Eur Respir J 1998; 11: 61–66
DOI: 10.1183/09031936.98.11010061
Printed in UK - all rights reserved

European Respiratory Journal
ISSN 0903 - 1936

Reliability and validity of a Swedish version of the St George's Respiratory Questionnaire

C-P. Engström*, L-O. Persson**, S. Larsson*, M. Sullivan+

Reliability and validity of a Swedish version of the St George's Respiratory Questionnaire. C-P. Engström, L-O. Persson, S. Larsson, M. Sullivan. ©ERS Journals Ltd 1998.
ABSTRACT: The St George's Respiratory Questionnaire (SGRQ) was designed to measure quality of life (QoL) in obstructive pulmonary disease. Its reliability, validity and sensitivity have been demonstrated. The aim was to develop a Swedish version of the SGRQ and to confirm its scaling and clinical properties.

The SGRQ was adapted for Swedish conditions following a translation-backtranslation procedure. The psychometric and clinical evaluation included 68 patients with chronic obstructive pulmonary disease (COPD). Supplementary QoL, clinical and physiological data were collected. A follow-up study was performed 1 yr later. Correlation analysis used a multitrait-multimethod model. Internal consistency reliability and discriminant validity were documented by performing a multitrait analysis.

The results confirmed expected levels of associations. Correlation coefficients between the SGRQ total score and the Sickness Impact Profile Total score (a generic health measure), forced expiratory volume in one second (FEV1) and 6 min walking distance were 0.69, -0.42 and -0.61 respectively. The pattern of correlations in the Swedish data set was very similar to that of the original. The stability of the SGRQ scores was confirmed at follow-up after 1 yr. The reliability was satisfactory, with Cronbach's alpha coefficients >0.80 for the SGRQ and its subdimensions.

In conclusion, the Swedish version of the St George's Respiratory Questionnaire is reliable, valid and compares well with the corresponding tests of the original version.
Eur Respir J 1998; 11: 61–66.

*Dept of Pulmonary Medicine, Sahlgrenska University Hospital, **Department of Nursing Göteborg College of Health Sciences, and +Health Care Research Unit, Dept of Internal Medicine, Sahlgrenska University Hospital, Göteborg, Sweden.

Correspondence: M. Sullivan
Health Care Research Unit
Dept of Internal Medicine
Sahlgrenska University Hospital
S-413 45 Göteborg
Sweden
Fax: 00 46 31826941

Keywords: Chronic obstructive pulmonary disease
cross-cultural adaptation
quality of life
translation

Received: December 10 1996
Accepted after revision August 28 1997

This study was supported by a grant from the Swedish Council for Social Research.

Impaired lung function in chronic obstructive pulmonary disease (COPD) leads to dyspnoea and to impaired exercise tolerance, and this in turn influences patients' quality of life (QoL) [1]. The importance of measuring QoL in COPD is no longer questioned today [2], and both generic and disease-specific questionnaires are applied. There is a clear advantage in assessing basic aspects of health by generic measures, but these carry the risk of not being able to detect small changes in health status [3, 4]. Therefore, disease-specific questionnaires have been developed. One of these, the St George's Respiratory Questionnaire (SGRQ), presented by Jones and co-workers in 1991 [5, 6], has been translated into a number of languages [7] and used in several studies including patients with COPD.

The aim of this study was to develop a Swedish version of the SGRQ and to confirm its basic psychometric properties, reliability and validity, and to compare it with the corresponding tests of the original version.

Materials and methods

Translation procedure

The translation followed an established forward-backward translation procedure, with independent translations and counter-translation. First, three independent translations (the authors C-P.E., M.S. and one psychologist) were pooled to a common version for back-translation into English by a translator whose native language was British English. The back-translation was almost identical to the source instrument. The final test version was then pilot tested in 10 patients with COPD and found to be well accepted and easy to fill in.

