



The responsiveness and validity of the CAMPHOR Utility Index

D.M. Meads*, S.P. McKenna*,[#], N. Doughty[†], C. Das⁺, W. Gin-Sing[§],
J. Langley[‡] and J. Pepke-Zaba[†]

ABSTRACT: The aim of the present study was to validate and determine the minimal important difference (MID) and responsiveness of the Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR) Utility Index, a new tool enabling cost utility analyses.

CAMPHOR, 6-min walking test (6MWT) and New York Heart Association (NYHA) data for 869 pulmonary hypertension patients (545 (63%) female; mean \pm SD age 56.6 ± 15.4 yrs) from three centres were analysed. Utility was correlated with 6MWT data and calculated by NYHA class to assess validity. Effect sizes were calculated for those with two CAMPHOR assessments. Distribution and anchor-based MIDs were calculated. Analyses were carried out in patients receiving bosentan in order to determine whether or not those remaining in NYHA class III following treatment improved.

The Utility Index distinguished between adjacent NYHA classes and correlated with 6MWT results. CAMPHOR subscales and utility were as responsive as the 6MWT (effect sizes ranged 0.31–0.69 for the CAMPHOR and 0.16–0.34 for the 6MWT). The within-group MID for the Utility Index was estimated to be ~ 0.09 . Patients remaining in NYHA class III experienced, on average, a significant improvement (CAMPHOR Utility Index and functioning), which exceeded the MID.

The CAMPHOR Utility Index is valid and responsive to change. Patients can experience significant and important improvements even if they do not improve on the basis of traditional outcomes, such as NYHA functional class.

KEYWORDS: Bosentan, Cambridge Pulmonary Hypertension Outcome Review, pulmonary hypertension, quality of life, responsiveness, utility

Pulmonary hypertension (PH) is a rare disease, affecting 2–5 per million population annually [1]. It is characterised by elevated pulmonary arterial pressure and pulmonary vascular resistance, which ultimately result in right heart failure and death [1, 2]. PH most commonly arises as a result of underlying cardiopulmonary disease, but may also be a consequence of pulmonary thromboembolic disease or disease of the pulmonary microcirculation [3].

Symptoms include dyspnoea, fatigue, palpitations, peripheral oedema, chest pain and syncope [2]. Available treatments include intravenous epoprostenol, subcutaneous and intravenous treprostinil and inhaled iloprost, and oral therapies, such as endothelin receptor antagonists (bosentan, sitaxsentan and ambrisentan) and the phosphodiesterase type 5 inhibitor sildenafil [4, 5]. However, the currently available treatments (with the exception of pulmonary endarterectomy for thromboembolic PH) do not cure the disease [6]. The current aim of therapy is to reduce pulmonary arterial pressure, improve

right heart function, improve exercise capacity and, ultimately, to lengthen survival time and improve quality of life (QoL).

Given the high cost of PH treatments (for example, epoprostenol costs £130–390 (GBP sterling) daily in the UK, and bosentan and sitaxentan each cost £55 daily in the UK [7]), there is a need to establish that the treatments are cost-effective. This necessitates a cost–utility analysis in which the cost of treatment is related to the benefit gained in terms of a parameter that expresses the quantity of life and QoL, the quality-adjusted life year (QALY). The QALY requires information relating to patients' preferences expressed in terms of utility, which is generally expressed as a value between 1 (representing perfect health) and 0 (death). To date, utility in PH populations, as in most other diseases, has been derived by asking patients to complete generic questionnaires, such as the European quality of life five-dimension (EQ-5D) questionnaire [8], which provide a utility score. Evidence suggests that disease-specific utility

AFFILIATIONS

*Galen Research, Manchester,
[#]School of Psychology, University of Central Lancashire, Preston,
[†]Pulmonary Vascular Disease Unit, Papworth Hospital NHS Trust, Papworth Everard,
[‡]Dept of Cardiology, Royal Free Hospital,
[§]Dept of Cardiology, Hammersmith Hospital, and
[¶]Actelion Pharmaceuticals UK, London, UK.

