ERJ Express. Published on September 3, 2008 as doi: 10.1183/09031936.00069708

Title: The responsiveness and validity of the CAMPHOR Utility Index.

D.M. Meads[1], S.P. McKenna[1,2], Doughty N[3], Das C[4], W. Gin-Sing [5], Langley J[6], Pepke-Zaba J[3]

[1]Galen Research, Enterprise House, Manchester Science Park, Lloyd Street North Manchester, M15 6SE, UK.

[2]School of Psychology, University of Central Lancashire, Preston, PR1 2HE, UK

[3]Pulmonary Vascular Disease Unit. Papworth Hospital NHS Trust, Papworth Everard, Cambridgeshire, CB3 8RE, UK.

[4] Department of Cardiology, Royal Free Hospital, Pond Street, London, NW3 2QG

[5] Department of Cardiology, Hammersmith hospital, Du Cane Road, London, W12 OHS

[6]Actelion pharmaceuticals UK, BSI Building 13th Floor, 389 Chiswick High Road, London, W4 4AL

Corresponding author:

David M Meads

Galen Research

Enterprise House

Manchester Science Park

Manchester M15 6SE

UK

Email: david_meads@hotmail.com

Telephone: +44 (0)161 226 4446

Fax: +44 (0)161 226 4478

Running title: Responsiveness of the CAMPHOR Utility Index

Abstract

Objective: Validate, determine the Mininal Important Difference (MID) and responsiveness of the Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR) Utility Index, a new tool enabling cost utility analyses.

Patients and methods: CAMPHOR, 6 Minute Walk Test (6MWT) and NYHA data from three centres for 869 PH patients (545 [63%] female; Mean age = 56.6 [±15.4]) were analysed. Utility was correlated with 6MWT 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 run on patients receiving Bosentan to determine whether those remaining in NYHA class III after treatment improved.

Results: The Utility Index distinguished between adjacent NYHA classes and correlated .0.49 with 6MWT. CAMPHOR subscales and Utility were as responsive as 6MWT (effect sizes ranged 0.31-0.69 for CAMPHOR and 0.16-0.34 for 6MWT). Within-group MID for the Utility Index was estimated to be approximately 0.09. Patients remaining in NYHA class III on average experienced a statistically significant improvement (CAMPHOR Utility Index and functioning) that exceeded the MID.

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

Key words: Bosentan, CAMPHOR, Pulmonary Hypertension, Quality of life, Responsiveness, Utility

Introduction

Pulmonary hypertension (PH) is a rare disease affecting between 2 and 5 per million population a year. [1] It is characterised by elevated pulmonary artery pressure (PAP) and pulmonary vascular resistance, which ultimately results 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 dyspnea, 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 (PDE-5) inhibitor, Sildenafil [4, 5]. However, currently available treatments (with the exception of pulmonary endarterectomy for thromboembolic PH) do not cure the disease [6]. The present aim of therapy is to, lower 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 per day in the UK [7] while Bosentan and Sitaxentan each cost £55 per day 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 and quality of life - 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 EQ-5D [8] which provide a utility score. Evidence suggests that disease-specific utility and health status measures may be more responsive to change in patients' health than generic measures [9,10]. Given this, a disease-specific utility measure – the Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR) Utility Index [11] - was recently developed to allow cost-utility analyses in PH. The derivation of the 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 allowedutility values to be ascribed to all possible combinations of responses to these CAMPHOR items on a scale from 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 in the CAMPHOR QoL scale. As well as being an outcome measure in its own right, the Utility Index also allows the derivation of the QALY. This metric allows comparisons to be made across diseases and aids clinicians, researchers and regulatory bodies make decisions about healthcare resource allocation while factoring in patients' views and preferences.

While 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 as it provides a way (beyond statistical significance, which is influenced by sample size [14]) of interpreting the relevance of changes in the patients' status over time.

Objectives

To validate further the CAMPHOR Utility Index, to determine the minimal important difference (MID) and to establish the responsiveness of the utility measure in a group of PH patients.

Methodology

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 from 2004 to 2006 either at first referral, for periodic assessment, exacerbations or surgery. The characteristics of those completing at least one CAMPHOR are

included in Table 1. Few patients had PH associated with left heart diseases (CHD) reflecting the specialities of those centres supplying data.

Assessments

As well as CAMPHOR responses, six-minute walk test (6MWT) and New York Heart Association (NYHA [15]) class data were available. Due to the nature of the data collection process it was not possible to collect all information from each centre or for each patient. The patient completed the CAMPHOR and 6MWT during the same visit at Papworth Hospital only. 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.

NYHA 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].

6MWT

The 6MWT is an exercise test employed as a clinical indicator of patients' functional capacity. The 6MWT is a practical test of how far the patient can walk unaided and at

their own pace in six minutes. The patient is permitted to slow down, to stop and to 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 100-foot hallway is needed to administer the test, down which patients walk back and forth [17].

