

Comparison of cardiac and pulmonary-specific quality of life measures in pulmonary arterial hypertension

Hubert Chen, MD, MPH¹; Teresa De Marco, MD²; Erin A. Kobashigawa, BA²; Patricia P. Katz, PhD³; Vivian W. Chang, BA²; Paul D. Blanc, MD, MSPH¹

Institutions:

¹ Division of Pulmonary Medicine, University of California San Francisco, San Francisco, USA

² Division of Cardiology, University of California San Francisco, San Francisco, USA

³ Institute for Health Policy Studies, University of California San Francisco, San Francisco, USA

Support: This study was supported by grant number K23 HL086585 from the National Heart, Lung and Blood Institute (H.C.) and the Flight Attendants Medical Research Institute (T.D.).

Keywords: pulmonary hypertension, quality of life

Corresponding Author:

Hubert Chen, MD, MPH

350 Parnassus Avenue, Suite 609

San Francisco, CA 94143-0924, USA

Tel: 415-502-6765

Fax: 415-476-6426

Email: hubert.chen@ucsf.edu

ABSTRACT

Studies systematically comparing the performance of health-related quality-of-life (HRQL) instruments in pulmonary arterial hypertension (PAH) are lacking. We sought to address this gap by comparing cardiac and respiratory-specific measures of HRQL in PAH.

We prospectively assessed HRQL in 128 patients with catheterization-confirmed PAH at baseline and at 6, 12, and 24+ months. Cardiac-specific HRQL was assessed using the Minnesota Living with Heart Failure Questionnaire (LHFQ); respiratory-specific HRQL using the Airways Questionnaire 20 (AQ20); and general health status using the 36-item Short Form physical component summary (SF-36 PCS).

The LHFQ and AQ20 were highly intercorrelated. Both demonstrated strong internal consistency and converged with the SF-36 PCS. Both discriminated among patients based on World Health Organization functional class (FC), 6-minute walk distance (6MWD), and Borg Dyspnea Index (BDI), except for a potential floor effect associated with low 6MWD. The LHFQ was more responsive than the AQ20 to changes over time in FC, 6MWD, and BDI. In multivariate analyses, the LHFQ and AQ20 were each longitudinal predictors of general health status, independent of FC, 6MWD and BDI.

In conclusion, both cardiac-specific and respiratory-specific measures appropriately assess HRQL in most patients with PAH. Overall, the LHFQ demonstrates stronger performance characteristics than the AQ20.

INTRODUCTION

Over the past 2 decades, clinical research in cardiopulmonary disease has broadened from a purely physiologic-based focus to a more comprehensive approach to health-related endpoints. In particular, there has been a rapid growth of interest in the development and application of patient-centered outcome measures.(1) The majority of such research has focused on common conditions, such as chronic obstructive pulmonary disease (COPD) or heart failure. The extent to which results from these studies can be extrapolated to more rare disease states is uncertain.

Pulmonary arterial hypertension (PAH) is a relatively uncommon disorder that manifests as both a pulmonary and a cardiac condition. Patients typically present with dyspnea on exertion that is indistinguishable from other more common pulmonary conditions. In fact, many patients are initially misdiagnosed with obstructive airway disease prior to establishing a diagnosis of PAH. Moreover, such patients also exhibit features of heart failure due to the decompensation of the right ventricle in the face of elevated pulmonary vascular resistance. Although traditionally considered to be a rapidly progressive and fatal disease, patients with PAH are now living longer due to the availability of effective medical therapies.(2-5) Consequently, the goals of PAH therapy have expanded from increasing survival to improving health-related quality of life (HRQL).

Despite the wealth of data on HRQL in COPD and heart failure, relatively few studies have addressed HRQL in PAH. Those studies that have been performed have utilized various types of

instruments, both generic and disease-specific.(6) Generic HRQL instruments can be advantageous in that they provide results that are comparable across a heterogeneous mix of conditions, but such approaches generally provide measures that are less sensitive to the nuances of the specific disease in question. In contrast, disease-specific instruments emphasize those aspects of health impacted by a particular type of disorder, for example, breathing-related symptoms and their potential negative impacts on a patient's sense of well being.