CORRESPONDENCE

D.M. Meads
Galen Research
Enterprise House
Manchester Science Park
Manchester
M15 6SE
UK
Fax: 44 1612264478
E-mail: david_meads@hotmail.com

Received:

May 06 2008

Accepted after revision:

August 08 2008

STATEMENT OF INTEREST

Statements of interest for all authors of this study and for the study itself can be found at www.erj.ersjournals.com/misc/statements.shtml

and health status measures may be more responsive to change in patients' health than generic measures [9, 10]. Given this fact, a disease-specific utility measure, the Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR) Utility Index [11], was recently developed in order to permit cost-utility analyses in PH. The derivation of this index involved conducting a valuation study in which a combination of six questions from the CAMPHOR QoL scale were presented to the general population in the form of health state scenarios. Patients' preferences for each scenario were gathered through a valuation exercise, which permitted utility values to be ascribed to all possible combinations of responses to these CAMPHOR items on a scale of 1 (perfect health) to 0 (death). As a result, a utility score for patients can be derived from their responses to the six relevant items of the CAMPHOR QoL scale. As well as being an outcome measure in its own right, the Utility Index also permits derivation of the QALY. This metric permits comparisons to be made across diseases, and aids clinicians, researchers and regulatory bodies in making decisions about healthcare resource allocation while also factoring in patients' views and preferences.

Although there is evidence for the reliability and validity of the CAMPHOR Utility Index [11], the responsiveness and minimal important difference (MID) of the scale have not been established for PH. Responsiveness is the ability of a scale to detect small but important changes over time [12]. The MID has been described as "the smallest difference in the outcome of interest that informed patients or informed proxies perceive as important, either beneficial or harmful, and which would lead the patient or clinician to consider a change in the management" [13]. The MID is valuable since it provides a means (beyond statistical significance, which is influenced by sample size [14]) of interpreting the relevance of changes in the patients' status over time.

The aim of the present study was to further validate the CAMPHOR Utility Index, determine its MID and establish the responsiveness of the utility measure in a group of PH patients.

METHODS

Several specialist PH centres in the UK collect patients' CAMPHOR responses routinely, together with exercise capacity and functional class data. Data from Papworth Hospital in Cambridge and the Royal Free and Hammersmith Hospitals in London were analysed.

Sample

The study sample consisted of patients attending the three participating centres in the UK either at first referral or for periodic assessment, exacerbations or surgery during the period 2004–2006. The characteristics of those completing at least one CAMPHOR are included in table 1. Few patients had PH associated with left heart diseases, reflecting the specialities of those centres supplying data.

Assessments

As well as CAMPHOR responses, 6-min walking test (6MWT) and New York Heart Association (NYHA) class [15] data were available. Owing to the nature of the data collection process, it was not possible to collect all information from each centre or

for each patient. Patients completed the CAMPHOR and 6MWT during the same visit only at Papworth Hospital. Consequently, only 6MWT data from Papworth Hospital are included in the analyses.

Patients at Papworth Hospital were also asked, at each follow-up visit, whether their QoL had changed since their previous visit on a seven-point scale (ranging from "very much worse" to "very much better"). This patient global rating was used as an anchor in the MID calculation.

New York Heart Association class

The NYHA functional class system places patients into one of four categories (I–IV) by taking into account their physical limitations and the symptoms brought on by physical activity. Physical limitations and activity-induced symptoms increase as the classes progress from I to IV (see the European Society of Cardiology guidelines for a full description of the classification system [16]).

6-min walking test

The 6MWT is an exercise test employed as a clinical indicator of patients' functional capacity. It is a practical test of how far the patient can walk unaided and at their own pace in 6 min. The patient is permitted to slow down, stop and rest when they want. Only standardised phrases of encouragement are used by the nurse or clinician administering the test. No exercise equipment is required, but a 30-m hallway, along which patients walk back and forth, is required to administer the test [17].

Cambridge Pulmonary Hypertension Outcome Review

The CAMPHOR [18] is a disease-specific suite of patient-reported outcome measures for use in PH. It comprises separate symptom (25 items), activity limitation (15 items) and QoL (25 items) scales. Scores range 0–25 for the symptom and QoL scales and 0–30 for the activity limitation scale. Higher scores indicate greater symptom experience, worse QoL and greater functional limitation, respectively. The CAMPHOR scales have been shown to be reliable (internal consistency $\alpha=0.90$ – 0.92 ; test-retest correlation = 0.86 – 0.92) and valid [18].

The Utility Index [11] consists of six CAMPHOR QoL items and permits derivation of PH-specific utility scores.