CAMPHOR

The CAMPHOR [18] is a disease-specific suite of patient–reported outcome (PRO) 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 α = 0.90-0.92; test-retest correlations = 0.86-0.92) and valid [18].

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

Analyses

Spearman correlations determined the level of association between the CAMPHOR scales and utility and between utility and the 6MWT.

CAMPHOR utility descriptive statistics were calculated for the whole group and by functional class and diagnostic groups. T-tests (for CAMPHOR utility) and Mann-

Whitney U tests (for CAMPHOR scales) were employed to test for differences between groups.

CAMPHOR responsiveness was examined by looking at change in patients' scores after treatment initiation. Only patients who had completed the CAMPHOR up to 2 months before and up to 1 month after starting treatment (Time 1), who had completed the CAMPHOR twice within a period of 21-365 days (Time 2), and had received ≥ 21 days of treatment between CAMPHOR completions were included in the analyses. Initially, all diagnoses, functional classes and treatment types were included. Cohen's effect sizes (ES) [19], the standardised response mean (SRM) and responsiveness statistic [20] were calculated for changes over time to assess responsiveness. Effect sizes were calculated by dividing the mean change in scale scores over time by; the baseline standard deviation (for ES); standard deviation of change scores (for SRM); and by the standard deviation of change scores for a group of stable patients - in this case patients whose NYHA class did not change (for the responsiveness statistic). Effect sizes were interpreted in the following way: <0.2 = minimal or no change; ≥0.2-<0.5 = a small change; ≥0.5-<0.8 = a moderate change and ≥0.8 = a large change [21]. Paired t-tests for utility and 6MWT and Wilcoxon Signed Ranks tests for CAMPHOR scales were employed to assess the significance of change over time.

To examine the relative responsiveness of CAMPHOR and NYHA class further analyses were conducted 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 PAH patients and those with PH associated with CTD and CHD in functional class III were entered into the analysis if they met the requirements of the responsiveness analysis above, if they had not changed functional class between CAMPHOR completions and if they were being treated with Bosentan.

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

The average time between completing two CAMPHORs in this subgroup was 85.1 days (SD = 51.3; range 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 Standard Error of Measurement (SEM)). The SEM has been proposed as a surrogate for the MID [22] and is a measure of the precision of a scale, taking into account its reliability. Although there are problems with these type of analyses (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 about 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 on the questionnaire.

Results

Scores by diagnosis

Unadjusted analyses suggested that there were significant differences in CAMPHOR scores between some diagnoses (Table 2). However, ANCOVAs controlling for age and gender found no significant differences, with gender revealed as the most important factor in each comparison. Separate ANCOVAs by gender 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 CTEPH. Female CTD patients had worse scores than other PH patients on the remaining measures but these differences were not significant. These analyses clearly showed that females achieved markedly worse scores than males on all outcome assessments in all diagnoses (with the exception of IPAH). After controlling for age, female patients with CTD and CTEPH scored significantly worse on all outcomes (including 6MWT) except for the CTD group in utility ($p=0.52$).

Validity of the Utility Index

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

Responsiveness

Significantly smaller sample sizes were available for these analyses as most patients had received treatment for a considerable 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). Effect sizes for the scale changes were small except for change in the symptom scale which was moderate. With the exception of the CAMPHOR functioning and QoL scales, these changes were statistically significant.

MID of Utility Index

The SEM for CAMPHOR Utility was 0.09 and the 1.96 SEM (which reflects the 95% confidence interval) value was 0.17. The mean change in utility for those reporting that their QoL as 'A little better' was .07 and .10 for those who reported being 'Moderately better' (Table 6). Utility changes required to achieve effect sizes of 0.2, 0.5 (which is half a standard deviation), and 0.8 were 0.05, 0.13, and 0.20, respectively. These different values suggest that a reasonable estimate of the within group MID for 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 actually experienced an improvement in mean CAMPHOR Utility Index and mean CAMPHOR activity limitation scores. The

improvement in utility, which was statistically significant, exceeded the proposed MID value.

Discussion

The data analyses reported above were designed to provide additional evidence of the validity of the CAMPHOR Utility Index, to establish its responsiveness and to help interpret change in the utility scores. These analyses – involving a reasonably large sample considering the rarity of the disease – have shown how utility, symptom, 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, class IV = 0.30) differ substantially from those derived by Highland and colleagues [27] who used an expert panel to derive hypothetical EQ-5D values. Values obtained with the EQ-5D in the McKenna et al study [11] were 0.69 for class II and 0.59 for class III suggesting 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 for the 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 for functional classes I to 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 moderate correlations with the 6MWT.