Since both cardiac-specific and pulmonary-specific measures focus heavily on dyspnea-related impairment, both types of measures have been used to assess HRQL in PAH.(7, 8) Nonetheless, studies systematically comparing the performance of different disease-specific instruments in PAH are lacking. Such information, juxtaposed against concurrent measures of generic HRQL, is of critical importance in order to interpret the results of studies that utilize one or another approach to HRQL in PAH. To compare the performance characteristics of cardiac-specific, pulmonary-specific, and generic HRQL measures in PAH, we co-administered all three types of instruments longitudinally within the context of a prospective, observational cohort study of individuals with well-defined disease.

METHODS

Study Subjects and Design

We consecutively screened and enrolled patients with an established diagnosis of PAH from the University of California San Francisco (UCSF) pulmonary hypertension clinic over a 5-year period from July 2003 to July 2008. Recruitment was performed in parallel with an existing prospective, observational cohort study of secondhand smoke exposure in patients with PAH.⁽⁹⁾ Inclusion criteria were: age 18 years or greater, PAH confirmed by right heart catheterization, a minimum of 3 months of PAH therapy prior to enrollment, and English fluency. Patients unable to complete a 6-minute walk test, diagnosed with an unstable psychiatric disorder, or actively smoking (a criterion of the parallel secondhand smoke study) were precluded from enrollment. Twenty-two subjects initially enrolled but with missing HRQL data were excluded from this analysis. Human subjects approval for this study was obtained through the institutional review board at UCSF in 2003 and has been renewed on an annual basis.

Subject characteristics, including demographic information, classification of PAH, and hemodynamic measurements from diagnostic right heart catheterization were abstracted from subjects' medical records using a standardized data collection form. World Health Organization (WHO) functional class, 6-minute walk distance (6MWD), Borg dyspnea index (post-walk), and HRQL (see instruments below) were prospectively assessed at baseline and at 6 month, 12 month, and post-24 month follow-up visits. Mean time elapsed from baseline to the follow-up time points was: 6.4, 12.9, and 47.7 months, respectively. Median time interval between

adjacent visits was 6.2 months (interquartile range 5.8 to 7.8 months). Of 128 subjects with HRQL data enrolled at baseline, 107 (84%), 65 (51%), and 56 (44%) completed 6, 12, and post-24 month visits, respectively. Overall mean observation time for the entire cohort was 27.6 months.

Study Measures

Three HRQL questionnaires were administered to each subject in random order at each visit: 1 generic measure and 2 different disease-specific measures (1 cardiac-specific and 1 pulmonary-specific).

General HRQL was assessed using the Medical Outcome Study 36-item Short Form Health Survey (SF-36, version 2). The SF-36 is a generic HRQL instrument comprised of 8 individual domains yielding 2 summary scores, a physical component summary (PCS) and a mental component summary (MCS).⁽¹⁰⁾ Summary component scores also can be calculated based on only 12 of the 36 items, yielding the SF-12 PCS and MCS. The possible range for each summary score and domain is 0 to 100, with a norm-based population mean of 50 (standard deviation [SD] = 10). Higher scores indicate better HRQL.

Cardiac-specific HRQL was assessed using an adaptation of the Minnesota Living with Heart Failure Questionnaire (LHFQ). The LHFQ was originally developed and validated for use in patients with congestive heart failure.⁽¹¹⁾ It has been adapted for use in pulmonary hypertension by substituting the term 'heart failure' with 'pulmonary hypertension'.⁽⁷⁾ The LHFQ is

comprised of 21 items utilizing a 6-point Likert response format. Individual item scores range from 0 to 5, with overall score ranging from 0 to 105. Higher scores indicate worse HRQL. Physical (8 items) and emotional (5 items) domains have been previously identified.(12)

Pulmonary-specific HRQL was assessed using the 20-item and 30-item Airways Questionnaire (AQ20, AQ30). The items of the AQ20 form a subset of the AQ30 items. The AQ20/30 was originally developed by the St. George's Respiratory Questionnaire (SGRQ) investigators to perform similarly but be less burdensome to administer.(13, 14) Its performance characteristics have been further adapted for use across a variety of respiratory conditions.(15) The original scale utilizes a 'Yes', 'No', or 'Not applicable' response format. The revised version we employed substitutes the term 'chest trouble' for 'breathing problem', and adds an additional response option, 'Unable', to items that relate to a specific activity.(15) Only positive responses (either 'Yes' or 'Unable') are scored and summed to provide a possible range of 0 to 20 for the AQ20 and 0 to 30 for the AQ30. As with the LFHQ, higher scores indicate worse HRQL. The AQ20 has been shown to yield similar results to the AQ30,(16) and its validity has been demonstrated in relation to well-established pulmonary-specific measures, such as the SGRQ and the Chronic Respiratory Disease Questionnaire.(14, 17)