Analyses

Spearman's correlation analysis determined the level of association between the CAMPHOR scales and utility and between utility and 6MWT results.

CAMPHOR utility descriptive statistics were calculated for the whole group and by functional class and diagnostic group. Differences between groups were tested using unpaired t-tests (for CAMPHOR utility) and Mann-Whitney U-tests (for CAMPHOR scales).

CAMPHOR responsiveness was evaluated by examining change in patients' scores after treatment initiation. Only patients who had completed the CAMPHOR ≤ 2 months before and ≤ 1 month after starting treatment (time 1), completed the CAMPHOR twice within a period of 21–365 days (time 2) and received ≥ 21 days of treatment between CAMPHOR completions were included in the analyses.

TABLE 1 Sample characteristics

Subjects n	869
Males	320 (36.8)
Females	545 (62.7)
Missing	4 (0.5)
Age yrs	56.6±15.4 (17.0–90.0)
Time since diagnosis yrs	1.9±1.8 (0.0–8.0)
Diagnosis	
Idiopathic PAH	160 (18.4)
Chronic thromboembolic PH	314 (36.1)
PH associated with left heart diseases	59 (6.8)
PAH associated with connective tissue disease	200 (23.0)
Other [#]	136 (15.7)
NYHA class[†]	
I	35 (4.5)
II	201 (26.0)
III	472 (61.1)
IV	64 (8.3)
Primary treatment[‡]	
Bosentan	351 (58.6)
Sildenafil	61 (10.2)
Pulmonary endarterectomy	102 (17.0)
Subcutaneous treprostinil	6 (1.0)
Epoprostenol/intravenous iloprost	34 (5.7)
Inhaled iloprost	39 (6.5)
Other	6 (1.0)
More than one treatment	86 (9.9)

Data are presented as mean ± SD (range) or n (%), unless otherwise indicated.

PAH: pulmonary arterial hypertension; PH: pulmonary hypertension; NYHA: New York Heart Association. [#]: including unclear diagnosis, PH related to HIV, sarcoidosis and respiratory disease; [†]: n=772; [‡]: n=599.

Initially, all diagnoses, functional classes and treatment types were included. Cohen's effect size [19], the standardised response mean and responsiveness statistic [20] were calculated for changes over time in order to assess responsiveness. They were calculated by dividing the mean change in scale scores over time by the baseline SD (for effect size), SD of change in scores (for standardised response mean) and SD of change in scores for a stable patient, patients whose NYHA class did not change in the present case, group (for the responsiveness statistic). Effect sizes were interpreted in the following way: <0.2: minimal or no change; ≥0.2–<0.5: small change; ≥0.5–<0.8: moderate change; and ≥0.8: large change [21]. Paired t-tests for utility and 6MWT result and Wilcoxon signed-rank tests for CAMPHOR scales were employed to assess the significance of change over time.

In order to examine the relative responsiveness of CAMPHOR and NYHA class data, further analyses were conducted in order to determine the extent to which improvements in health and functional status might occur in those whose functional class did not improve. Analyses were conducted on patients who remained in NYHA class III, the class supplying the largest sample, following treatment. All idiopathic pulmonary arterial hypertension (PAH; IPAH) patients and those with PH associated with connective tissue disease (CTD) and congenital heart disease in functional class III were entered into the

analysis if they met the requirements of the responsiveness analysis above, had not changed functional class between CAMPHOR completions and were being treated with bosentan.

This last criterion was included since patients with these types of PH were initially prescribed bosentan following diagnosis at the centres included in the present study. In addition, the number of patients on the other treatments was too small to permit separate analyses.

The mean ± SD time between completing two CAMPHORs in this subgroup was 85.1 ± 51.3 days (21–203 days).

The MID of the Utility Index was determined by calculating the mean change in score between the two assessments for patients reporting feeling “a little better” on the QoL global rating of change item (which represents the anchor) and by distributional statistics (scores required to achieve certain effect sizes and the SEM). The SEM has been proposed as a surrogate for the MID [22], and, taking into account its reliability, is a measure of the precision of a scale. Although there are problems with these types of analysis (particularly that of anchoring questionnaire scores to a global rating of change [23, 24]), this approach to the determination of the MID is still regarded as the most appropriate [25, 26]. The anchor-based and distributional values are “triangulated” in order to arrive at the MID threshold value [25]. This involves taking into account the values from multiple approaches and making a judgement regarding what represents a reasonable convergence value. If changes in group scores over time reach the MID, then it can be claimed that the group in question has experienced a noticeable and important improvement, one that is beyond the random variation in scores obtained using the questionnaire.

