The Polycystic Ovary Syndrome Health-Related Quality of Life Questionnaire (PCOSQ): A validation

Article *in* Human Reproduction · February 2004

DOI: 10.1093/humrep/deh048 · Source: PubMed

13 authors, including:

Georgina Jones
The University of Sheffield
111 PUBLICATIONS **3,924** CITATIONS
SEE PROFILE

Ksenija Benes
The Royal Wolverhampton NHS Trust
16 PUBLICATIONS **580** CITATIONS
SEE PROFILE

Manisha Palep-Singh
Creation, Centre for Infertility & Assisted Reproduction, Fortis Hospitals, India
16 PUBLICATIONS **504** CITATIONS
SEE PROFILE

Adam Balen
Leeds Teaching Hospitals NHS Trust
379 PUBLICATIONS **32,619** CITATIONS
SEE PROFILE

Human Reproduction Vol.19, No.2 pp. 371–377, 2004 DOI: 10.1093/humrep/deh048

The Polycystic Ovary Syndrome Health-Related Quality of Life Questionnaire (PCOSQ): a validation

G.L.Jones[1,5], K.Benes[2], T.L.Clark[2], R.Denham[2], M.G.Holder[2], T.J.Haynes[2], N.C.Mulgrew[2], K.E.Shepherd[2], V.H.Wilkinson[2], M.Singh[3], A.Balen[3], H.Lashen[4] and W.L.Ledger[4]

[1]University of Sheffield Institute of General Practice and Primary Care, Community Sciences Building, Herries Rd, Northern General Hospital, Sheffield S5 7AU, [2]University of Sheffield School of Medicine and Biomedical Sciences, Beech Hill Rd, Sheffield S10 2RX, [3]The General Infirmary at Leeds, Clarendon Wing, Belmont Grove, Leeds LS2 9NS and [4]University of Sheffield Academic Unit of Reproductive and Developmental Medicine, The Jessop Wing, Tree Root Walk, Sheffield S10 2SF, UK

[5]To whom all correspondence should be addressed. E-mail: G.L.Jones@sheffield.ac.uk

BACKGROUND: We wished to evaluate the psychometric properties of the Polycystic Ovary Syndrome Questionnaire (PCOSQ), a questionnaire developed to measure the health-related quality of life (HRQoL) of women with polycystic ovary syndrome. METHOD: To assess reliability and validity, women recruited from an outpatient gynaecology clinic at the Jessop Wing, Royal Hallamshire Hospital, Sheffield completed two copies of the PCOSQ and the Short Form-36 (SF-36). Secondary factor analysis was carried out to verify the composition of the dimensions. Semi-structured interviews were conducted to assess face validity. RESULTS: Of the 92 women who consented, 82 women (89%) returned questionnaires at time 1, and 69 women (75%) returned questionnaires at time 2. All five PCOSQ dimensions were internally reliable with Cronbach's α scores ranging from 0.70 to 0.97. Intra-class correlation coefficients to evaluate test–retest reliability were high (range 0.89–0.95, $P < 0.001$). Construct validity was demonstrated by high correlations for all comparisons of similar scales of the SF-36 and PCOSQ (0.49 and 0.54). Acne was identified as an important area of HRQoL missing from the questionnaire. CONCLUSIONS: The PCOSQ is a reliable instrument for measuring the HRQoL in women with PCOS. However, the validity of the questionnaire needs to be improved by incorporating a dimension on acne into the instrument.

Key words: health-related quality of life/health status/polycystic ovary syndrome/psychometric properties/questionnaires

Introduction

Polycystic ovary syndrome (PCOS) affects 5–10% of women in the developed world, making it the most common endocrine disorder among women of reproductive age (for review see Franks, 1995; Solomon, 1999). It is typically defined as the association of hyperandrogenism with chronic anovulation in women without specific underlying disease of the adrenal or pituitary glands (Franks, 1995). PCOS is diagnosed on the clinical picture, supported in some women by biochemical abnormalities and/or polycystic ovaries on ultrasonography (Zawadski and Dunaif, 1992).

The symptoms typically associated with PCOS—amenorrhoea, oligomenorrhoea, hirsutism, obesity, subfertility, anovulation and acne—can lead to a significant reduction in quality of life. For example, hirsutism has been shown to cause marked psychological stress (Sonino *et al*., 1993) and infertility issues can cause tensions within the family, altered self-perception, and problems at work (Paulson *et al*., 1988; Downey *et al*., 1989). Despite this, a recent systematic review revealed that limited research had been carried out to assess the impact that the symptoms and associated treatments for PCOS have upon the quality of life of women with the condition (Jones *et al*., 2002).