In addition to the 3 measures above, a PAH-specific quality of life (QoL) measure, the Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR), was added to the study in 2007 as a protocol modification. The CAMPHOR was originally developed in the United Kingdom(18) and its validity recently reevaluated in PAH patients in the United States.(19) It is comprised of 3 scales intended to assess symptoms (25 items), activity (15 items), and QoL (25

items). Scores are calculated based on the sum of each individual scale. Higher scores indicate worse HRQL (symptoms and activity scales) and QoL.

Statistical Analysis

Mean (\pm SD) and median scores at baseline were calculated for each HRQL instrument. To assess for normality and possible floor or ceiling effects, we examined the frequency distribution for each instrument individually. In addition, the number and percent of subjects scoring either the minimum or maximum possible score were also calculated. Internal consistency was evaluated for both disease-specific measures using Cronbach's alpha.

Correlations among the different HRQL measures were examined using data from all subject visits. To account for repeated measures data, correlations were based on standardized beta coefficients obtained using general estimating equations (GEE) and robust variance estimates. The strength of associations between measures were reported using thresholds commonly used in the behavioral sciences: trivial (≤ 0.10), weak ($>0.1-0.3$), moderate ($>0.3-0.5$), strong ($>0.5-0.7$), very strong ($>0.7-0.9$) and near perfect ($>0.9-1.0$).⁽²⁰⁻²²⁾ The SF-12 demonstrated nearly identical results to the SF-36 for the PCS and MCS was not studied further. Near perfect correlations were also observed between the AQ20 and the AQ30 ($r=0.98$, $p<0.0001$, $n=110$) and thus the AQ30 was not further analyzed either.

Relationships among the HRQL measures and other clinical variables (WHO functional class, 6MWD and Borg dyspnea index) were assessed using all available visit data. GEE was used to

account for repeated measures. To avoid strong assumptions regarding the structure of the longitudinal data, we specified an independent correlation matrix with robust variance estimates. Differences in HRQL among levels of each clinical variable were first examined by an overall F-test, followed by testing for trend (i.e., monotonic relationship) across levels using linear contrasts. In the case of 6MWD and the Borg dyspnea index, categories were defined to approximate quartiles.

Relationships between *change* in HRQL measures and *change* in other clinical variables (WHO functional class, 6MWD and Borg score) were assessed in a similar fashion using paired visit data. Change was categorized as: better, same, or worse for each clinical variable. Change was defined as a difference of: 1 or more class for WHO functional class; >40 meters for 6MWD (equivalent to approximately 0.5 standard effect size); and 2 or more for the Borg dyspnea index. Testing for trend across categories of change was performed using F-tests with linear contrasts.

To assess predictors of general physical health status, we used standard linear regression to model the relationship between measures obtained at baseline and the SF-36 PCS (the dependent variable) at the last study visit for which such data were available (median follow up time: 12.3 months). Univariate models were constructed for each predictor variable. Separate multivariate models were constructed for the AQ20 and LHFQ as independent predictors of SF-36 PCS, adjusting for age, sex, race-ethnicity, etiology of PAH, and all other clinical variables. Among those subjects with CAMPHOR data at follow up, parallel models were constructed using the CAMPHOR as the dependent variable. All analyses were conducted in SAS 9.2 (Cary, NC) or STATA/IC 11.0 (College Station, TX).

RESULTS

Demographic and clinical characteristics for the 128 subjects included in this analysis are shown in **Table 1**. Subjects were approximately 50 years in age, predominantly female, and predominantly white, non-Hispanic. English-speaking Hispanic and Asian subjects represented substantial minorities. Idiopathic PAH and PAH associated with connective tissue disease accounted for the majority of subjects. The overall distributions of PAH etiology, hemodynamic parameters, WHO functional class, and 6-minute walk results are not substantially different when compared to those reported by other recent pulmonary hypertension registries.(23, 24)

Summary statistics for the SF-36, AQ20, and LHFQ at baseline are shown in **Table 2**. There was substantial impairment in the SF-36 PCS, but only minimal impairment in the MCS. Scores for both subscales approximated a normal distribution without evidence of any ceiling or floor effects (**Figure 1**, panels a and b). The domains most severely affected were: Physical Functioning, General Health, and Role-Physical (**Figure 2**).

