Comparison of the standard gamble, rating scale, AQLQ and SF-36 for measuring quality of life in asthma

E.F. Juniper*, G.R. Norman*, F.M. Cox#, J.N. Roberts¶

Comparison of the standard gamble, rating scale, AQLQ and SF-36 for measuring quality of life in asthma. E.F. Juniper, G.R. Norman, F.M. Cox, J.N. Roberts. ©ERS Journals Ltd 2001.
ABSTRACT: With interest in health economics growing, there is a demand for valid methods for measuring health-related quality of life (HRQL) in asthma using utilities. The aims of this study were to develop disease-specific versions of the standard gamble and rating scale, to compare their measurement properties with those of the Asthma Quality of Life Questionnaire (AQLQ) and the Medical Outcomes Survey Short-Form 36 (SF-36), as well as to determine their validity for assessing asthma-specific quality of life.

Forty adults with symptomatic asthma participated in a 9-week observational study. Participants completed the standard gamble, rating scale, AQLQ, SF-36 and other measures of clinical asthma status at baseline and after 1, 5 and 9 weeks.

In patients whose asthma was stable between assessments, reliability was good for the rating scale (intraclass correlation coefficient (ICC)=0.89) and the AQLQ (ICC=0.95) but more modest for the SF-36 mental score (ICC=0.68), SF-36 physical score (ICC= 0.65) and standard gamble (ICC=0.59). The responsiveness index was highest in the AQLQ (1.35), followed by the rating scale (0.74), the physical score of the SF-36 (0.61) and the standard gamble (0.31). Construct validity (correlation with other indices of health status) was strongest for the AQLQ and the rating scale.

In conclusion, both the disease-specific rating scale and the Asthma Quality of Life Questionnaire have strong measurement properties for measuring asthma-specific quality of life; the Short-Form 36 health survey physical summary score has more modest properties. Although the disease-specific standard gamble has acceptable discriminative properties, its evaluative properties are too inadequate for it to be used in cost/utility analyses. Poor correlation between the standard gamble and the rating scale indicates that utilities cannot be derived from rating scale data.
Eur Respir J 2001; 18: 38–44.

*Dept of Clinical Epidemiology and Biostatistics, McMaster University Faculty of Health Sciences, Hamilton, Ontario, Canada, #GlaxoWellcome Inc., Research Triangle Park, North Carolina, USA, ¶Dept of Epidemiology and Biostatistics, University of Western Ontario, London, Ontario, Canada.

Correspondence: E. Juniper
20 Marcuse Fields
Bosham, West Sussex
PO18 8NA
UK
Fax: 44 1243573680

Health-related quality of life (HRQL) has been defined as "the functional effect of an illness and its consequent therapy on a patient, as perceived by the patient" [1]. There is no "best" instrument for measuring HRQL in patients with asthma since each approach has its strengths and weaknesses. Generic health profiles, such as the Medical Outcome Survey Short-Form 36 (SF-36) [2], have been designed to measure the burden of illness experienced by patients with a wide range of medical conditions. However, they often lack sufficient depth to capture the small but important changes that may occur in a particular illness, either spontaneously or as the result of an intervention [3]. This lack of responsiveness in generic profiles has led to the development of disease-specific questionnaires, such as the Asthma Quality of Life Questionnaire (AQLQ) [4], which focus on the functional impairments associated with a single illness.

With the growing recognition of the importance of health economics, many investigators are interested in using HRQL instruments to estimate utilities [5]. This approach is founded in modern utility theory, a normative rational model of decision under uncertainty. The standard gamble is a utility and it measures the value that patients place on their own health state. The technique was developed by VON NEUMANN and MORGENSTERN [6] and adapted for clinical use by TORRANCE [7].

The rating scale (feeling thermometer) [8] measures patients' perception of their own health state but it is not a utility since it does not meet the utility theory requirement of "decision under uncertainty". Nevertheless, it is being used with increasing frequency in clinical studies because it is much easier than the standard gamble, both to administer and to understand. However, neither the original standard gamble nor the original rating scale have shown adequate measurement properties for confident use in adults with asthma [3].