Health-Related Quality of Life (HRQoL) is a multi-dimensional, dynamic concept that encompasses physical, psychological and social aspects that are associated with a particular disease or its treatment (Naughton and McBee, 1997; Colwell *et al*., 1998). Although HRQoL measurement has an important role in evaluative research, the reliable assessment of quality of life depends upon the psychometric properties of the questionnaire (i.e. the tests underlying the construction and evaluation of the questionnaire) and the statistical methods employed to analyse and interpret the data (Fayers and Machin, 2000). It is important therefore that any HRQoL questionnaire to be used is based upon these psychometric properties. Although there are many tests which can be performed to evaluate these properties, the general consensus is that they should be reliable, valid and sensitive to change (Nunally, 1978).

At present, one reason for the limited research on the impact of PCOS upon quality of life may be because no validated health outcome measure exists to measure the health status of

women with the condition. One disease specific questionnaire has been developed, the Polycystic Ovary Syndrome Questionnaire (PCOSQ) (Cronin *et al*., 1998). It contains 26 items, measuring the following five areas of HRQoL: emotions (eight items, e.g. moody as a result of having PCOS?), body hair (five items, e.g. growth of visible hair on chin?), weight (five items e.g. had trouble dealing with your weight?), infertility problems (four items, e.g. concerned with infertility problems?) and menstrual problems (four items, e.g. irregular menstrual periods?). However, only the content validity of the instrument had been evaluated, thus preventing its use in clinical settings and limiting the research that can be carried out to evaluate the impact of PCOS-associated symptoms and their treatment upon quality of life. While generic questionnaires exist to measure HRQoL, such as the SF-36, they may not be sensitive enough to measure changes in specific illnesses as they were designed to measure health status across a wide variety of diseases (Streiner and Norman, 2000).

Consequently, the aim of this study was to evaluate the other psychometric properties of the PCOSQ, in particular the reliability, validity and factor structure of the domains when assessing the HRQoL in women with PCOS.

Materials and methods

Ethical approval for this study was obtained from the South Sheffield Research Ethics Committee.

A total of 186 women of reproductive age with PCOS was recruited from a gynaecology clinic at the Jessop Wing, Royal Hallamshire Hospital, Sheffield. We defined the inclusion criteria as two out of three of the following: a physical symptom, a biochemical abnormality or polycystic ovaries visualized on ultrasound scan (Lewis, 2001). Physical symptoms included hirsutism, oligo/amenorrhoea, infertility, body mass index >28 kg/m^2 or acne (Heineman, 1997). Biochemical abnormalities included an LH/FSH ratio >1.5, a testosterone level >2.0 or sex hormone binding globin <30 (Balen, 1999). Women were excluded from the study if they had another major illness that substantially influenced their quality of life or another cause of androgen excess, e.g. congenital adrenal hyperplasia. The patients' notes were reviewed to identify those who met the study inclusion criteria, from which 172 women were eligible. These women were sent a consent form and a letter inviting them to participate, to which 92 (53.5%) responded.

The PCOSQ was then administered by postal survey to the 92 women who had consented (time 1). Included with the questionnaire was the SF-36 (Ware *et al*., 1992). The SF-36 was used for two reasons. First, it has been argued that different health status measures (ideally generic and disease-specific health questionnaires) should be used in studies concerned with quality of life measurement (Fitzpatrick *et al*., 1993). Second, it was necessary to include another instrument to evaluate the construct validity of the PCOSQ.

Construct validity is a powerful measure of evaluating the validity of an instrument (Kline, 1986). It is usually evaluated by testing the instrument against hypotheses concerning the scores in the test (Kline, 1986). The test is said to have demonstrated construct validity if the hypothesis is supported but poor construct validity if it is rejected (Kline, 2000). In the absence of another disease-specific PCOS questionnaire, the SF-36 was chosen. It is a well-validated generic questionnaire (Kosinski *et al*., 1999) and contained two scales which could be used to help evaluate the construct validity of the PCOSQ.

A second copy of the PCOSQ and SF-36 was also included in a sealed envelope and respondents were requested to complete the second copies 3–6 days after they had completed and returned the first (time 2). This was to evaluate the test–retest reliability of the questionnaire. Test–retest reliability (i.e. how stable the questionnaire is over time) is fundamental if the purpose of the instrument is to measure outcome (McDowell and Newell, 1996). Usually a test is administered to a set of subjects on two occasions (given that there has been no change during this time) and then the scores obtained from the test and the retest are correlated (Kline, 1986). A 'change' letter was also included in the sealed envelope for the respondent to report any important changes in their health status. This was because test–retest reliability would only be analysed on those patients who reported 'no change' to their health status during that time.