Pulmonary-specific (AQ20) and cardiac-specific (LHFQ) measures also demonstrated evidence of impairment, with mean scores falling within the middle of the possible range for each scale (**Table 2**). Both disease-specific measures demonstrated deviations from a normal distribution, in particular at their tails (**Figure 1**, panels c and d). Minor ceiling effects (best possible [lowest] score) were observed for both measures, in particular the AQ20, for which 6% of subjects had the best possible score at baseline. Internal consistency, as measured by Cronbach's α (**Table 2**),

was very high for both pulmonary- and cardiac-specific measures, indicating a strong unidimensional construct in each case.

PAH-specific QoL was assessed among 67 subjects who completed the CAMPHOR at baseline (n=17) or at a subsequent visit (n=50). Mean \pm SD scores were 9.1 ± 6.4 for symptoms (possible range 0 to 25), 8.3 ± 5.1 for activity (possible range 0 to 30), and 7.7 ± 6.3 for QoL (possible range 0 to 25). Observed variance for the CAMPHOR was similar to that which has been described in other PAH cohorts.(18, 19) Differences in data collection for the CAMPHOR, and the proprietary nature of the instrument, precluded further psychometric evaluation of its performance relative to other measures.

Correlations among the SF-36, AQ20, and LHFQ are shown in **Table 3**. All correlations were statistically significant ($p < 0.01$) in the anticipated direction. Convergent validity of both disease-specific measures was supported by strong correlations with the SF-36. The physical and emotional domains of the LHFQ correlated most strongly with the SF-36 PCS and MCS subscales ($r = -0.73$ and -0.69 , respectively), providing further evidence of convergent validity. Pulmonary-specific HRQL, as measured by the AQ20, correlated very strongly with cardiac-specific HRQL, as measured by the total LHFQ ($r = 0.75$) as well as each of its domains.

Relationships among the 3 main HRQL measures and functional status, exercise capacity, and dyspnea are shown in **Table 4**. All 3 HRQL measures studied showed reasonable discrimination of subjects over categories of severity (i.e., known groups validity) based on WHO functional class, 6MWD, and Borg dyspnea index. Although a statistically significant monotonic trend was

detected in each case, both disease-specific measures showed evidence of poor discrimination between the 2 lowest quartiles of 6MWD. No significant associations were observed between any of the hemodynamic data obtained at the time of initial right heart catheterization and any of the 3 HRQL measures assessed at baseline (data not shown).

All 3 HRQL measures studied demonstrated responsiveness to change in functional class, exercise capacity, and dyspnea (standardized effect sizes shown in **Table 5**). All measures were less responsive to worsening than improvement, particularly the AQ20, suggesting possible floor effects. This phenomenon was least evident for the LHFQ. Among the 3 instruments, the LHFQ was the most responsive overall, although differences in performance were not substantial. Averaged anchor-based estimates of change derived using WHO functional class and 6MWD corresponded to a point change of: +3.1 (improved) to -1.9 (worsened) for the SF-36 PCS; -1.7 (improved) to +0.8 (worsened) for the AQ20; and -11.5 (improved) and +8.0 (worsened) for the LHFQ. Distribution-based estimates of change derived using the 1 standard error of measurement (SEM) criterion were: ± 3.1 for the SF-36 PCS, ± 1.8 for the AQ20, and ± 6.1 for the LHFQ.