RESULTS

Scores by diagnosis

The unadjusted analyses suggested that there were significant differences in CAMPHOR scores between certain diagnoses (table 2). However, ANCOVAs controlling for age and sex found no significant differences, with sex revealed as the most important factor in each comparison. Separate ANCOVAs by sex controlling for age found no differences between PH types on any scale for males. Female patients with PH associated with CTD scored significantly higher on the symptom scale than patients with either IPAH or chronic thromboembolic PH. Female CTD patients exhibited worse scores than other PH patients for the remaining measures, but these differences were nonsignificant. These analyses clearly showed that females achieved markedly worse scores than males for all outcome assessments and all diagnoses (with the exception of IPAH). After controlling for age, female patients with CTD and chronic thromboembolic PH scored significantly worse for all outcomes (including 6MWT result) except for the CTD group for utility (p=0.52).

Validity of the Utility Index

Table 3 shows moderate correlations between utility, CAMPHOR scores and 6MWT result. There were significant differences in scale and utility scores between each adjacent functional class (table 4).

TABLE 2 Cambridge Pulmonary Hypertension Outcome Review scores by pulmonary hypertension (PH) diagnosis

	Subjects n	Utility Index	Symptoms	Functioning	QoL
IPAH	158	0.54±0.28	11.8±5.9	11.4±6.7	10.9±6.5
CTEPH	308	0.56±0.29	11.3±6.7	11.2±6.9	10.3±7.0
PH associated with left heart diseases (CHD)	59	0.57±0.31	11.9±6.2	10.5±5.9	10.3±7.3
PAH associated with CTD	185	0.48±0.28	13.5±6.2	13.6±6.3	12.1±6.6
Comparison (unadjusted)		CTEPH>CTD**	PPH<CTD**; CTEPH<CTD**	PPH<CTD**; CHD<CTD**; CTEPH<CTD**	CTEPH<CTD**

Data are presented as mean ±SD. QoL: quality of life; IPAH: idiopathic pulmonary arterial hypertension; CTEPH: chronic thromboembolic pulmonary hypertension; CHD: congenital heart disease; CTD: connective tissue disease; PPH: primary pulmonary hypertension. **: p<0.01.

Responsiveness

Significantly smaller sample sizes were available for the responsiveness analyses since most patients had received treatment for a considerable period of time before completing their first CAMPHOR. Table 5 includes effect size statistics for changes in the Utility Index and CAMPHOR scale scores between time 1 (baseline) and time 2 (post-treatment). The effect sizes for the scale changes were small, except for the change in the symptom scale, which was moderate. With the exception of the CAMPHOR functioning and QoL scales, these changes were significant.

Minimal important difference of Utility Index

The SEM of CAMPHOR utility was 0.09, and 1.96 SEM (which reflects the 95% confidence interval) were 0.17. The mean change in utility for those reporting their QoL as “a little better” was 0.07, and 0.10 for those who reported being “moderately better” (table 6). The utility changes required to achieve effect sizes of 0.2, 0.5 (half an SD) and 0.8 were 0.05, 0.13 and 0.20, respectively. These various values suggest that a reasonable estimate of the within-group MID of the Utility Index would be 0.09.

Table 7 indicates that the group of patients remaining in NYHA functional class III who had been treated with bosentan experienced an improvement in mean CAMPHOR Utility Index and mean CAMPHOR activity limitation scores. The improvement in utility, which was significant, exceeded the proposed MID.

TABLE 3 Correlations between Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR) utility and CAMPHOR scores and 6-min walking test (6MWT) result

	CAMPHOR			6MWT
	Symptoms	Functioning	QoL	
CAMPHOR Utility Index	-0.69	-0.63	-0.88	0.49

QoL: quality of life.

DISCUSSION

The data analyses reported above were designed to provide additional evidence of the validity of the CAMPHOR Utility Index, establish its responsiveness and help interpret changes in utility scores. These analyses, involving a reasonably large sample considering the rarity of the disease, have shown how utility, symptoms, functioning and QoL scores relate to 6MWT performance and highlighted differences in these outcomes according to PH diagnoses and functional class.

