 and the utility scores.

Responsiveness refers to the sensibility that an instrument has to identify a change in people's health conditions after an intervention, whether positive or negative in its effect on the score. It refers to a very relevant characteristic, mainly in clinical trial environments.

Table 5 shows the dimensional responsiveness in the Portuguese WHQ version and in the Kupperman menopausal index with values that ranged from 0.91 for menstrual problems to 5.79 for vasomotor symptoms. The higher the ratio – average change/standard error –, the higher is the alteration in conditions in the various indexes and dimensions after clinical intervention.

DISCUSSION

Studies on quality of life have their roots primarily based in social sciences. They recognize that although there is still convergence of perception in different populations, there is a basic presumption that every people or culture has their peculiar ways of feeling and understanding quality of life according to their beliefs, attitudes, and religious rites. Increasingly unequal socioeconomic situations lead to different perceptions, behaviors, and opportunities within different communities.³

The WHQ has basically been developed for assessment purposes, aiming at quantifying its change over the years. Therefore, there is the need to carry out procedures and tests, which are more complex than a simple translation, but validate its application in specific cultural contexts.⁵

During translation and cultural adaptation, the WHQ required some simple modifications before it was ready for use.

The intraclass correlation coefficient of each

Table 4 - Pearson correlation coefficient between the sub-scales and overall score of Portuguese WHQ version and the Kupperman index and numerical scale.

| WHQ | Somatic symptoms | Depressed mood | Cognitive difficulties | Anxiety | Sexual functioning | Vasomotor symptoms | Sleep problems | Menstrual problems | Attraction |
|--------------------|------------------|----------------|------------------------|----------|--------------------|--------------------|----------------|--------------------|------------|
| Vasomotor symptoms | 0.3652 | 0.5167 | 0.3562 | 0.4318 | 0.3030* | 0.7818 | 0.3037** | 0.2327* | 0.1305* |
| Paresthesia | 0.5445 | 0.2823** | 0.2612** | 0.3393** | 0.2863* | 0.0380* | 0.2472** | 0.3447*** | 0.2407** |
| Nervousness | | | | | | | | | |
| Melancholy | 0.3129** | 0.7628 | 0.4415 | 0.5217 | 0.2613* | 0.3343*** | 0.3639*** | 0.3409*** | 0.2660** |
| Insomnia | 0.3245 | 0.3487 | 0.2913 | 0.2789 | 0.1772* | 0.4063*** | 0.7396 | 0.2253* | 0.0866* |
| Kupperman Index | 0.6981 | 0.5538 | 0.4906 | 0.6539 | 0.4036** | 0.6280 | 0.5773 | 0.3415*** | 0.3396*** |

p<0.001

*Not significant; **p<0.05; ***p<0.01

Table 5 - Responsiveness of the Portuguese WHQ version subscales and the Kupperman menopausal index, assessed before and after hormonal replacement therapy.

| | Mean change | Standard error | Mean change/ Standard error |
|------------------------|-------------|----------------|--------------------------------|
| Kupperman index | -11.90 | 2.21 | 5.38 |
| WHQ | | | |
| Vasomotor symptoms | -3.05 | 0.52 | 5.79 |
| Sleep problems | -2.40 | 0.66 | 3.60 |
| Depressed mood | -4.15 | 1.18 | 3.51 |
| Sexual functioning | -2.13 | 0.60 | 3.50 |
| Somatic symptoms | -2.73 | 0.95 | 2.85 |
| Cognitive difficulties | -1.45 | 0.56 | 2.54 |
| Anxiety/ fear | -1.55 | 0.96 | 1.60 |
| Attraction | -0.85 | 0.65 | 1.29 |
| Menstrual problems | -0.57 | 0.62 | 0.91 |

The sign (-) means reduced symptoms

subscale for the WHQ Portuguese Version, which ranged from 0.69 to 0.92, and total of 0.92 (95% CI: 0.86-0.96) in test-retest and from 0.18 to 0.87, and total of 0.84 (95% CI: 0.71-0.92) in absolute agreement, has been considered to be good and its values clinically satisfactory.

The internal consistency reliability coefficient between a wide range of dimensions of the WHQ Portuguese version showed a Cronbach alpha coefficient ranging from -0.02 (sexual problems and attraction) to 0.73 (somatic symptoms and anxiety), with most ranging between 0.40 and 0.60, which indicate good values.

Only 36 among 87 women responded to the sexual functioning item, suggesting low sexual activity – even for women with partners – or some kind of situational or cultural constraint they may have been exposed to.

The high correlation between items suggests some redundancy. Low rates of agreement, or non-agreement, suggest that a specific question would preferably be used to assess some other dimension, rather than that one under scrutiny.