Linear regression models, intended to assess the extent to which pulmonary-specific (AQ20) and cardiac-specific (LHFQ) measures predict general physical health status (SF-36 PCS) at follow up, are shown in **Table 6**. In our univariate predictor models, WHO functional class, 6-minute walk distance, Borg dyspnea index, the AQ20, and the LHFQ were all predictive of general physical health status; however, class of PAH therapy was not, including: intravenous or subcutaneous prostacyclin ($p=0.52$), inhaled prostacyclin ($p=0.42$), endothelin receptor

antagonists ($p=0.99$), and phosphodiesterase-5 inhibitors ($p=0.42$) (not shown in table). In our multivariate analyses, both the AQ20 and the LHFQ were independently predictive of general physical health status, even after taking into account functional class, 6-minute walk distance, Borg dyspnea index, class of PAH therapy, and subject characteristics. WHO functional class, 6-minute walk distance and Borg dyspnea index were not independently predictive of general physical health status in multivariate analyses, but a non-significant trend was observed for function class. In a model including both the AQ20 and LHFQ together (not shown in table), the LHFQ remained a significant predictor of the general physical health status ($p=0.006$), whereas the AQ20 did not.

A similar pattern of findings was observed using multiple linear regression to predict PAH-specific QoL at follow up, as assessed by the CAMPHOR QoL scale among a subgroup of 52 subjects (Online depository, **Table S1**). Both the AQ20 and the LHFQ were independently predictive of PAH-specific QoL, even after taking into account functional class, 6-minutes walk distance, Borg dyspnea index and subject characteristics. In a model including both the AQ20 and LHFQ together, the LHFQ remained a significant predictor of the PAH-specific QoL ($p<0.02$), whereas the AQ20 did not.

DISCUSSION

In this study, we compared the performance characteristics of cardiac-specific (LHFQ) and pulmonary-specific (AQ20) HRQL measures in patients with PAH with respect to traditional endpoints. Our results largely support the validity of both disease-specific instruments, but also indicate potential limitations for each. Both the LHFQ and AQ20 demonstrated difficulty discriminating between patients in the next to lowest and lowest quartiles of 6-minute walk distances. Although both instruments were sensitive to *improvement* in functional class, 6-minute walk distance, and Borg dyspnea index over time, they appeared less sensitive to *worsening* in the same parameters. Nonetheless, the LHFQ and AQ20 were highly intercorrelated and were each strong predictors of the general physical health and PAH-specific QoL measured longitudinally, even after taking functional class, exercise capacity, and perceived dyspnea into account. Overall, the LHFQ performed in a slighter better fashion than the AQ20.

In general, both cardiac- and respiratory-specific instruments performed in the expected manner. The mild ceiling effect observed for the AQ20 could be related to its limited number of activity-based items along with its “yes/no” response format. A possible floor effect for both the AQ20 and LHFQ is suggested by lack of discrimination among patients with shorter walk distances and decreased sensitivity to worsening over time, but alternative explanations exist. Lack of a difference in disease-specific HRQL scores between the two lower quartiles of 6MWD could be due to *non-dyspnea-related* impairment (e.g., pain, fatigue). Supporting this explanation is the observation that the SF-36 PCS was able to discriminate patients with shorter walk distances, and that similar floor effects were not observed when measures based predominantly on dyspnea

(functional class and Borg score) were used as the reference. Decreased sensitivity of the HRQL measures to worsening over time could reflect a true floor effect, but may be confounded by a healthy-survivor effect, whereby those with the worst HRQL have died, thus attenuating the observed decrease in HRQL among those who survived.

An alternative interpretation of our results could be that the SF-36 PCS outperformed either disease-specific instrument insofar as it did not demonstrate any ceiling/floor effects and was responsive to underlying changes in health status. It is important to keep in mind, however, that the SF-36 is a general measure of physical health status, and not dyspnea-related impairment. Its strong performance in this cohort is likely due to the fact that patients had few comorbidities and therefore the major driver of their physical health status was due to a cardiopulmonary limitation. Care must be taken in extrapolating our findings to other PAH populations, such as connective-tissue disease related PAH, where general measures of physical health status may be confounded by other sources of impairment.

Studies of PAH-related HRQL published to date have included either a cardiac-specific(7, 25) or respiratory-specific instrument,(8) but not both. Existing studies are also limited in size and follow up. In the only analysis to undertake a direct comparison of HRQL measures, Chua and colleagues studied 83 patients using pooled data from 3 different clinical trials.(26) Their study examined the LHFQ, SF-36, and Australian Quality of Life questionnaire (a preference-based utility measure). No respiratory-specific measure was included. Their analysis was primarily restricted to the detection of simple correlations, did not allow for the modeling of asymmetric

bidirectional change, and did not standardize parameter estimates, thereby precluding any comparisons of effect size.