Design

A cross-sectional sample of patients with COPD was recruited, stratified according to severity of disease and examined extensively by clinical, physiological and QoL measures. All tests were performed in 1 day. A longitudinal follow-up was scheduled 12 months later. The design of the study was approved by the local Ethics Committee.

Patients

Details of procedures, patient selection and patient characteristics have been reported elsewhere [8]. Sixty eight patients with COPD were recruited from the Department of Pulmonary Medicine, Sahlgrenska University Hospital, Göteborg, Sweden. All of the patients were current smokers or former smokers with a consumption of at least 10 pack-years. Exclusion criteria were disabling or severe

For editorial comment see page 5.

diseases other than COPD or co-existence of other causes of impaired pulmonary function such as asthma. No patient was in an acute exacerbation phase of the disease at the time of investigation. The patients were stratified to three groups to give about equal representation of forced expiratory volume in one second (FEV_1), using the groups FEV_1 <30% predicted, FEV_1 30–50% pred FEV_1 50–80% pred. At follow-up, 52 patients (76.5%) participated. Among the 16 drop-outs, three patients had died, two had moved and 11 felt too tired to go through with the study or de-clared no interest.

In addition to the patient group, a "risk group" of 13 smokers or former smokers with spirometrical signs of COPD was included in the study design. All of these subjects considered themselves as healthy and had consequently never seen a doctor because of pulmonary symptoms. They were recruited from a study of spirometrical data in the population and underwent the same investigations as the patients. Data received from this risk group have been used in the present study of psychometric properties of the SGRQ.

Physiological measurement

Routine spirometry was performed with a Vitalograph spirometer (Selefa, Buckingham, Ireland) before and 15 min after inhalation of 1 mg terbutaline to reach optimal standardization. Transfer factor of the lung for carbon dioxide ($T_{L,CO}$) was measured with the single-breath method. Values for prediction were those described by Berglund *et al.* [9] and Salorinne [10], respectively. Arterial blood gases (partial pressure of oxygen (P_{O_2}) and partial pressure of carbon dioxide (P_{CO_2})) were measured in patients with an FEV_1 <50% pred. Finally, a 6 min walking distance test (6MWD) was performed following standardized instructions [11].

Clinical measurements

The clinicians' opinion on the health of the patients was expressed by the Karnofsky Performance Status scale (KPS) [12]. It has a range from 0–100, where 0 equals dead and 100 means normal. Days in hospital during the last 12 months were counted from hospital records.

QoL questionnaires

The SGRQ. This has 76 items divided into three sections: Symptoms (problems caused by specific respiratory symptoms); Activity (restriction of activity by dyspnoea); and Impacts (impact on everyday life caused by the disease). Every item has a predetermined weight. Component scores are calculated for each of the three sections and a total score is derived including all items. The scores range from 0–100% of possible distress. Thus a low score indicates good health. Recently, reference scores have been obtained from subjects with no history of respiratory disease (P.W. Jones, personal communication).

The Sickness Impact Profile (SIP). Functional status was measured by an extensively validated Swedish version of the SIP [13, 14]. It is a well-known, generic health status questionnaire, constructed to facilitate comparisons between different health conditions over a range of important functional aspects. It consists of 136 weighted items, grouped into 12 categories: ambulation (A); body care and movement (BCM); mobility (M); emotional behaviour (EB); social interaction (SI); alertness behaviour (AB); communication (C); work (W); sleep and rest (SR); eating (E); home management (HM); and recreation and pastimes (RP). The scale scores are expressed as a percentage of maximum dysfunction to form a 0–100 scale. A score of 0 indicates no dysfunction, a score 0–10 indicates slight to moderate dysfunction and a score >10 marked dysfunction. The scores of the categories A, BCM and M form a physical dimension (Physical) and categories EB, SI, AB and C form a psychosocial dimension (Psychosocial). All 12 categories are included in an overall SIP score.