In a recent study, the original generic versions of the standard gamble and the rating scale were modified to be disease-specific for children with asthma [9]. Although neither instrument had measurement properties as strong as the disease-specific Paediatric

Asthma Quality of Life Questionnaire, they were considerably better than might have been expected had their generic forms been used. In the present study, both the standard gamble and the rating scale have been modified to be specific for adults with asthma and their measurement properties compared with those of the AQLQ and the SF-36.

Recognizing the weakness of the measurement properties of the standard gamble in some medical conditions, TORRANCE [5] has suggested that there is a power curve relationship between standard gamble and rating scale responses so that utilities may be interpolated from rating scale data. In this study, the relationship between the standard gamble and the rating scale has been examined to determine whether utilities can be interpolated from the rating scale data in asthma.

Methods

Subjects

Forty adults (18–65 yrs) with symptomatic asthma were enrolled from patients who had participated in previous studies or who responded to notices in the local media. Patients were excluded if they had evidence of fixed airway obstruction (forced expiratory volume in one second (FEV_1) postbronchodilator) <60% predicted)) or other illness that might have an impact on HRQL. All participants communicated in English and were able to make reliable measurements of peak expiratory flow (PEF). They all signed a consent that had been approved by the McMaster University Faculty of Health Sciences' Ethics Committee.

Study design

In this 9-week, observational study, patients were assessed at baseline and after 1, 5 and 9 weeks. At each visit, a trained interviewer administered the standard gamble and the rating scale and patients completed the self-administered version of the AQLQ and the SF-36. In addition, patients completed the Asthma Control Questionnaire (ACQ) [10] and pre-bronchodilator spirometry was measured. The order in which patients completed the tests was kept consistent at each visit. For 1 week before each follow-up visit, patients recorded morning PEF and daily β_2-agonist use in a diary.

Patients whose asthma was adequately controlled, continued on their established asthma medications throughout the study. Patients whose asthma was not adequately controlled at week 1 and/or week 5, were advised to increase their medication in the manner recommended by their asthma physician.

Outcome measures

Rating scale. The rating scale [8] (feeling thermometer) looks like a thermometer with clearly defined end-points: 0=least preferred health state (death) and 100=most preferred health state (perfect health). In this study, patients first read aloud a description of three hypothetical asthma health states which were derived from the AQLQ. The order of presentation was randomized (Appendix 1).

Patients were asked to place three markers on the thermometer to reflect their feelings about each state. Patients then read the following:

"Think about how your asthma has bothered you during the last two weeks. Try to remember whether your asthma made you wheeze or feel short of breath and whether it limited you in your activities such as work, sports, social activities or hobbies. Think about whether you were bothered by such things as cigarette smoke, dust or pollution. Remember whether your asthma woke you up at night. Consider whether asthma has made you feel frustrated or afraid."

Patients were asked to place a marker on the thermometer to reflect their feelings about their asthma during the previous 2 weeks. This value was considered to represent the patient's preference rating for their own health state.

Standard gamble. The standard gamble [8] is administered to patients using a decision board designed by TORRANCE 7. Initially, patients are asked to think about a particular health state and then to consider whether they would prefer to remain in that health state for the next 10 yrs or take a chance with a new (imaginary) treatment. They are told that the new treatment has the ability to return them to perfect health immediately with no side-effects, but also has the ability to cause instant death. Initially, the probability of returning to perfect health if they took the new treatment is set at 100% with absolutely no chance of death. Usually, all those who understand the concept choose to take the new treatment rather than stay in the present health state. The probability of returning to perfect health on taking the new treatment is then gradually reduced (and the chance of death increased) until the patient decides to remain in the current health state rather than take a chance on the new treatment. The indifference point represents the value that the patient places on that health state. In this study, patients first completed the standard gamble with each of the three hypothetical asthma states used for the rating scale (Appendix 1) and then it was administered according to how their own asthma had been during the previous 2 weeks.

Asthma Quality of Life Questionnaire. The AQLQ [4] is a 32-item disease-specific questionnaire that has been designed to measure the functional impairments that are most troublesome to adults with asthma [4]. Patients are asked to recall their experiences during the previous 2 weeks and to score each item on a 7-point scale. The overall AQLQ score is the mean response to all 32 questions. Four independent studies have established that the AQLQ has strong measurement properties and validity [3, 11–13].

Medical Outcomes Survey Short-Form 36. This 36-item generic health status questionnaire [2] provides scores in nine functional domains which may be combined

into two summary scores (physical and mental health) using an algorithm provided by the developer. In asthma, it has been shown to have acceptable internal consistency and cross-sectional validity [14].