Face validity is concerned with how appropriate, relevant and understandable items on a questionnaire are to the focus or aim of the questionnaire (Jenkinson and McGee, 1998). Although on its own, face validity does not ascertain the true validity of a questionnaire, it is important to establish as it can improve the co-operation of respondents completing a questionnaire (Kline, 2000), identify any ambiguities in the wording of items (Jenkinson and McGee, 1998) and identify any irrelevant or missed-out items.

To check the face validity of the PCOSQ, individual interviews with 12 women with PCOS recruited from an outpatient gynaecology clinic at the Jessop Wing Hospital, Sheffield were carried out. These explored whether the PCOSQ was addressing the relevant issues regarding the impact of PCOS upon the health status of women suffering from the condition. This sample size was determined at the point where no new issues emerged regarding the face validity of the questionnaire (Peto *et al*., 1998).

Analysis

The PCOSQ consists of five domains, each relating to a common symptom of PCOS; body hair, emotions, infertility, menstrual problems and infertility. Each question on the PCOSQ is associated with a 7-point scale in which 7 represents optimal function and 1 the poorest function. In order to compare it with the SF-36, each question was re-coded from 0 to 6 in which 6 represents optimal function and 0 the poorest. Each scale was then transformed on a range from 0 (indicating worst health status) to 100 (best health status) enabling the extent of ill health to be measured (scale score = total of raw scores for each item in the scale/maximum possible raw score$\times$100). This was repeated for time 2 responses. Similar tables were constructed for the SF-36.

To verify the factor structure and compositions of the PCOSQ dimensions, secondary factor analysis was used. Factor analysis is a statistical procedure which enables the underlying dimensions (or scales) of a questionnaire to be determined (Kline, 2000). The data from the questionnaires returned at time 1 were analysed using principal component analysis (varimax rotation) as used in the development of the original questionnaire (Cronin *et al*., 1998). To reduce statistical error, it has been postulated that $\geqslant$100 subjects are needed for factor analysis or a $\geqslant$2:1 ratio of subjects to items (Kline, 2000).

To measure the internal consistency reliability of the questionnaire, Cronbach's α statistic was used. Internal consistency reliability is an indication of how well the items within a scale are associated with each other or their 'homogeneity' (Velikova *et al*., 1999) and Cronbach's α is the measure which is most frequently used for establishing this. Scores >0.7 usually indicate that scale items are measuring related constructs (Cronbach, 1951). Item-total consistency was also calculated to check the internal reliability of a dimension. This is the extent to which there is a linear relationship between an

item and its scale score, which has been corrected for overlap (Gandek *et al*., 1998). To correct for overlap, the item which is to be correlated with the scale is omitted from the scale total. A correlation coefficient of ⩾0.40 indicates satisfactory item-total consistency (Ware *et al*., 1980).

To analyse test–retest reliability, the Wilcoxon signed rank test (nonparametric) was used to calculate if there were any statistical differences between the scores at times 1 and 2 on those patients who had reported 'no change' to their health status during that time. The intra-class correlation coefficient was calculated to examine the relationship between the scale scores at time 1 and 2. The results were taken to be significant if $P < 0.05$.

To assess construct validity, we hypothesized a significant correlation (Spearman's ρ non-parametric coefficient) would be found between scales with a similar content on the PCOSQ and SF-36 (the 'emotions' scale of the PCOSQ with the domains 'mental health' and 'role-emotional' of the SF-36). We also hypothesized that these two domains of the SF-36 would correlate more strongly with the 'emotions' domain of the PCOSQ than any of its other four scales. All statistical analyses were performed using SPSS v 11.5.

Table I. Clinical features of the patient sample

Clinical features	No. of total data set affected (%) (n = 186)	No. of sample set affected (%) (n = 92)
Symptoms		
Excessive body hair	94 (50.6)	46 (50.0)
BMI >28 kg/m^2	100 (53.8)	48 (52.2)
Acne	54 (29.0)	31 (33.7)
Infertility	120 (64.5)	59 (64.1)
Oligo/amenorrhoea	160 (86.0)	80 (87.0)
Biochemical abnormalities		
LH/FSH ratio > 1.5	89 (47.8)	50 (54.3)
Testosterone > 2.0	43 (23.1)	40 (43.5)
SHBG < 30	19 (10.2)	28 (30.4)
PCO diagnosed on ultrasound	133 (71.5)	75 (81.5)

BMI = body mass index; SHBG = sex hormone-binding globulin; PCO = polycystic ovaries.

Results

Of the 92 women who agreed to take part in the study, 82 returned questionnaires at time 1 (89.1%) and 69 at time 2 (75.0%). The mean age of the sample was 29.4 ± 5.7 (range 20–41). Ethnicity was also recorded; 78 (84.9 %) of the respondents were white, five (5.4%) were Pakistani, 1 (1.1%) was Chinese, and no ethnicity data were provided for seven (7.6%) respondents. The clinical features of the sample are reported in Table I.