Emotional distress, the Hospital Anxiety and Depression scale (HAD). A previously tested Swedish version of the HAD [15, 16] was used to assess emotional status. The HAD scale is specifically developed for the detection of anxiety and depression in patients with somatic conditions. It consists of 14 items and gives separate scores for anxiety and depression. Higher scores indicate more emotional distress.

Emotional wellbeing, the Mood Adjective Check List (MACL). The MACL measures emotional wellbeing [17]. We used a shortened 38 item version covering three basic dimensions of mood: pleasantness/unpleasantness; activation/deactivation; and calmness/tension. The scores from all items form an overall MACL score. In this scale, higher scores indicate a more positive emotional state.

Global health rating. At the follow-up 12 months later the patients indicated their perceived health on a seven-point categorical scale (excellent health = 0, very bad health = 6).

Analysis methods

To be considered reliable, the measurement error of the Swedish SGRQ should be small. We assessed the internal consistency reliability by calculating the Cronbach's α coefficient. According to conventional rules, it should at least exceed 0.70. To be considered valid, the Swedish SGRQ should measure QoL in COPD. To document this, we performed correlation analyses between data from the SGRQ measure and established QoL instruments and clinical and physiological variables. The anticipated results included: 1) correlation in the expected direction with the SIP, KPS and global health rating, and with clinical and physiological data (concurrent validity); 2) stronger correlations between variables measuring related phenomena, for instance SGRQ activity and 6MWD or SIP Physical score (convergent validity), and weaker between variables not directly related, for instance SGRQ activity and MACL or SIP Psychosocial score (discriminant validity); 3) a similar pattern of correlations in the original and Swedish versions (measurement equivalence); and 4) an invariant pattern of correlations in the baseline and follow-up data sets (stability). Both parametric and nonparametric (Spearman) correlations were calculated. As no major discrepancies were found, Pearson's correlation coefficients are reported throughout.

Scaling properties of the Swedish SGRQ were tested with the revised Multitrait Analysis Program [18]. This gives items and scales descriptive statistics, scales internal consistency estimates, item scale correlations and scale-to-scale correlations based on raw data. Item internal consistency is conventionally supposed to be satisfactory if the correlation between the item and its scale is at least 0.40. Item discriminant validity is the correlation between the item and the other scales. It should ideally be lower than the correlation between the item and its own scale (scaling success). In this test, the 13 patients in the risk group with mild COPD detected in a spirometry screening were included.

Results

Patients

Baseline characteristics and QoL data are given in table 1. Fifty four per cent of the patients were former smokers and all others were current smokers. Sixty three per cent were males. Blood gases were measured in 55 patients. Mean (SD) PO_2 was 9.6 (1.2) kPa. Only eight patients had a PO_2 <8 kPa and only five of these received home oxygen. Data on days in hospital were skewed. Forty eight patients had none and the remainder had 1–90 days in hospital during the preceding year. The mean SGRQ total score was 46% of the possible maximum score, range 6.4–90.6. The mean of the SIP Overall score was 8.5, range 0–41.9. The risk group comprised 13 patients. Their mean (SD) age was 64.7 (6.8) yrs and FEV1 and vital capacity (VC) were: 75.5 (5.3) and 73.3 (12.7) % pred, respectively.

Table 1. – Baseline characteristics and quality of life scores of the study group (n=68)

Characteristic	Mean (SD)	Quality of life scale	Mean (SD)
Age yrs	64.6 (6.8)	SGRQ Symptoms	54.5 (24.9)
FEV1 % pred	39.9 (17.0)	SGRQ Activity	60.7 (20.9)
FEV1 reversibility %	12.6 (14.1)	SGRQ Impacts	34.5 (19.8)
VC pred	66.6 (18.9)	SGRQ Total	46.0 (18.3)
$T_{L,CO}$ % pred	33.9 (22.5)	SIP Physical	7.8 (8.9)
6MWD m	260.7 (95.7)	SIP Psychosocial	6.8 (9.1)
KPS	71.8 (11.1)	SIP Overall	8.5 (8.1)
Days in hospital*	6.1 (16.3)	MACL Total	3.0 (0.62)
Disease duration yrs	7.6 (4.7)	HAD Anxiety	5.2 (4.5)
Pack-years n	37.3 (16.1)	HAD Depression	4.4 (3.8)