Statistical analysis

Comparison of the measurement properties (reliability, responsiveness and both cross-sectional and longitudinal construct validity) [15] of the four instruments required defining a group of patients whose asthma remained stable between consecutive clinic visits (weeks 1–5 and 5–9) and a group of patients whose asthma changed. Categorization was done using ACQ data in which a change in score of <0.5 on the 7-point scale can be considered clinically important [10]. Patients who either improved or deteriorated by a score $\geqslant 0.5$ were considered "changed" and all other patients were considered "stable".

The reliability of the instruments was determined from patients in the "stable" group. If a patient contributed more than one data point to this category, a single point was selected blindly using a random number generator. Test-retest reliability has been estimated as the within-subject variance and related to the total variance as an intraclass correlation coefficient (ICC). This statistic provides evidence of the instrument's ability to discriminate between patients of different levels of impairment.

The responsiveness of the instruments was determined in three ways. First, for patients in the "changed" group, it was determined whether the instruments could detect change using a paired t-test. Secondly, it was determined whether the instrument could detect differences between patients in the "stable" and "changed" groups using an unpaired t-test. Thirdly, the responsiveness index ($\Delta/sd\Delta$, where Δ=change in score between visits) was calculated [16]. The difference in responsiveness indices between instruments was tested using a paired t-test. Some patients experienced a change in their asthma during both study periods and therefore, contributed two observations to the "changed" group. To ensure that this did not result in an overestimation of the precision of responsiveness, the variance was inflated to take into account within-subject correlations by the quantity $1+(n-1)\rho$ where ρ is the ICC of the change scores and n=2 (the number of observations per subject).

The authors considered that the validity of the AQLQ for measuring asthma-specific quality of life has already been established [3, 11–13] and that if the SF-36 rating scale and standard gamble are capable of measuring an aspect of asthma-specific HRQL, they should correlate moderately-to-strongly with the AQLQ. In addition, the correlations between the four instruments and other measures of asthma clinical status were examined and compared to expected values derived from previous studies.

Results

Thirty-nine patients completed the study; one patient failed to return for the final visit. Demographic and health status data at enrolment are shown in table 1.

Thirty-five patients remained stable between two consecutive clinic visits and thus provided data for the reliability analysis. The highest degree of reliability was observed with the AQLQ (ICC=0.95), followed by the rating scale (ICC=0.89), the SF-36 mental score (ICC=0.68), SF-36 physical score (ICC=0.65) and standard gamble (ICC=0.59) (table 2).

Twenty-six patients contributed 35 observations to the "changed" group and 35 patients contributed 44 observations to the "stable" group. All instruments except the standard gamble and the mental summary score of the SF-36, were able to detect change in the "changed" group (table 3) and a similar pattern was seen for the ability of instruments to differentiate between the "stable" and "changed" groups. The highest responsiveness index was seen in the AQLQ (1.35). Statistical differences between responsiveness indices are shown in table 4.

Correlations between the trial instruments ranged from weak to strong (table 5) (fig. 1). The strongest association was between the AQLQ and the rating scale (r=0.67). Correlations, both cross-sectional and longitudinal, between the trial instruments and other measures of health status show that those of the AQLQ tended to be closest to the expected values (tables 6 and 7).

Table 1. – Demographics and baseline Scores

Subjects n	40
Age	38±11
Sex M:F	12:28
Asthma Medications: previous 30 days	
No medications	1
Short-acting bronchodilators (bd) alone	7
Inhaled steroids (is)+bd	29
Long-acting bronchodilator+is+bd	3
FEV1 prebronchodilator % pred	77.8±20.2
Asthma Quality of Life Questionnaire	5.43±0.83
SF-36 Physical	47.3±7.8
SF-36 Mental	48.6±10.8
Rating Scale	72.6±19.4
Standard Gamble	0.87±0.13

Data are presented as n or mean±SD. M: male; F: female; FEV1: forced expiratory volume in one second; SF-36: Medical Outcomes Survey Short-Form 36.