Secondary factor analysis was carried out on the 82 questionnaires returned at time 1, producing a ratio of 3:1 items to subjects. The results from the secondary factor analysis (principal component analysis, varimax rotation) are shown in Table II. Initially, only factors which gained an eigenvalue (raw sum of the squares) of ⩾1 were retained. This procedure identified six factors which accounted for 78.8% of

Table II. Secondary factor analysis on the Polycystic Ovary Syndrome Health-Related Quality of Life Questionnaire (PCOSQ)

Items	Rotated component matrix mactors					
	1	2	3	4	5	6
W3. Concerned about being overweight	**0.841**					
W10. Had trouble dealing with weight	**0.851**					
W12. Felt frustration trying to lose weight	**0.877**					
W22. Feel not sexy because overweight	**0.704**					
W24. Difficulties staying at ideal weight	**0.847**					
BH1. Growth of visible hair on chin		**0.823**				
BH9. Growth of visible hair on upper lip		**0.793**				
BH15. Growth of visible hair on face		**0.873**				
BH16. Embarrassment of excess body hair		**0.898**				
BH26. Growth of visible body hair		**0.881**				
EM2. Depressed having PCOS			**0.756**			
EM4. Easily tired			0.345			
EM6. Moody as a result of having PCOS			**0.567**			
EM17. Worried about having PCOS			**0.765**			
EM18. Self-conscious having PCOS			**0.744**			
EM11. Had low self-esteem having PCOS			**0.801**			
INF23. Feel a lack of control over PCOS			0.652			
EM14. Felt frightened of getting cancer				0.497		
INF5. Concerned with infertility problems				**0.914**		
INF13. Felt afraid of not having children				**0.823**		
INF25. Feel sad because of infertility				**0.902**		
MEN7. Headaches					**0.757**	
MEN19. Abdominal bloating					**0.722**	
MEN21. Menstrual cramps					**0.682**	
EM20. Late menstrual period						0.797
MEN8. Irregular menstrual periods						0.789

Extraction method: prinicipal component analysis; rotation method: varimax with Kaiser normalization.
A factor indicates the relationships between a set of items. A factor is therefore defined by the items which load on it or the factor loadings. Loadings >0.5 are considered satisfactory.
Bold type indicates the items which loaded on original factors as determined by Cronin *et al*. (1998).
W = items from weight domain; BH = body hair; EM = emotions; INF = infertility; MEN = menstrual; PCOS = polycystic ovary syndrome.

Table III. Distribution of the scale scores for the Polycystic Ovary Syndrome Health-Related Quality of Life Questionnaire (PCOSQ) and Short Form-36 (SF-36) from the questionnaires returned at time 1

Scale	*n*	Mean	SD	Range of scores	% scoring minimum (floor)	% scoring maximum (ceiling)
PCOSQ						
Body hair	81	55.6	34.1	0–100	3.7	18.5
Emotions	78	49.8	30.0	8–100	2.6	2.6
Infertility	81	39.3	31.4	0–100	13.6	4.9
Menstrual problems	81	44.3	24.7	0–100	3.7	2.5
Weight	82	32.4	33.1	0–100	22.0	4.9
SF-36						
Bodily pain	82	69.7	22.7	27–100	4.9	20.7
Energy and vitality	82	51.5	16.7	17–100	2.4	1.2
General health	80	64.4	18.3	30–100	1.3	5.0
Mental health	82	63.4	16.3	23–100	1.2	2.4
Physical functioning	80	89.8	11.0	47–100	1.3	23.8
Role limit/emotion	82	50.4	40.3	0–100	29.3	30.5
Role limit/physical	82	69.5	36.7	0–100	12.2	50.0
Social functioning	82	67.4	23.4	20–100	1.2	23.2
Change in health	82	60.0	17.3	20–100	3.7	6.1

The scales run from 0 (poor health) through to 100 (good health).

Table IV. Internal reliability and test–retest reliability correlation of scales generated from time 1 and time 2 on the Polycystic Ovary Syndrome Health-Related Quality of Life Questionnaire (PCOSQ) and the Short Form-36 (SF-36)