FEV1: forced expiratory volume in one second; % pred: percentage of predicted value; VC: vital capacity. $T_{L,CO}$: transfer factor of the lung for carbon monoxide; 6MWD: 6 min walking distance; KPS: Karnofsky Performance Status; SGRQ: St George's Respiratory Questionnaire; SIP: Sickness Impact Profile; MACL: Mood Adjective Check List; HAD: Hospital Anxiety and Depression scale; *: days in hospital during the last 12 months.

Internal consistency reliability and other scale properties

Table 2 shows results from the multitrait analysis of the Swedish SGRQ. The Cronbach's α coefficient was 0.81 for the Symptoms section, 0.88 for the Activity section, 0.88 for the Impacts section and 0.91 for the whole SGRQ. Table 2 also shows that the items of each SGRQ component correlated most with other items of the same scale (scaling success). On the other hand, this table shows that all items in the three SGRQ components did not contribute to the scales (item-internal consistency) and there was an overlap between the component scales (item-discriminant validity).

Correlation pattern

Table 3 shows the pattern of correlations from the present study. All correlations were significant in the expected direction. The SGRQ Total score correlations were strong for the SIP Overall score (r=0.69), walking distance (r=-0.61) and HAD Depression score (r=0.59). Modest associations were noted for the pulmonary physiological variables (range -0.37 to -0.44). No correlation was found between PO_2 and the SGRQ Total score nor the SIP Overall score (data not shown). Age and gender did not correlate with the SGRQ Total or component scores with one exception. A weak correlation was seen between female gender and the SGRQ Activity score, r=0.25 ($p<0.05$).

Table 2. – Results from tests of scaling properties of the Swedish version of the St George's Respiratory Questionnaire (n=81)

	St George's Respiratory Questionnaire			
	Symptoms	Activity	Impacts	Total
Items n	8	16	26	50
Scale levels n	29	16	31	76
Incomplete scale scores % +	0	0	0	0
Theoretical range	0–100	0–100	0–100	0–100
Observed range	6–100	6–93	0–86	5–91
At ceiling %	1.2	0	0	0
At floor %	0	0	8.6	0
Item-internal consistency#	0.55 (0.40–0.67)	0.55 (0.34–0.68)	0.44 (0.17–0.75)	-
Item-discriminant validity§	0.40 (0.22–0.68)	0.42 (0.02–0.64)	0.34 (-0.06–0.70)	-
Scaling success % ‡	81	88	85	-
Cronbach's alpha coefficient	0.81	0.88	0.88	0.91

+: half-scale criterion (at least 50% of items within a scale filled in); #: mean (range) correlation between items and hypothesized scale corrected for overlap; §: mean (range) correlation between items and other scales. ‡: per cent correlations that are higher with hypothesized scale.

Comparison with the original SGRQ

In table 3, corresponding values are presented from the original report by JONES and co-workers [5, 6]. A striking similarity could be seen between the two sets of correlations except for the HAD Anxiety dimension and, to some extent, FEV_1.

Convergent and discriminant validity

Table 4 is a multitrait-multimethod correlation matrix and displays the correlations between the SGRQ component scores *versus* the SIP dimensional scores, the SIP independent category scores, mood scores and physiological and clinical variables. The SGRQ activity scores covaried the most with the other physical functioning measures: SIP Physical dimension, walking distance and SIP Home Management (r=0.61, -0.63, 0.59 respectively). As expected, the correlations were lower with data measuring other aspects of QoL. The SGRQ Impacts score correlated substantially with the SIP Physical score (r=0.65) and the SIP Overall score (data not shown, r=0.69, p<0.0001). With only few exceptions (walking distance, SIP Sleep/Rest and HAD Depression), the SGRQ Activity and Impact components showed the same patterns of associations with other measures. The SGRQ Symptoms score, however, was more specific and, consequently, a different pattern was seen, with only weak correlations with other variables. The SGRQ and the generic SIP showed almost equal correlations with physiological and clinical variables in patients with COPD (data not shown).