Table 2. – Reliability

Instrument	Within subject SD	Between subject SD	ICC
Rating Scale	6.04	17.59	0.89
Standard Gamble	0.11	0.13	0.59
AQLQ	0.18	0.8	0.95
SF-36 Physical	4.76	6.5	0.65
SF-36 Mental	6.07	8.76	0.68

ICC: intraclass correlation coefficient; AQLQ: Asthma Quality of Life Questionnaire; SF-36: Medical Outcome Survey Short-Form 36; SD: standard deviation. n=35.

Table 3. – Responsiveness: change in score between consecutive visits

Instrument	Stable group[+]	Changed group[#]	Difference p-value	Responsiveness index
Rating Scale	0.80±13.86	13.25±17.26***	0.0009	0.74
Standard Gamble	0.03±0.16	0.05±0.13	0.64	0.31
AQLQ	0.08±0.38	0.89±0.66***	<0.0001	1.35
SF-36 Physical	-0.66±6.69	5.02±6.49***	0.002	0.61
SF-36 Mental	1.38±8.67	-0.42±8.83	0.45	-0.06

Data are presented as mean±SD. AQLQ: Asthma Quality of Life Questionnaire; SF-36: Medical Outcome Survey Short-Form 36. [+]: n=44; [#]: n=35; ***: p<0.001.

Table 4. – Statistical comparison of responsiveness indices

Instrument (responsiveness index)	AQLQ (1.35)	Rating Scale (0.74)	SF-36 Physical (0.61)	Standard Gamble (0.31)
Rating Scale (0.74)	p<0.05			
SF-36 Physical (0.61)	p<0.05	NS		
Standard Gamble (0.31)	p<0.001	p<0.05	NS	
SF-36 Mental (-0.06)	p<0.001	p<0.01	p<0.05	NS

AQLQ: Asthma Quality of Life Questionnaire; SF-36: Medical Outcome Survey Short-Form 36; NS: not significant.

Twenty-three of the 40 patients valued their life at the highest utility (0.95) on the standard gamble but of these, 16 scored 45–100 on the rating scale (fig. 1). This was the same range covered by the 17 patients who scored less than the maximum on the standard gamble. With such a poor distribution, it was inappropriate to attempt to fit a power relationship.

Discussion

The original rating scale (feeling thermometer), which contains generic health marker states and asks patients to rate their overall health, has only modest measurement properties in adults with asthma [3]. In this study, it was found that when using asthma-specific marker states and rephrasing the patient's own health state in relation to asthma, the measurement properties appear to be stronger. In addition, the strong correlation between the AQLQ and the rating scale, and similar correlations between these instruments and other measures of clinical status, suggest that the rating scale is measuring the same construct (concept) as the AQLQ. Although the rating scale is not a utility and cannot in itself be used for cost/utility analyses or for calculating Quality Adjusted Life Years [17], this study raises the question as to whether

Table 5. – Pearson correlation coefficients between instruments at the end of week 1

Instrument	SF-36 Physical	SF-36 Mental	Rating Scale	Standard Gamble
AQLQ	0.53	0.48	0.67	0.49
SF-36 Physical			0.36	0.19
SF-36 Mental			0.52	0.38
Rating Scale				0.6

AQLQ: Asthma Quality of Life Questionnaire; SF-36: Medical Outcome Survey Short-Form 36. n=40.

it can be used as an alternative to the AQLQ for measuring asthma specific quality of life. The rating scale is very quick and easy to administer, though probably not as quick as the new Mini AQLQ [18]. However, the AQLQ: 1) tends to have stronger evaluative and discriminative measurement properties; 2) is self-administered, whereas the rating scale requires a trained administrator; 3) allows the different components of quality of life impairment (symptoms, emotional function, activity limitation and problems with environmental stimuli) to be examined and evaluated independently; 4) allows easy identification of an individual patient's problems; and 5) allows the minimal important difference to be established [19].