Scale	Internal reliability, time 1[a] α (*n*)	Internal reliability, time 2[b] α (*n*)	Differences between time 1 and time 2[c] Wilcoxon signed rank test (*Z*)*	Test–retest reliability[c] Intra-class correlation**
PCOSQ				
Body hair	0.95 (81)	0.97 (69)	–1.44	0.95 (60)
Emotions	0.88 (78)	0.91 (67)	–0.86	0.93 (56)
Infertility	0.87 (81)	0.91 (68)	–1.82	0.92 (59)
Menstrual problems	0.70 (81)	0.73 (67)	–1.35	0.89 (56)
Weight	0.96 (82)	0.95 (68)	–1.48	0.95 (54)
SF-36				
Bodily pain	0.91 (82)	0.90 (68)	–0.29	0.67 (61)
Energy and vitality	0.79 (82)	0.83 (68)	–0.97	0.79 (61)
General health	0.81 (80)	0.83 (68)	–0.20	0.88 (59)
Mental health	0.77 (82)	0.79 (66)	–1.46	0.80 (59)
Physical functioning	0.85 (80)	0.83 (67)	–1.63	0.84 (59)
Role limit/emotion	0.74 (82)	0.88 (68)	–0.19	0.89 (61)
Role limit/physical	0.82 (82)	0.85 (68)	–1.42	0.75 (61)
Social functioning	0.83 (82)	0.85 (68)	–1.27	0.81 (60)

Values in parentheses are numbers (*n*).
Cronbach's α-values indicate how well the items within a scale are associated with each other and are used to establish internal reliability; α-values >0.7 are usually considered satisfactory.
[a]Everyone who completed questionnaires from the survey at time 1.
[b]Everyone who completed questionnaires from the survey at time 2.
[c]Everyone who completed questionnaires at time 1 and time 2 and reported no significant change in their health/life during that period.
*$P > 0.05$ for all values.
**$P < 0.001$ for all values.

the variance. Only those questions which obtained a value of ⩾0.50 on any of the factors were initially retained and any factors which scored less than this were omitted.

From this analysis, two dimensions (weight and body hair) were identical to the initial composition of the scales on the PCOSQ, with the same items loading on each factor.

The infertility scale was identical except that item 23, 'feel a lack of control over the situation with PCOS', loaded on the emotions scale instead. The remaining two scales were the same except that two original items (frightened of getting cancer, and late menstrual period) did not load on the emotions scale and one item (irregular menstrual periods) did not load on the menstrual problems domain. Only one item failed to obtain a value of 0.50 which was originally in the emotions scale (easily tired) but this was close at 0.40.

The distribution of the scale scores for the five domains on the PCOSQ and SF-36 are shown in Table III. The mean scale scores for the PCOSQ showed that 'weight' (32.4) and 'infertility' (39.3) scored the lowest, indicating worst health in these dimensions. This was in comparison to the body hair domain which had the highest mean score (55.6) suggesting that this symptom was causing the least negative impact upon

quality of life. The SF-36 questionnaire results indicated that the 'role limitation–emotional' and 'energy and vitality' domains were the areas of poorest health with mean scores of 50.4 and 51.5 respectively. The highest mean score was evident for physical functioning (89.8), indicating the best area of health as measured on the SF-36.

The internal reliability consistency of the scales was assessed using Cronbach's α. As shown in Table IV, all the scales on the PCOSQ reached the required 0.7 at both time 1 (range 0.70–0.96) and time 2 (range 0.73–0.97). All the item-total correlations of the individual items exceeded the minimum value of 0.40 except for item 14 at time 1 on the emotions scale (felt frightened of getting cancer). However, this was close at 0.39. At time 2, both item 17 (worried about having PCOS) and item 8 (irregular menstrual periods) achieved an item correlation <0.40 with the scores reaching 0.13 and 0.37 respectively.

The SF-36 was tested in the same way. The Cronbach's α value results are shown in Table IV. At both times 1 and 2, all of the domains showed internal reliability with Cronbach's α scores >0.7 (range 0.74–0.91). The item-total correlation scores were >0.40 for all individual questions, except for three items in the physical functioning domain which achieved item-total correlations of 0.39 ('vigorous activities', time 1), 0.31 ('lifting or carrying groceries', time 2) and 0.12 ('bathing or dressing', time 2).

Sixty-nine women (75.0%) returned the questionnaires at time 1 and 2. Of these, 57 (82.61%) women reported no change in their health status during this time. No significant differences (Wilcoxon signed rank test) between the scores at times 1 and 2 were found on the five scales of the PCOSQ or SF-36 (Table IV).

The intra-class correlation coefficients to evaluate the test–retest reliability of the PCOSQ were high for all domains. As shown in Table IV, they ranged from 0.89 to 0.95 ($P < 0.001$). The SF-36 was less well correlated (0.67–0.89) with a greater P-value for some domains, showing less significance (Table IV).