Table 3. – Relationships between SGRQ Total scores and physiological data, quality of life scores and global health rating. Correlation coefficients from the Swedish and UK data sets

	Correlation coefficient	
	Swedish SGRQ	JONES and co-workers [5, 6]
FEV_1 % pred	-0.42***	-0.30†
VC % pred	-0.44***	-0.42
$T_{L,CO}$ % pred	-0.37**	NA
6MWD m	-0.61***	-0.61
HAD Anxiety	0.36**	0.58
HAD Depression	0.59***	0.59
SIP Physical	0.64***	0.69
SIP Psychosocial	0.56***	0.62
SIP Overall	0.69***	0.71
Health rating††	0.50***	0.63

†: personal communication; ††: seven-point health rating scale and five-point health rating scale respectively. For all self-assessments: a high score represents worse health/quality of life. **: p<0.01; ***: p<0.001. For definitions, see legend to table 1.

Changes over one year

For the 52 patients that took part in the follow-up, the mean differences in QoL scores, and in clinical and physiological variables, between the two occasions were nonsignificant or small (table 5). There were correlations between the SGRQ Total score difference and the corresponding differences of the SIP Overall, Physical and Psychosocial scores: r=0.49 (p=0.0002), 0.43 (p=0.0015) and 0.39 (p= 0.0045) respectively. No association was found between the change of score of the SGRQ and the other measured variables (data not shown).

Comparison with data from the follow-up

Table 6 compares the cross-sectional correlation series for the SGRQ total scores from the two data sets 12 months apart. The pattern of correlations was similar on both occasions.

Table 4. – Multitrait-multimethod correlation matrix. Coefficients of correlation between the Total and component scores of the SGRQ and the SIP dimension and category scores and scores representing emotional distress (HAD), wellbeing (MACL), clinical variables and physiological data

	Correlation coefficient			
SGRQ components	Symptoms	Activity	Impacts	Total
FEV_1 % pred	NS	-0.49***	-0.4***	-0.42***
VC % pred	NS	-0.42***	-0.44***	-0.44***
$T_{L,CO}$ % pred	NS	-0.47***	-0.32*	-0.37**
6MWD m	-0.37**	-0.63***	-0.53***	-0.61***
Days in hospital†	NS	0.21 NS	0.27*	0.25*
KPS	-0.42***	-0.56***	-0.57***	-0.62***
SIP Physical	0.25*	0.61***	0.65***	0.64***
SIP Eating	NS	0.46***	0.50***	0.49***
SIP Sleep/Rest	0.37**	0.48***	0.63***	0.62***
SIP Home Management	0.27*	0.59***	0.56***	0.59***
SIP Work	NS	NS	NS	NS
SIP Recreation/Pastimes	0.26*	0.44***	0.43***	0.46***
SIP Psychosocial	0.30*	0.50***	0.55***	0.56***
HAD Anxiety	0.31**	0.27*	0.33**	0.36**
HAD Depression	0.38**	0.49***	0.58***	0.58***
MACL Total	-0.39***	-0.51***	-0.48***	-0.54***

†: days in hospital during the last 12 months. For KPS: a high score indicates a better performance status. *: p<0.05; **: p<0.01; ***: p<0.001; NS: nonsignificant. For further definitions, see legend to table 1.