Similarly, the results of this study suggest that the disease-specific version of the standard gamble has slightly stronger measurement properties than those of the generic version [3]. Although the discriminative properties (reliability and cross-section validity) for
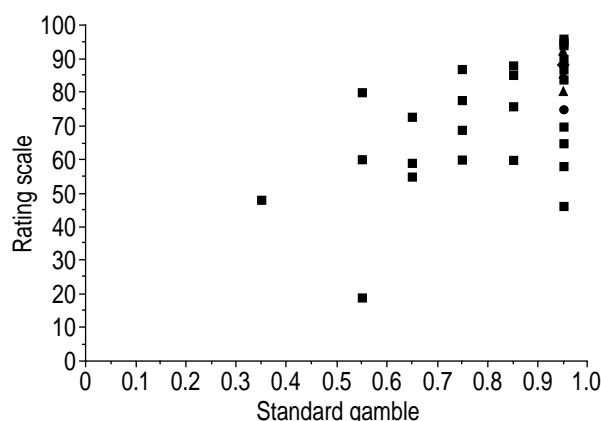

Fig. 1. – Association between rating scale and standard gamble scores at the end of the first week. Number of patients scoring 0.95 (maximum) on the standard gamble was 23 of 40. ■: 1 observation; ▲: 2 observations; ●: 3 observations; ◆: 4 observations.

Table 6. – Cross-sectional validity: Pearson correlation coefficients at the end of week 1

Instrument	Rating Scale	Standard Gamble	AQLQ	SF-36 Physical	SF-36 Mental
ACQ[#]	0.54	0.46	0.64	0.34	0.31
FEV1[¶]	0.05	0.11	0.18	0.09	0.06
PEF[+]	0.27	0.29	0.2	0.18	0.12
β_2-agonist use	0.07	0.13	0.03	0.08	0.01

ACQ: Asthma Control Questionnaire; FEV1: forced expiratory volume in one second. PEF: peak expiratory flow; AQLQ: Asthma Quality of Life Questionnaire; SF-36: Medical Outcome Survey Short-Form 36. n=40. Expected correlations based on previous observations [3, 11–13]. [#]: $r>0.6$; [+]: $r>0.2$; [¶]: $r<0.2$. No data available for β_2-agonist use. All r-values $\geqslant 0.33$ are statistically significant ($p<0.05$).

the disease-specific standard gamble can only be considered modest, they do indicate that the instrument is able to detect differences between patients. However, the evaluative properties (responsiveness and longitudinal validity) are weak and indicate that the standard gamble is insensitive to within-patient change and that it should be used circumspectly in clinical trials and cost/utility analyses. This is a serious difficulty because the standard gamble is the only true "patient-based" utility. Recognizing the weak measurement properties of the standard gamble in some medical conditions, TORRANCE [5] has suggested that there is a power curve relationship between the standard gamble and the rating scale and that one can interpolate utilities from rating scale data. In this study, more than half the patients scored the maximum on the standard gamble (i.e. no impairment), even though they tended to represent patients at the more severe end of the asthma spectrum (32 of the 40 patients required regular inhaled steroids). This very poor distribution of the data prohibits power curve fitting and indicates that utilities should not be interpolated from rating scale data in asthma (fig. 1).

Concerning the validity of the standard gamble as an HRQL instrument, TORRANCE [5] argues that "the standard gamble measurement technique is valid by definition". This follows from the argument that "the standard gamble technique by VON NEUMANN and MORGENSTERN [6] is a widely used and well-respected method for the measurement of utilities and is taken as the "well-accepted measure" (criterion)" [7] (i.e. it is the gold standard). Data from this study challenge this argument. It would be expected that an instrument that purports to be the gold standard for measuring HRQL would correlate moderately-to-strongly with other valid measures of HRQL and health status. Since the standard gamble showed only modest correlations with the AQLQ, the SF-36 and clinical asthma status and was unresponsive to change in

clinical status, the assumption of the standard gamble′s criterion validity in asthma is questionable.

To minimize the burden to the patients at each clinic visit and to prevent the completion of similar instruments influencing patients′ responses, neither the time trade-off [8] nor any measure of societal utility were included. The time trade-off is a simpler tool than the standard gamble but it is not a utility and its measurement properties tend to be between those of the standard gamble and the rating scale [7]. An alternate method of measuring utilities is with a generic health status questionnaire to which weighting has been given according to the value that society places on various health states. Such questionnaires include the Health Utilities Index [20], the Quality of Wellbeing Scale [21] and the EuroQol [22]. It was considered that the SF-36 would give a good estimate of the measurement properties of generic questionnaires in general (the societal weights do not alter the measurement properties of the basic generic questionnaire) and thus address the usefulness of these utility questionnaires in asthma.