The utility values by functional class obtained in the present sample (class I=0.89, class II=0.71, class III=0.46 and class IV=0.30) differ substantially from those derived by HIGHLAND *et al.* [27], who used an expert panel to derive hypothetical EQ-5D scores. The scores obtained using the EQ-5D in the study of MCKENNA *et al.* [11] were 0.69 for class II and 0.59 for class III, suggesting that the CAMPHOR Utility Index is better at discriminating between these classes. Utility scores by NYHA class also appear larger with the CAMPHOR than they do with the six-dimensional health state classification derived from the 36-item short-form health survey (SF-6D) [28]. This is supported by a recent study that found utility in PAH patients to be 0.73, 0.67, 0.60 and 0.52, respectively, in functional classes I–IV using the SF-6D [29].

Evidence of the Utility Index's validity was provided by its ability to distinguish between functional class groups and the

TABLE 4 Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR) scores by New York Heart Association functional class

Class	Subjects n	Utility Index	Symptoms	Functioning	QoL
I	35	0.89±0.17 [#]	1.5±2.5	1.9±2.7	2.1±3.9
II	196	0.71±0.26 [#]	8.5±5.3 [†]	7.6±4.9 [†]	7.2±5.8 [†]
III	458	0.46±0.26 [#]	14.0±5.5 [†]	13.6±5.9 [†]	12.7±6.1 [†]
IV	62	0.30±0.18	17.2±4.8 [†]	20.3±5.2 [†]	16.3±5.5 [†]

Data are presented as mean ±SD. QoL: quality of life. [#]: p<0.01 (higher) versus class below; [†]: p<0.01 (higher) versus class above.

TABLE 5 Effect sizes for within-group change over two visits following treatment

	CAMPHOR				6MWT
	Utility Index	Symptoms	Functioning	QoL	
Subjects n	55	56	55	55	25
Effect size	0.36	0.69	0.33	0.43	0.16
SRM	0.27	0.44	0.31	0.33	0.24
Respon stat	0.30	0.57	0.33	0.53	0.34

CAMPHOR: Cambridge Pulmonary Hypertension Outcome Review; 6MWT: 6-min walking test; QoL: quality of life; SRM: standardised response mean; respon stat: responsiveness statistic.

moderate correlations with 6MWT results. However, it remains necessary to examine how utility scores relate to clinical outcomes, such as assessments of haemodynamics. The derivation of the MID for the Utility Index should aid researchers in interpreting changes in utility scores and defining a responder. However, replication of these results is desirable given the small sample that completed the global rating questions.

Differences in symptom scores between PH diagnoses were consistent with previous research, indicating that CTD patients show more severe symptoms [30]. These differences may need to be accounted for in clinical studies including patients with different PH aetiologies, and it could be argued that data from CTD patients should be analysed separately. The present results suggest that there may also be a need to control for sex differences.

The CAMPHOR scales appeared to be at least as responsive as the 6MWT. However, given the uneven sample sizes used in the present comparison, these results should be interpreted with caution. The favourable responsiveness of CAMPHOR functioning may be explained by the fact that the 6MWT result represents only one aspect of functioning, whereas the CAMPHOR scale covers wider activities of daily living. Other researchers have found the 6MWT less responsive than patient-reported outcome measures [31]. The unresponsiveness

TABLE 6 Change in Utility Index associated with response to global rating of change item

Response	Subjects	Change
Very much worse	6	-0.09 ± 0.30
Moderately worse	8	-0.07 ± 0.25
A little worse	18	-0.04 ± 0.25
No change	29	-0.01 ± 0.19
A little better	21	0.07 ± 0.28
Moderately better	14	0.10 ± 0.27
Very much better	13	0.08 ± 0.25

Data are presented as n or mean ± SD.

of the NYHA classification has been reported in atrial fibrillation [32] and congestive heart failure [33] patients. This problem was confirmed in the present study by analysing changes in patients who remained in NYHA class III. The analyses indicated that mean Utility Index and CAMPHOR functioning improved with treatment to an extent that was significant and could be considered important by patients.