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 will help researchers interpret changes in utility scores and define 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 have 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. Results suggest that there may also be a need to control for gender differences.

The CAMPHOR scales appeared to be at least as responsive as the 6MWT. However, given the uneven sample sizes used for this comparison these results should be treated with caution. The favourable responsiveness of CAMPHOR functioning may be explained by the fact that the 6MWT 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 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. Analyses indicated that mean Utility Index and CAMPHOR functioning improved with treatment to an extent that was statistically significant and could be considered important by patients.

Of the 56 patients who had NYHA class and QoL global rating question data at follow-up reporting a change in QoL, only 14% had also changed functional class. In addition, an improvement by one NYHA class required a mean utility improvement of 0.20 (SD = .30) - over twice the proposed MID value. To illustrate the relative insensitivity of the NYHA; an improvement of 1 class as the definition of a responder (a patient who has responded to treatment), the number needed to treat (NNT) would be over 10 in this sample. This compares to a NNT of less than 3 if the Utility Index MID is used as a definition of treatment success.

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

The study has a number of limitations. As this 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 made. It was also not possible to ensure a standard period 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 CAMPHOR has previously been shown to be valid, reliable and responsive and is recommended for use in PH clinical studies alongside traditional clinical outcomes. As utility values can now be derived directly from responses to the CAMPHOR it will be possible to conduct cost-utility analyses based on responses to the measure.

Acknowledgements

The authors would like to thank Dr Gerry Coghlan (Royal Free Hospital, London) and Dr Simon Gibbs (Hammersmith Hospital, London) for providing additional data.

The authors would also like to thank Actelion Pharmaceuticals UK Ltd for supporting the research.

References

1. Rubin LJ. Primary pulmonary hypertension. N Engl J Med 1997; 336: 111-7.

2. McLaughlin VV, McGoon MD. Pulmonary arterial hypertension. Circulation. 2006 Sep 26;114(13):1417-31.

3. 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; 126:14S-34S.

4. Chin KM, Rubin LJ. Pulmonary arterial hypertension. J Am Coll Cardiol. 2008 Apr 22;51(16):1527-38.

5. 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 Jun;131(6):1917-28.

6. Humbert M, Sitbon O, Simonneau G. Treatment of pulmonary arterial hypertension. N Engl J Med. 2004 Sep 30;351(14):1425-36.

7. British Medical Association and the Royal Pharmaceutical Society of Great Britain. British National Formulary 55. Access date: 20 June 2008, URL: www.bnf.org

8. EuroQoL Group. EuroQoL: A new facility for the measurement of health-related quality of life. Health Policy 1990; 16: 199-208.

9. 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.

10. 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 Mar;52(3):187-92.

11. 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 and Quality of Life Outcomes (*in press*).

12. 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 Information Journal*, 2002; 36: 209-238.

13. Schunemann HJ, Puhan M, Goldstein R, Jaeschke R, Guyatt GH. Measurement properties and interpretability of the Chronic respiratory disease questionnaire (CRQ). COPD. 2005 Mar;2(1):81-9.

14. Juniper EF. Quality of life questionnaires: does statistically significant = clinically important? *J Allergy Clin Immunol*. 1998 Jul;102(1):16-7.

15. Fisher JD. New York Heart Association Classification. Arch Intern Med. 1972 May;129(5):836.

16. Galiè N, Torbicki A, Barst R, Dartevelle P, Haworth S, Higenbottam T, Olschewski H, Peacock A, Pietra G, Rubin LJ, Simonneau G, Priori SG, Garcia MA, Blanc JJ, Budaj A, Cowie M, Dean V, Deckers J, Burgos EF, Lekakis J, Lindahl B, Mazzotta G, McGregor K, Morais J, Oto A, Smiseth

OA, Barbera JA, Gibbs S, Hoeper M, Humbert M, Naeije R, Pepke-Zaba J; Task Force. 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 Dec;25(24):2243-78.

17. 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 Jul 1;166(1):111-7.

18. 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 Feb;15(1):103-15.

19. Cohen J. *Statistical Power Analysis for the Behavioural Sciences*, 1977; Academic Press, New York.

20. Guyatt GH, Bombadier C, Tugwell PX. Measuring disease-specific quality of life in clinical trials. *CMAJ*, 1986; 134; 889-95.

21. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. Med Care 1989; 27:S178.

22. Wyrwich KW. Minimal important difference thresholds and the standard error of measurement: is there a connection? J Biopharm Stat. 2004 Feb;14(1):97-110.

23. Norman GR, Stratford P, Regehr G. Methodological problems in the retrospective computation of responsiveness to change: the lesson of Cronbach. J Clin Epidemiol. 1997 Aug;50(8):869-79.

24. van Stel HF, Maille 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 Mar;12(2):133-45.