The measurement properties of the AQLQ are very similar to those reported in a number of other validation studies [3, 11–13] and confirm that the AQLQ has strong evaluative and discriminative properties and can be used with confidence to measure asthma-specific quality of life in clinical trials, clinical practice and epidemiological surveys.

The generic SF-36 health profile is used extensively in a wide range of medical conditions and has been shown to be an excellent instrument for measuring burden of illness and for detecting problems not usually associated with a particular disease. In asthma, it has been shown to have good internal consistency and acceptable cross-sectional validity [14]. In this study, the physical summary score of the SF-36 showed only modest reliability and responsiveness but this is not surprising. Categorizing patients

Table 7. – Longitudinal construct validity: Pearson correlation coefficients

Instrument	ΔRating Scale	ΔStandard Gamble	ΔAQLQ	ΔSF-36 Physical	ΔSF-36 Mental
ΔACQ[#]	0.7	0.35	0.84	0.42	0.04
ΔFEV1[¶]	0.33	0.27	0.23	0.06	0.12
ΔPEF[+]	0.46	0.18	0.45	0.36	0.18
Δβ_2-agonist use[+]	0.45	0.12	0.52	0.4	0.13

ACQ: Asthma Control Questionnaire; FEV1: forced expiratory volume in one second; PEF: peak expiratory flow; AQLQ: Asthma Quality of Life Questionnaire; SF-36: Medical Outcome Survey Short-Form 36. Expected correlations based on previous observations [3, 11, 23]. [#]: $r>0.6$; [+]: $r=0.4–0.6$; [¶]: $r=0.2–0.4$. All r-values $\geqslant 0.23$ are statistically significant ($p<0.05$).

into "stable" and "changed" groups was based on changes in the ACQ. Patients who reported asthma stability between clinic visits may have experienced changes in other medical conditions (*e.g.* sprained ankle, common cold) which would have been detected by the SF-36 and thus lowered the estimate of reliability. Similarly, for responsiveness, the SF-36 would have picked up changes in conditions unrelated to asthma. In addition, the SF-36 probably does not have the depth of focus to detect small but clinically important changes in asthma-specific quality of life. The mental summary score of the SF-36 did not perform well. However, this too is not surprising because the questions that contribute to this score focus predominantly on mood, such as "down in the dumps", "downhearted and blue" and "a happy person" and adults with asthma tend have more problems with anxieties and fears than they do with depression [4].

One of the goals of validation is to determine whether instruments measure the concept (construct) of interest. In this study, the authors could have chosen to evaluate whether the instruments are capable of measuring generic HRQL (*i.e.* the impact of overall health status on patients' quality of life) or whether they are capable of measuring disease-specific HRQL (*i.e.* the problems specifically associated with asthma). It was decided *a priori* to evaluate the ability of the instruments to measure asthma-specific quality of life. Both cross-sectional and longitudinal correlations suggest that the AQLQ and the rating scale are valid for this task. The SF-36 would not be expected to measure asthma-specific quality of life as well as these two instruments. Nevertheless, the correlations for the physical summary score are sufficiently strong to suggest that this component of the SF-36 has modest but acceptable construct validity.

In conclusion, the results of this study show that both the disease-specific rating scale and the Asthma Quality of Life Questionnaire have strong discriminative and evaluative properties for measuring asthma-specific quality of life. The Medical Outcome Survey Short-Form 36 physical summary score has modest measurement properties and validity. Although the disease-specific standard gamble has modest discriminative properties, its evaluative properties are too inadequate for it to be used in cost/utility analyses. The weakness of the relationship between the standard gamble and the rating scale indicates that utilities should not be interpolated from rating scale data in asthma.

Appendix 1: marker states

"You have asthma but it hardly bothers you at all. Occasionally, you wheeze and become a little short of breath. Sometimes cigarette smoke, dust and air pollution can be quite troublesome. Your asthma only limits you in strenuous activities. You have no problems at work because of your asthma. Your asthma does not bother you at night. Your asthma is never frightening."

"You have asthma and it bothers you quite a bit. You wheeze quite often and get short of breath. Your asthma limits you in many things you like doing, such as sports and hobbies. Quite often your asthma interferes with your work and social life. Cigarette smoke, dust and air pollution bother you. Your asthma sometimes wakes you up at night. You get frustrated by your asthma and sometimes an attack is a little frightening."