Of the 56 patients with NYHA class and QoL global rating question data at follow-up reporting a change in QoL, only 14% had changed functional class. In addition, an improvement by one NYHA class required a mean ± SD utility improvement of 0.20 ± 0.30 , more than twice the proposed MID. To illustrate the relative insensitivity of NYHA class, for an improvement of one class as the definition of a responder (a patient who has responded to treatment), the number needed to treat would be >10 in the present sample. This compares to a NNT of <3 if the Utility Index MID is used as the definition of treatment success.

Despite these limitations, NYHA class continues to be used as an outcome measure and to determine whether or not patients receive treatment. The present analyses suggest that determination of the outcome of clinical trials solely in terms of NYHA classification (and improvement in 6MWT result) is unsafe. Patients may not receive treatment when their disease severity suggests that they should, and researchers and regulatory bodies may erroneously conclude that no improvement in the patient's condition has occurred with treatment.

TABLE 7 Changes in scores[#] for those in functional class III at both visits following bosentan

	CAMPHOR				6MWT
	Utility Index	Symptoms	Functioning	QoL	
Subjects n	23	24	24	22	5
Mean ± SD	0.13 ± 0.27	2.08 ± 4.68	3.21 ± 5.50	1.32 ± 3.90	24.8 ± 34.82
Median	0	1	3.5	1	30
Range	-0.34–0.66	-6.0–13.0	-5.0–22.0	-8.0–9.0	60.0– -20.0
p-value	0.031 [*]	0.054 ⁺	0.005 ⁺	0.093 ⁺	0.225 ⁺

CAMPHOR: Cambridge Pulmonary Hypertension Outcome Review; 6MWT: 6-min walking test; QoL: quality of life. p-values present within-group change as follows. [#]: between time 1 and time 2; ^{*}: paired t-test; ⁺: Wilcoxon signed-rank test.

The present study has a number of limitations. Since the sample was a convenience sample, it could not be ensured that patients completed the CAMPHOR before starting treatment or at the same time as clinical assessments were performed. It was also not possible to ensure a standard period of time between visits. The study was not powered to determine true treatment effects in the responsiveness analyses, and no placebo control group was available. It was only possible to examine the effect of bosentan on patients remaining in NYHA class III given the small numbers receiving other treatments in this group. Finally, the sample sizes for some of the analyses were small despite the large initial number of patients completing the CAMPHOR.

The Cambridge Pulmonary Hypertension Outcome Review has previously been shown to be valid, reliable and responsive, and is recommended for use in pulmonary hypertension clinical studies alongside traditional clinical outcome measures. Since utility values can now be derived directly from responses to the Cambridge Pulmonary Hypertension Outcome Review, it is possible to conduct cost-utility analyses based on responses to this measure.

ACKNOWLEDGEMENTS

The authors would like to thank G. Coghlan (Royal Free Hospital, London, UK) and S. Gibbs (Hammersmith Hospital, London, UK) for providing additional data.