Particular strengths of our study are not only the inclusion of different disease-specific measures, but also the prospective evaluation of a sizable longitudinal cohort of PAH patients. Moreover, we systematically and thoroughly examined the psychometric performance of both disease-specific instruments for normality/distribution effects, internal consistency (reliability), convergence and discrimination (construct validity), sensitivity to change over time (responsiveness), and ability to predict future health status (predictive validity). To increase interpretability, we provide distributional and anchor-based estimates of meaningful change for each measure. Our estimates, which approximate those obtained in other disease states,(27-30) should serve to inform future investigators when developing trial-specific responder definitions for PAH.

One weakness of our study is the lack of direct psychometric comparisons between the CAMPHOR and other disease-specific measures. The CAMPHOR was released for use in the United States in 2008,(19) at which point >80% of our study cohort had already completed a baseline visit. In response, we modified our study protocol to include its administration for all newly enrolled subjects and at follow up for remaining subjects. Our preliminary data support the predictive validity of existing measures in relation to the CAMPHOR, but further study is required. Another limitation of our study is, despite its comparatively large size in PAH, the lack of condition-specific subsets that might have allowed for stratified sub-analyses. The lack of

culturally-adapted instruments in multiple languages and utility indices also limit cross-cultural comparisons or preference-based evaluation of health states.

Bearing these limitations in mind, our results do show that existing cardiac- and respiratory-specific HRQL measures perform reasonably well in most situations. Use of less strictly defined dyspnea-based measures, such as the LHFQ or AQ20, could have potential advantages over an instrument exclusively focused on PAH. For example, the LHFQ has been used extensively in left-sided heart failure,(11, 12) and therefore might be a valuable tool for assessing dyspnea-based impairment in patients with pulmonary hypertension due to left ventricular dysfunction (WHO Group 2). Likewise, the AQ20 has been used extensively in patients with mixed obstructive disease,(14, 15) and therefore might be better suited for those patients with pulmonary hypertension in the setting of obstruction (WHO Group 3). In contrast, instruments exclusively focused on PAH cannot be applied in other disease states, and thus have limited value when comparing dyspnea-related impairment across conditions or when evaluating populations in which diagnostic heterogeneity may exist.

In conclusion, both cardiac-specific and respiratory-specific measures can be used to assess HRQL in PAH, but may be limited by lack of responsiveness to deterioration in health status over time. Overall, the cardiac-specific LHFQ demonstrates stronger performance characteristics than the respiratory-specific AQ20. Future comparison of these instruments with newer PAH-specific measures is imperative to understanding how best to utilize these important investigative tools.

ACKNOWLEDGEMENTS

We wish to acknowledge our study coordinator, Carla M. Teehankee, RN, for her dedication to enrolling subjects, conducting study visits, and managing the study database.

FIGURE LEGENDS

Figure 1. Health-related quality of life impairment across 8 domains of the SF-36. Higher scores indicated better quality of life. A score of 50 represents a norm-based population mean.