"You have bad asthma and it is really troublesome. You feel wheezy and it is difficult to breathe a lot of the time. You cannot take part in any sports and your social life is very restricted. You are very bothered by cigarette smoke, dust and air pollution. At night it is difficult to sleep because of your asthma. Quite often you get bad asthma attacks and these are very frightening."

References

1. Schipper H, Clinch J, Powell V. Definitions and conceptual issues. *In*: Spilker B, ed. Quality of Life and Pharmacoeconomics in Clinical Trials. Philadelphia, Lippincott-Raven Publishers, 1996; 11–23.
2. Stewart AL, Hays R, Ware JE. The MOS short-form general health survey. Reliability and validity in a patient population. *Med Care* 1988; 26: 724–732.
3. Rutten-van Molken MPMH, Clusters F, Van Doorslaer EKA, *et al.* Comparison of performance of four instruments in evaluating the effects of salmeterol on asthma quality of life. *Eur Respir J* 1995; 8: 888–898.
4. Juniper EF, Guyatt GH, Epstein RS, Ferrie PJ, Jaeschke R, Hiller TK. Evaluation of impairment of health-related quality of life in asthma: development of a questionnaire for use in clinical trials. *Thorax* 1992; 47: 76–83.
5. Torrance GW. Utility approach to measuring health-related quality of life. *J Chron Dis* 1987; 40: 593–600.
6. von Neumann J, Morgenstern O. Theory of games and economic behavior. New York, Wiley, 1953.
7. Torrance GW. Social preferences for health states: an empirical evaluation of three measurement techniques. *Socio-Econ Plan Sci* 1976; 10: 129–136.
8. Torrance GW. Measurement of health state utilities for economic appraisal. *J Health Econom* 1986; 5: 1–30.
9. Juniper EF, Guyatt GH, Feeny DH, Griffith LE, Ferrie PJ. Minimum skills required by children to complete health-related quality of life instruments: comparison of instruments for measuring asthma-specific quality of life. *Eur Respir J* 1997; 10: 2285–2294.
10. Juniper EF, O'Byrne PM, Guyatt GH, Ferrie PJ, King DR. Development and validation of a questionnaire to measure asthma control. *Eur Respir J* 1999; 14: 902–907.
11. Juniper EF, Guyatt GH, Ferrie PJ, Griffith LE. Measuring quality of life in asthma. *Am Rev Respir Dis* 1993; 147: 832–838.
12. Rowe BH, Oxman AD. Performance of an asthma quality of life questionnaire in an outpatient setting. *Am Rev Respir Dis* 1993; 148: 675–681.
13. Leidy NK, Coughlin C. Psychometric performance of the Asthma Quality of Life Questionnaire in a US sample. *Qual Life Res* 1998; 7: 127–134.

14. Bousquet J, Knani J, Dhivert H, *et al.* Quality of life in asthma. 1. Internal consistency and validity of the SF-36 questionnaire. *Am J Respir Crit Care Med* 1994; 149: 371–375.

15. Guyatt GH, Kirshner B, Jaeschke R. Measuring health status: what are the necessary measurement properties? *J Clin Epidemiol* 1992; 45: 1341–1345.

16. Guyatt GH, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chron Dis* 1987; 40: 171–178.

17. Torrance GW, Feeny D. Utilities and quality-adjusted life years. *Intl J Tech Ass Health Care* 1989; 5: 559–575.

18. Juniper EF, Guyatt GH, Cox FM, Ferrie PJ, King DR. Development and validation of the Mini Asthma Quality of Life Questionnaire. *Eur Respir J* 1999; 14: 32–38.

19. Juniper EF, Guyatt GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific quality of life instrument. *J Clin Epidemiol* 1994; 47: 81–87.

20. Feeny D, Furlong W, Barr RD, Torrance GW, Rosenbaum P, Weitzman S. A comprehensive multi-attribute system for classifying the health status of survivors of childhood cancer. *J Clin Oncol* 1992; 10: 923–928.

21. Kaplan RM, Anderson JP, Wu AW, Matthews WC, Kozin F, Orenstein D. The Quality of Well-being Scale: application in AIDS, cystic fibrosis and arthritis. *Med Care* 1989; 27: S27–S43.

22. The EuroQol Group. A new facility for the measurement of health-related quality of life. *Health Policy* 1990; 16: 199–208.