REFERENCES

- Rubin LJ. Primary pulmonary hypertension. *N Engl J Med* 1997; 336: 111–117.
- McLaughlin VV, McGoon MD. Pulmonary arterial hypertension. *Circulation* 2006; 114: 1417–1431.
- McGoon M, Gutterman D, Steen V, et al. Screening, early detection, and diagnosis of pulmonary arterial hypertension: ACCP evidence-based clinical practice guidelines. *Chest* 2004; 126: Suppl. 1, 14S–34S.
- Chin KM, Rubin LJ. Pulmonary arterial hypertension. *J Am Coll Cardiol* 2008; 51: 1527–1538.
- Badesch DB, Abman SH, Simonneau G, Rubin LJ, McLaughlin VV. Medical therapy for pulmonary arterial hypertension: updated ACCP evidence-based clinical practice guidelines. *Chest* 2007; 131: 1917–1928.
- Humbert M, Sitbon O, Simonneau G. Treatment of pulmonary arterial hypertension. *N Engl J Med* 2004; 351: 1425–1436.
- British Medical Association, Royal Pharmaceutical Society of Great Britain. British National Formulary 55. www.bnf.org Date last updated: March 2008. Date last accessed: 20 June 2008.
- EuroQoL Group. EuroQoL: a new facility for the measurement of health-related quality of life. *Health Policy* 1990; 16: 199–208.
- Terwee CB, Dekker FW, Wiersinga WM, Prummel MF, Bossuyt PM. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. *Qual Life Res* 2003; 12: 349–362.
- Guyatt GH, King DR, Feeny DH, Stubbing D, Goldstein RS. Generic and specific measurement of health-related quality of life in a clinical trial of respiratory rehabilitation. *J Clin Epidemiol* 1999; 52: 187–192.
- McKenna SP, Ratcliffe J, Meads DM, Brazier JE. Development and validation of a preference based measure derived from the Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR) for use in cost utility analyses. *Health Qual Life Outcomes* 2008; 21: 65.
- Chassany O, Sagnier P, Marquis P, Fullerton S, Aaronson N. Patient-reported outcomes: the example of health-related quality of life – a European guidance document for the improved integration of health-related quality of life assessment in the drug regulatory process. *Drug Inf J* 2002; 36: 209–238.
- Schunemann HJ, Puhan M, Goldstein R, Jaeschke R, Guyatt GH. Measurement properties and interpretability of the Chronic Respiratory Disease Questionnaire (CRQ). *COPD* 2005; 2: 81–89.
- Juniper EF. Quality of life questionnaires: does statistically significant=clinically important? *J Allergy Clin Immunol* 1998; 102: 16–17.
- Fisher JD. New York Heart Association Classification. *Arch Intern Med* 1972; 129: 836.
- Galiè N, Torbicki A, Barst R, et al. Guidelines on diagnosis and treatment of pulmonary arterial hypertension. The Task Force on Diagnosis and Treatment of Pulmonary Arterial Hypertension of the European Society of Cardiology. *Eur Heart J* 2004; 25: 2243–2278.
- ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med* 2002; 166: 111–117.
- McKenna SP, Doughty N, Meads DM, Doward LC, Pepke-Zaba J. The Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR): a measure of health-related quality of life and quality of life for patients with pulmonary hypertension. *Qual Life Res* 2006; 15: 103–115.
- Cohen J. *Statistical Power Analysis for the Behavioural Sciences*. New York, Academic Press, 1977.
- Guyatt GH, Bombardier C, Tugwell PX. Measuring disease-specific quality of life in clinical trials. *CMAJ* 1986; 134: 889–895.
- Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care* 1989; 27: Suppl. 3, S178–S189.
- Wyrwich KW. Minimal important difference thresholds and the standard error of measurement: is there a connection? *J Biopharm Stat* 2004; 14: 97–110.
- Norman GR, Stratford P, Regehr G. Methodological problems in the retrospective computation of responsiveness to change: the lesson of Cronbach. *J Clin Epidemiol* 1997; 50: 869–879.
- van Stel HF, Maillé AR, Colland VT, Everaerd W. Interpretation of change and longitudinal validity of the quality of life for respiratory illness questionnaire (QoLRIQ) in inpatient pulmonary rehabilitation. *Qual Life Res* 2003; 12: 133–145.
- Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008; 61: 102–109.

- 26** Crosby RD, Kolotkin RL, Williams GR. An integrated method to determine meaningful changes in health-related quality of life. *J Clin Epidemiol* 2004; 57: 1153–1160.
- 27** Highland KB, Strange C, Mazur J, Simpson KN. Treatment of pulmonary arterial hypertension: a preliminary decision analysis. *Chest* 2003; 124: 2087–2092.
- 28** Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002; 21: 271–292.
- 29** Keogh AM, McNeil KD, Wlodarczyk J, Gabbay E, Williams TJ. Quality of life in pulmonary arterial hypertension: improvement and maintenance with bosentan. *J Heart Lung Transplant* 2007; 26: 181–187.
- 30** Coghlan JG, Handler C. Connective tissue associated pulmonary arterial hypertension. *Lupus* 2006; 15: 138–142.
- 31** Demers C, McKelvie RS, Negassa A, Yusuf S, RESOLVD Pilot Study Investigators. Reliability, validity, and responsiveness of the six-minute walk test in patients with heart failure. *Am Heart J* 2001; 142: 698–703.
- 32** Ganiats TG, Browner DK, Dittrich HC. Comparison of Quality of Well-Being scale and NYHA functional status classification in patients with atrial fibrillation. New York Heart Association. *Am Heart J* 1998; 135: 819–824.
- 33** Eurich DT, Johnson JA, Reid KJ, Spertus JA. Assessing responsiveness of generic and specific health related quality of life measures in heart failure. *Health Qual Life Outcomes* 2006; 4: 89.