25. 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 Feb;61(2):102-9.

26. Crosby RD, Kolotkin RL, Williams GR. An integrated method to determine meaningful changes in health-related quality of life. J Clin Epidemiol. 2004 Nov;57(11):1153-60.

27. Highland KB, Strange C, Mazur J, Simpson KN. Treatment of pulmonary arterial hypertension: a preliminary decision analysis. Chest. 2003 Dec;124(6):2087-92.

28. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ. 2002 Mar;21(2):271-92.

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 Feb;26(2):181-7. The Journal of Heart and Lung Transplantation. 2007 Feb;26(2):181-187.

30. Coghlan JG., Handler C. Connective tissue associated pulmonary arterial hypertension. *Lupus.* 2006 vol 15, no 3, p.183-42.

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 Oct;142(4):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 May;135(5 Pt 1):819-24.

33. Dunselman PH, Kuntze CE, van Bruggen A, Beekhuis H, Piers B, Scaf AH, 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 Nov 24; 4:89.

Table 1: Sample characteristics (Total N=869)

Number (%) males	320 (36.8)
Number (%) females	545 (62.7)
Missing	4 (0.5)
Mean (SD) age; years	56.6 (15.4)
Range (years)	17.0 – 90.0
Time (SD) since diagnosis; years	1.9 (1.8)
Range (years)	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 [including unclear diagnosis, PH related to HIV, sarcoidosis, respiratory disease] (%)	136 (15.7)
NYHA class (N = 772)	
Number (%) I	35 (4.5)
Number (%) II	201 (26.0)
Number (%) III	472 (61.1)
Number (%) IV	64 (8.3)
Primary treatments (N = 599)	
Number (%) Bosentan	351 (58.6)
Number (%) Sildenafil	61 (10.2)
Number (%) Pulmonary endarterectomy	102 (17.0)
Number (%) Subcutaneous Treprostinil	6 (1.0)
Number (%) Epoprostinil / intravenous Iloprost	34 (5.7)
Number (%) Inhaled Iloprost	39 (6.5)

Number (%) Other	6 (1.0)
Number receiving more than one treatment	86 (9.9)

Table 2: CAMPHOR scores by PH diagnosis

Diagnosis	n	Utility Index	Symptoms	Functioning	QoL
Idiopathic PAH (IPAH)	158	0.54 (0.28)	11.8 (5.9)	11.4 (6.7)	10.9 (6.5)
Chronic thromboembolic PH (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 connective tissue disease (CTD)	185	0.48 (0.28)	13.5 (6.2)	13.6 (6.3)	12.1 (6.6)
Diagnosis comparison (Unadjusted)		CTEPH > CTD*	PPH<CTD* CTEPH<CTD*	PPH<CTD* CHD<CTD* CTEPH<CTD*	CTEPH<CTD*

*Statistically significant p <0.01

Table 3: Correlations between CAMPHOR utility, CAMPHOR scores and 6MWT

	CAMPHOR symptoms	CAMPHOR functioning	CAMPHOR QoL	6MW
CAMPHOR Utility Index	-.69	-.63	-.88	.49

Spearman correlations

Table 4: CAMPHOR mean (SD) scores by functional class

NYHA class	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)***

^ Significantly higher than class II ($p<0.01$); ^^Significantly higher than class III ($p<0.01$);

^^^ Significantly higher than class IV ($p<0.01$);

*Significantly higher than class I ($p<0.01$); **Significantly higher than class II ($p<0.01$);

***Significantly higher than class III ($p<0.01$)

Table 5: Effect sizes for within group change over two visits following treatment

	CAMHOR Utility Index	CAMPHOR symptoms	CAMPHOR functioning	CAMPHOR QoL	6MWT
N	55	56	55	55	25
Effect size	0.36	0.69	0.33	0.43	0.16
Standardised Response Mean (SRM)	0.27	0.44	0.31	0.33	0.24
Responsiveness statistic	0.30	0.57	0.33	0.53	0.34

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

	N	Mean change	SD
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

Table 7: Change scores for those in functional class III at both visits after receiving Bosentan

	Change scores Time 1 to Time 2				
	CAMPHOR Utility Index	CAMPHOR symptoms	CAMPHOR functioning	CAMPHOR QoL	6MWT
N	23	24	24	22	5
Mean	0.13	2.08	3.21	1.32	24.8
SD	0.27	4.68	5.50	3.90	34.82
Median	0.0	1.0	3.5	1.0	30.0
Minimum / Maximum	-0.34/0.66	-6.0 / 13.0	-5.0 / 22.0	-8.0 / 9.0	60.0 / -20.0
P value for test of within group change over time	.031*	.054†	.005†	.093†	.225†

*Paired t-test

†Wilcoxon Signed Ranks test