Table 5. – Differences (Δ) of means between the second and first measurement points (n=52)

	ΔMeans	Δ_{SD}	p-value
SGRQ Symptoms	-1.7	23.3	NS
SGRQ Activity	1.3	14.0	NS
SGRQ Impacts	2.4	10.5	NS
SGRQ Total	1.4	9.6	NS
SIP Physical	0.7	6.1	NS
SIP Psychosocial	1.9	6.0	<0.05
SIP Total	1.4	5.1	NS
MACL Total	0.0	0.4	NS
FEV_1 % pred	-2.7	7.8	<0.05
VC % pred	-4.4	13.1	<0.05
6MWD m	-4.3	56.5	NS
KPS	-0.8	7.2	NS
Days in hospital†	0.9	12.2	NS

†: Registered during the last 12 months. NS: nonsignificant (t-tests for paired samples). For further definitions, see legend to table 1.

Table 6. – Relationships between SGRQ Total scores and clinical variables, physiological data and quality of life scores. Coefficients of correlation from the baseline data set and the 12 month follow-up

	Correlation coefficients	
	Baseline	Follow-up
FEV_1 % pred	-0.42***	-0.49***
VC % pred	-0.44***	-0.42**
6MWD	-0.61***	-0.53***
HAD Anxiety	0.36**	0.49***
HAD Depression	0.59***	0.63***
MACL Total	-0.54***	-0.55***
SIP Physical	0.64***	0.64***
SIP Psychosocial	0.56***	0.62***
SIP Overall	0.69***	0.74***
KPS	-0.62***	-0.61***
Days in hospital†	0.25*	0.34*

For the SGRQ, HAD and SIP; a high score indicates worse health/quality of life. For MACL and KPS: a high score indicates a better wellbeing/performance status. *: p<0.05; **: p<-0.01; ***: p<0.001. For definitions, see legend to table 1.

Discussion

Significant and strong correlations were found between the Swedish SGRQ and a widely used, generic health status questionnaire, the SIP. Significant, but more modest, correlations were found with measures of emotional disturbance (HAD) and wellbeing (MACL). The same holds true for spirometric data. The SGRQ also correlated with the patients' own global health rating and with the assessment of performance status by a clinician (KPS).

Thus, our study confirms the validity of the Swedish adaptation of the SGRQ. The absence of correlation between PO_2 and the SGRQ Total score is probably due to the small proportion of patients with a low PO_2. In a population of patients where almost all had a low PO_2 such a correlation was found [19].

The good agreement between correlations found in this study and those presented previously from the original English studies [5, 6] confirms the measurement equivalence between the two versions of the SGRQ. Only small discrepancies were found. In the report of Jones *et al.* [5], the correlation with the HAD Anxiety scale was somewhat stronger. The results are, however, difficult to compare as their study also included asthmatics. The small difference in correlations with FEV_1 probably depends on the stratification in our study, with rather strong representation of patients with both very low and high FEV_1.

Our results demonstrated a satisfactory internal consistency reliability (alpha levels above 0.80) of the Swedish SGRQ, although there seem to be many items in the Activity and Impacts components not contributing to the reliability of the measure for patients with COPD. We conclude that the Swedish SGRQ is reliable and valid and compares well with data from the English source version.

There was also good agreement between the correlation patterns from the baseline study and the follow-up 12 months later. This indicates expected stability of the Swedish SGRQ as the physical condition of the study sample was not altered. One important aspect of the validity of a questionnaire, its sensitivity to change, was thus not examined in our study, *i.e.* differences between the variables from the baseline study and the follow-up were nonsignificant or small, as were the difference in FEV_1 % pred. A small change in spirometric values does not influence QoL to a large extent [4]. Two more important predictors of QoL, walking distance and depression score, did not change significantly.

Finally, the differences in scoring between the Sickness Impact Profile and the St George's Respiratory Questionnaire, with comparatively low values of the Sickness Impact Profile and much higher values in the St George's Respiratory Questionnaire (several patients had a Sickness Impact Profile overall score of zero but the lowest St George's Respiratory Questionnaire total score was 6.4), do not influence the correlation pattern. The St George's Respiratory Questionnaire and the generic Sickness Impact Profile showed almost equal correlations with physiological and clinical variables in patients with chronic obstructive pulmonary disease.