Construct validity was assessed by correlating the 'emotions' scale of the PCOSQ with the 'mental health' and 'role–emotional' scales of the SF-36. As hypothesized, a good positive correlation was found for both analyses. Emotions (PCOSQ) correlated with role–emotional (SF-36) ($r = 0.49$, $P < 0.01$, $n = 55$). This correlation was greater than with the weight domain ($r = 0.37$, $P < 0.05$, $n = 47$), the infertility domain ($r = 0.37$, $P > 0.05$, $n = 58$) and the body hair domain ($r = 0.41$, $P > 0.05$, $n = 60$), but not for the menstrual domain ($r = 0.50$, $P < 0.01$, $n = 56$). Emotions (PCOSQ) correlated with mental health (SF-36) ($r = 0.62$, $P < 0.01$, $n = 54$) and this was more strongly correlated than with the other four scales of the PCOSQ, including infertility ($r = 0.43$, $P < 0.01$, $n = 53$), body hair ($r = 0.22$, $P > 0.05$, $n = 50$), weight ($r = 0.34$, $P < 0.05$, $n = 47$) and menstrual problems ($r = 0.59$, $P < 0.01$, $n = 50$).

Face validity

Of the 12 women interviewed to assess face validity, 11 felt that overall the questionnaire did address the areas of HRQoL negatively affected as a result of PCOS. However, some limitations with the instrument were reported.

Three women (25%) mentioned that questions relating to their symptom of acne were missing from the questionnaire. Three women (25%) were worried by the question that asked 'during the past two weeks, have you felt frightened of getting cancer?' (no. 16). One woman (8%) expressed concern that the wording of the questionnaire implied that PCOS was a short-term condition. Two women (17%) felt the questionnaire did not address their feelings of frustration about the lack of available information on PCOS. One woman (8%) mentioned the omission of a question about weight-related prejudice from infertility services.

Discussion

At present, no validated health outcome measure exists to measure the quality of life of women with PCOS. Although a disease-specific questionnaire has been developed, only the content validity of the instrument has been established, therefore preventing its use in clinical and research settings. While generic questionnaires exist to measure HRQoL, such as the SF-36, they do not collect information on all the areas of well-being and functioning that may be important to women with PCOS. For example, infertility and hirsutism can place a considerable strain on the emotional well-being and personal relationships of women with this disease. However, this information is not collected on the SF-36 generic health measure.

The mean scale scores on the PCOSQ reflect the negative impact PCOS can have upon the quality of life of women with the condition. Perhaps not surprisingly, weight and infertility appeared to be the most significant aspects of the illness. Other studies have reported the negative impact infertility can have upon women and their personal relationships (Epstein and Rosenberg, 1997; Leiblum and Greenfeld, 1997). The finding that weight caused the most negative impact on quality of life has implications for the management of the condition, especially as it has been estimated that ~50% of women with PCOS suffer from obesity or are overweight (Gambineri *et al*., 2002), yet the best way to manage and aid weight loss in this group of women is unclear (Moran and Norman, 2002).

The secondary factor analysis carried out in this new data set would suggest that overall the structure of the PCOSQ domains are verified, especially for weight, body hair, menstrual problems and infertility. The composition of the emotions scale was less supported. Two original items failed to load on this scale, and an item originally from the infertility scale (lack of control over PCOS) did, suggesting that there may be limitations with this scale. In addition, the item 'felt frightened of getting cancer' loaded on the infertility domain. This was consistent with the original factor analysis carried out by Cronin *et al*. (1998), although the authors subsequently moved this item to the emotions scale as they felt it to be more appropriate to that domain.

One limitation with the results produced from our factor analysis may be that the ratio of respondents to items was not large enough. Although it has been shown that a ratio of 2:1 of

respondents to variables is satisfactory and can produce similar results to a larger ratio (Kline, 2000), it has been argued that the ratio of the number of subjects to the number of items on the instrument should be ⩾5:1 (McDowell and Newell, 1996). A study containing such a ratio of respondents to items would be recommended for the future before amending the composition of the emotions scales on the PCOSQ.

The internal reliability of the PCOSQ was found to be high, with all the scales exceeding the accepted α value of 0.70, although the menstrual problems scale was weaker than the other domains with an α of just 0.7. This indicates that this scale may benefit from further analysis. This is further supported as item 20 (late menstrual period) and item 8 (irregular menstrual period) loaded on a new factor, suggesting that a new scale referring to menstrual periods specifically may be required.

The test–retest reliability of the PCOSQ was found to be high. All scales achieved high intra-class correlations >0.8 and were overall higher than for the SF-36. This indicates that the questionnaire produces consistent results from subjects at different times, when no evidence of change in health status exists.

The small sample size, and the absence of another questionnaire which contained similar domains to the PCOSQ, limited the testing of construct validity in this study. However, because scales on the PCOSQ correlated significantly with similar scales on the SF-36 as hypothesized and overall less with the other scales on the questionnaire, it suggests that the construct validity of the instrument is supported. Further analysis on the construct validity of the questionnaire needs to be carried out to verify this further.