In validation tests, which correlated it with other quality of life assessment tools and with previously mentioned clinical parameters, statistically relevant correlations were found in the great majority of scores and dimensions, both with SF-36 and utility. However, there were found repeated low sexual function correlations with SF-36 dimensions, which correlated significantly only with vitality -0.41 (95% CI: -0.2 to -0.59) and mental health, -0.41 (95% CI: -0.19 to -0.59). Unexpectedly, however, there was also a positive correlation of vasomotor symptoms that barely reached statistical significance in relation to bodily pain, -0.26 (95% CI: -0.22 to -0.47).

A likely explanation for this is that, since hot flushes represent the most pronounced and expected

symptoms of menopause, women each in their own way, prepare themselves to face them sooner or later, accepting them as an unavoidable manifestation of the process.

By means of the Kupperman menopausal index, statistically significant correlations were found in all dimensions, thus underscoring its clinical validity.

Unlike reliability validity assessments, a consensus has yet to be reached as to the best way to assess responsiveness, mostly due to the difficulty in determining what could be considered a minimally significant difference, clinically speaking, following a given intervention.⁶

In this study, the determination of relative responsiveness was applied, and as shown in Table 5, and as expected for this particular clinical intervention (HRT), vasomotor symptoms showed the most significant change, superior even to the usual index, the Kupperman menopausal index, and the questionnaire had adequate sensitivity to detect it.

Other dimensions showing good change rates were: sleep problems, depressed moods, and sexual functioning.

Therefore, current methodological evidence suggests the high quality of the scale in measuring and comparing climacteric women's quality of life over time or following intervention. It shows high reliability and high validity in the process of construct validation.

Despite the brevity of the Portuguese WHQ version, it covers the full range of relevant climacteric factors.

A significant effort in perfecting HRT-trial measurement scales, psychological interventions, and general prevention for middle-aged and older women is called for.

REFERENCES

1. Barr JT. The outcomes movement and health status measures. *J Allied Health* 1995;24:13-28.
2. Carr AJ, Thompson PW, Kirwan JR. Quality of life measures. *Br J Rheumatol* 1996;35:275-81.
3. Ciconelli RM. Tradução para o português e validação do questionário genérico de avaliação de qualidade de vida "medical outcomes study 36-item short-form health survey (SF-36)" [dissertação]. São Paulo: Universidade Federal de São Paulo, Escola Paulista de Medicina; 1997.
4. Ferraz MB, Oliveira LM, Araújo PMP, Atra E, Tugwell P. Crosscultural reliability of the physical ability dimension of the health assessment questionnaire. *J Rheumatol* 1990;17:813-7.
5. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures literature review and proposed guidelines. *J Clin Epidemiol* 1993;46:1417-32.
6. Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis* 1987;40:171-8.
7. Hunter M, Battersby R, Whitehead M. Relationships between psychological symptoms, somatic complaints and menopausal status. *Maturitas* 1986;8:217-28.
8. Hunter MS. The Women's Health Questionnaire (WHO): Frequently Asked Questions (FAQ). *Health Qual Life Outcomes* 2003;1:41.
9. Kupperman HS, Blatt MHG, Wiesbader H, Filler W. Comparative clinical evaluation of estrogenic preparations by the menopausal and amenorrheal indices. *J Clin Endocrinol Metab* 1953;13:688-703.
10. Limousin-Lamothe M-A, Mairon N, Joyce CRB, Le Gal M. Quality of life after the menopause: influence of hormonal replacement therapy. *Am J Obstet Gynecol* 1994;170:618-24.
11. Medina AP, Rodríguez Malagón MN, Gil Laverde JFA, Ramírez Rodríguez GA. Tamaño de la Muestra [monografía en CD-ROM]. Bogotá: Pontificia Universidad Javeriana; 2001.
12. Sechrest L, Fay TL, Hafeez Zaid SM. Problems of translation in cross-cultural research. *J Cross-cult Psychol* 1972;3:41-56.
13. Silva Filho CR. Qualidade de vida no climatério [dissertação]. São Paulo: Universidade Federal de São Paulo, Escola Paulista de Medicina; 1998.
14. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med* 1998;17:101-10.
15. Ware JE Jr, Gandek B. Overview of the SF-36 health survey and the International Quality of Life Assessment (IQOLA) Project. *J Clin Epidemiol* 1998;51:903-12.
16. Wiklund I, Karlberg J, Lindgren R, Sandin K, Mattsson L-A. A Swedish version of the women's health questionnaire: a measure of postmenopausal complaints. *Acta Obstet Gynecol Scand* 1993;72:648-55.