Acknowledgements: The authors gratefully acknowledge the assistance of A-M. Hilmersson, clinical research nurse, in monitoring the study.

References

1. Curtis JR, Deyo RA, Hudson LD. Health-related quality of life among patients with chronic obstructive pulmonary disease. *Thorax* 1994; 49: 162–170.
2. Fishman AP. Pulmonary rehabilitation research. *Am J Respir Crit Care Med* 1994; 149: 825–833.
3. MacKenzie CR, Charlson ME, DiGioia D, Kelley K. Can the Sickness Impact Profile measure change? An example of scale assessment. *J Chron Dis* 1986; 39: 429–438.
4. Jones PW, Baveystock CM, Littlejohns P. Relationships between general health measured with the Sickness Impact Profile and respiratory symptoms, physiological measures and mood in patients with chronic airflow limitation. *Am Rev Respir Dis* 1989; 140: 1538–1543.
5. Jones PW, Quirk FH, Baveystock CM, Littlejohns P. A self-complete measure of health status for chronic airflow limitation. *Am Rev Respir Dis* 1992; 145: 1321–1327.
6. Jones PW, Quirk FH, Baveystock CM. The St George's Respiratory Questionnaire. *Respir Med* 1991; 85: (Suppl. B), 25–31.
7. Ferrer M, Alonso J, Prieto L, *et al*. Validity and reliability

of the St George's Respiratory Questionnaire after adaptation to a different language and culture: the Spanish example. *Eur Respir J* 1996; 9: 1160–1166.

8. Engström CP, Persson LO, Larsson S, Rydén A, Sullivan M. Functional status and well being in chronic obstructive pulmonary disease with regard to clinical parameters and smoking, a descriptive and comparative study. *Thorax* 1996; 51: 825–830.
9. Berglund E, Birath G, Bjure J, *et al.* Spirometric studies in normal subjects. I. Forced expirograms in subjects between 7 to 70 years of age. *Acta Med Scand* 1963; 173: 185–191.
10. Salorinne Y. Single-breath pulmonary diffusing capacity. Reference values and application in connective tissue diseases and in various lung diseases. *Scand J Respir Dis* 1976; Suppl. no 96: 12s–16s.
11. Butland RJA, Pang J, Gross ER, Woodcock AA, Geddes DM. Two-, six-, and 12-minute walking test in respiratory disease. *BMJ* 1982; 284: 1607–1608.
12. Karnofsky D, Burchenal JH. Clinical evaluation of chemotherapeutic agents in cancer. *In*: McLeod CM, ed. Evaluation of Chemotherapeutic Agents. New York, Columbia University Press, 1949; pp. 191–205.
13. Bergner M, Bobbit RA, Carter WB, Gilson BS. The Sickness Impact Profile: development and final revision of a health status measure. *Med Care* 1981; 19: 787–805.
14. Sullivan M. The Sickness Impact Profile (SIP): an instrument for overall health assessment, a basic evaluation. *J Drughter Res* 1988; 13: 167–169.
15. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983; 286: 171–173.
16. Sullivan M, Karlsson J, Sjöström L, *et al.* Swedish obese subjects (SOS) - an intervention study of obesity. Baseline evaluation of health and psychological functioning in the first 1743 subjects examined. *Int J Obes* 1993; 17: 503–512.
17. Sjöberg L, Svensson E, Persson LO. The measurement of mood. *Scand J Psychol* 1979; 20: 1–18.
18. Hays RD, Hayashi T. Beyond internal consistency: rationale and users' guide for Multitrait Analysis Program on the microcomputer. *Beh Res Meth Instrum Comp* 1990; 22: 167–175.
19. Okubadejo AA, Jones PW, Wedzicha JA. Quality of life in patients with chronic obstructive pulmonary disease and severe hypoxaemia. *Thorax* 1996; 51: 44–47.