In terms of face validity, the women interviewed felt that, on the whole, the questionnaire was addressing the relevant issues to women with PCOS. However, the lack of questions about acne was raised as a serious omission. Acne is recognized as a common symptom of the condition. The finding that 34% of respondents in the study suffered from acne would support this and suggest that the addition of a new acne domain to the PCOSQ would be important if the instrument is to be used in a clinical setting.

One explanation for the omission of an acne domain may be due to the item selection phase of the PCOSQ. During this phase, only 10 PCOS patients were interviewed. Although it was reported that these women had the full range of complaints, it is not known exactly what the symptom profile or presenting symptoms of these patients were. This small number was justified by the fact that no further items were generated after the first five interviews (Cronin *et al*., 1998). Although this may have been an acceptable number, other studies have interviewed more subjects to generate items for disease-specific questionnaires (Jenkinson *et al*., 1998; Jones *et al*., 2001).

There are two methodological issues which are important to discuss. First, the definition of PCOS used in our study differed from that in the original paper by Cronin *et al*. (1998). For this validation study, a definition of the syndrome commonly used in Europe and the UK was used (Balen *et al*., 1999; Lewis *et al*., 2001) instead of the National Institutes of Health (NIH) definition as used in the development of the PCOSQ. The aim of our study was to validate the use of the questionnaire in UK practice and we therefore have applied it to a group of women who have PCOS using the definition currently accepted in this country. For this reason, it was felt that using the NIH definition to recruit patients would not help ascertain the validity of the PCOSQ for measuring the HRQoL of PCOS patients within Europe and the UK.

Another potential limitation to the study is that patients were recruited from a gynaecology clinic. Consequently, it could be argued that there was a bias towards PCOS patients with menstrual disturbance and infertility and not those presenting with other PCOS-related conditions (i.e. weight increase and acne) that are often referred to endocrinology and dermatology clinics. However, it was felt that sufficient patients with complaints of a dermatological or endocrinological nature to allow us to validate the PCOSQ were seen in the gynaecology clinics. For example, in our group of patients 50.6% of the population had hirsutism and 29% of the population had acne. The findings are therefore of interest to dermatologists and endocrinologists and the validation of the questionnaire seems relevant to women with PCOS, not withstanding which subspecialist sees them. Furthermore, our finding that a domain on acne needs to be included was generated by the composition of our group of patients, which, in distinction from the patients selected to develop the instrument, did include a large number of women with acne and is therefore more representative of the PCOS population than used to develop the initial questionnaire.

In conclusion, the PCOSQ is a promising tool for measuring the HRQoL of women with PCOS. The next phase of our research is to develop and incorporate an acne domain into the PCOSQ. The amended PCOSQ will be tested by the same criteria on a new sample of PCOS patients, along with the 'sensitivity to change' of the new instrument. This is the ability of a questionnaire to detect and describe changes in patients' health status over time and whether these changes are clinically relevant (Kazis *et al*., 1989). An improved questionnaire could potentially be used to assess the full impact of treatment regimes, for example, in randomized controlled trials for patients with PCOS and be used to generate more understanding into the impact that symptoms and treatments for this condition have upon HRQoL.

References

Balen A (1999) Pathogenesis of polycystic ovary syndrome—the enigma unravels? Lancet 354,966–967.

Colwell H, Mathias SD, Pasta DJ, Henning JM and Steege JF (1998) A health-related quality of life instrument for symptomatic patients with endometriosis: a validation study. Am J Obstet Gynecol 179,47–55.

Cronbach L (1951) Coefficient alpha and the internal structure of tests. Psychometrika 16,297–334.

Cronin L, Guyatt G, Griffith L, Wong E, Azziz R, Futterweit W, Cook D and Dunaif A (1998) Development of a health-related quality-of-life questionnaire (PCOSQ) for women with polycystic ovary syndrome (PCOS). J Clin Endocrinol Metab 83,1976–1987.

Downey J, Yingling S, McKinney M, Husami N, Jewelewicz R and Maidman J (1989) Mood disorders, psychiatric syptoms and distress in women presenting for infertility evaluation. Fertil Steril 52,425–432.

Epstein YM and Rosenberg HS (1997) He does, she doesn't; she does, he doesn't: couple conflicts about infertility. In Leiblum SR (ed) Infertility:

Psychological Issues and Counselling Strategies. Wiley, New York, pp 129–148.

Fayers PM and Machin D (2000) Quality of Life: Assessment, Analysis and Interpretation. Wiley, Chichester, UK.

Fitzpatrick R, Ziebland S, Jenkinson C, Mowat A and Mowat A (1993) A comparison of the sensitivity to change of several health status instruments in rheumatoid arthritis. J Rheumatol 20,429–436.

Franks S (1995) Polycystic ovary syndrome. N Engl J Med 333,853–861.

Gambineri A, Pelusi C, Vicennati V, Pagotto U and Pasquali R (2002) Obesity and the polycystic ovary syndrome. Int J Obes Rel Metab Disord 26,883–896.

Gandek B Ware JE, Aaronson NK, Alonso J, Apolone G, Bjorner J, Brazier J, Bullinger M, Fukuhara, S, Kaasa, S et al (1998) Test of data quality, scaling assumptions, and reliability of the SF-36 in eleven countries: results from the IQOLA Project. International Quality of Life Assessment. J Clin Epidemiol 51,1149–1158.

Heineman MJ (1997) Introduction to polycystic ovary syndrome. Gynaecol Forum 2,3–4.

Jenkinson C and McGee H (1998) Health Status Measurement: A Brief But Critical Introduction. Radcliffe Medical Press, Oxford.

Jenkinson C, Fitzpatrick R and Peto V (1998) The Parkinson's Disease Questionnaire: user manual for the PDQ-39, PDQ-8, and PDQ summary index. Health Services Research Unit, University of Oxford, Oxford, UK.

Jones G, Kennedy S, Barnard A and Wong J (2001) Development of an endometriosis quality-of-life instrument: the Endometriosis Health Profile-30. Obstet Gynecol 98,258–264.

Jones G, Kennedy S and Jenkinson C (2002) Health-related quality of life measurement in women with common benign gynecological conditions: a systematic review. Am J Obstet Gynecol 187,501–511.

Kazis LE, Anderson JJ and Meenan RF (1989) Effect sizes for interpreting changes in health status. Med Care 27,S178–189.

Kline P (1986) A Handbook of Test Construction. Methuen & Co., London, UK.

Kline P (2000) A Psychometrics Primer. Free Association Books, London, UK.

Kosinski M, Keller SD, Hatoum HT, Kong SX and Ware JE (1999) The SF-36 Health Survey as a generic outcome measure in clinical trials of patients with osteoarthritis and rheumatoid arthritis: tests of data quality, scaling assumptions and score reliability. Med Care 37,MS10–22.

Leiblum SR and Greenfeld DA (1997) The course of infertility: immediate and long-term reactions. In Leiblum SR (ed) Infertility: Psychological Issues and Counselling Strategies. Wiley, New York, pp 83–102.

Lewis V (2001) Polycystic ovary syndrome: a diagnostic challenge. Obstet Gynecol Clin North Am 28,1–20.

McDowell I and Newell C (1996) Measuring Health: A Guide to Rating Scales and Questionnaires. Oxford University Press, Oxford, UK.

Moran LJ and Norma RJ (2002) The obese patient with infertility: a practical approach to diagnosis and treatment. Nutr Clin Care 5,290–297.

Naughton MJ and McBee WL (1997) Health-related quality of life after hysterectomy. Clin Obstet Gynecol 40,947–957.

Nunally JC (1978) Psychometric Theory. Tate McGraw Hill, New Dehli.

Paulson JD, Haarmann BS, Salerno RL and Asmar P (1988) An investigation of the relationship between emotional maladjustment and infertility. Fertil Steril 49,258–62

Peto V, Jenkinson C and Fitzpatrick R (1998) PDQ-39: a review of the development, validation and application of a Parkinson's disease quality of life questionnaire and its associated measures. J Neurol 245(Suppl 1),S10–14.

Solomon CG (1999) The epidemiology of polycystic ovary syndrome. Prevalence and associated disease risks. Endocrinol Metab Clin North Am 28,247–263.

Sonino N, Fava GA, Mani E, Bellurdo P and Boscaro M (1993) Quality of life in hirsuit women. Postgrad Med J 69,186–189.

Streiner DL and Norman G (2000) Health Measurement Scales: A Practical Guide to their Development and Use. 2nd edn, Oxford University Press, Oxford, UK.

Velikova G., Stark D and Selby P (1999) Quality of life instruments in oncology. Eur J Cancer 35,1571–1580.

Ware JE and Sherbourne EC (1992) The MOS 36-Item Short Form Health Survey 1: Conceptual Framework and Item Selection. Med Care 30,473–483.

Ware JE, Brook RH, Williams KN, Stewart AL and Davies-Avery A (1980) Conceptualisation and Measurement of Health for Adults in the Health Insurance Study, Volume 1: Model of Health and Methodology. RAND Corporation, Santa Monica, USA.

Zawadski JK and Dunaif A (1992) Diagnostic criteria for polycystic ovary syndrome: Towards a rational approach. In Dunaif A, Givens JR, Haseltine FP, Merriam GE (eds) Polycystic Ovary Syndrome (Series ed Hershman SM) Current Issues in Endocrinology and Metabolism. Blackwell Scientific Publications, Boston, USA, 377pp.

Submitted on April 3, 2003; resubmitted on June 30, 2003; accepted on September 26, 2003