Figure 1

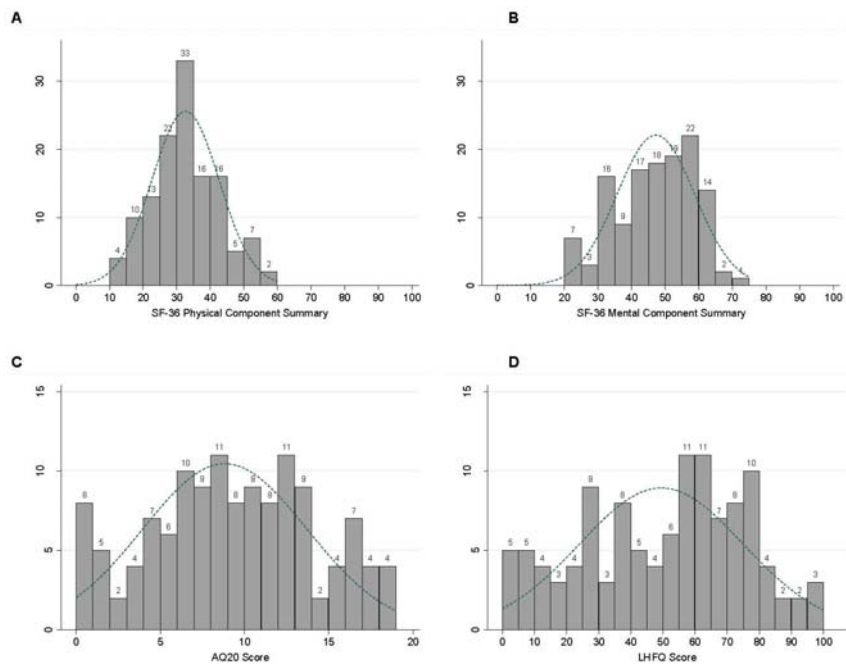
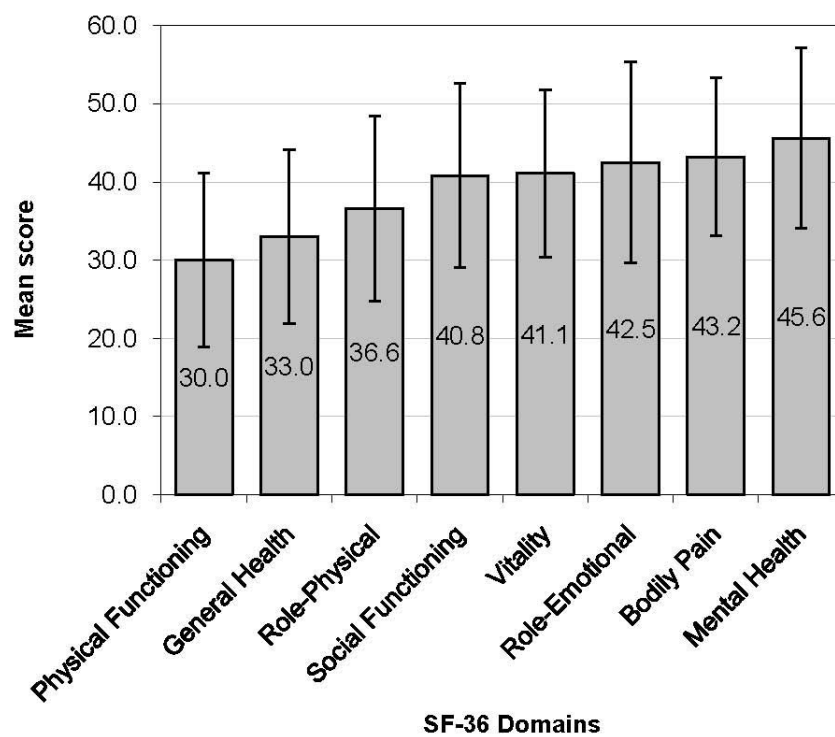


Figure 2. Frequency distribution of health-related quality of life scores by instrument. Panel A: SF-36 Physical Component Summary; Panel B: SF-36 Mental Component Summary; Panel C: AQ20; Panel D: LHFQ. Dashed line represents a normal distribution for comparison.

Figure 2.



REFERENCES

1. Garratt A, Schmidt L, Mackintosh A, Fitzpatrick R. Quality of life measurement: bibliographic study of patient assessed health outcome measures. *Bmj*. 2002;324:1417.
2. McLaughlin VV, Shillington A, Rich S. Survival in primary pulmonary hypertension: the impact of epoprostenol therapy. *Circulation*. 2002;106:1477-82.
3. Sitbon O, Humbert M, Nunes H, Parent F, Garcia G, Herve P, et al. Long-term intravenous epoprostenol infusion in primary pulmonary hypertension: prognostic factors and survival. *J Am Coll Cardiol*. 2002;40:780-8.
4. McLaughlin VV, Sitbon O, Badesch DB, Barst RJ, Black C, Galie N, et al. Survival with first-line bosentan in patients with primary pulmonary hypertension. *Eur Respir J*. 2005;25:244-9.
5. Galie N, Manes A, Negro L, Palazzini M, Bacchi-Reggiani ML, Branzi A. A meta-analysis of randomized controlled trials in pulmonary arterial hypertension. *Eur Heart J*. 2009;30:394-403.
6. Chen H, Taichman DB, Doyle RL. Health-related quality of life and patient-reported outcomes in pulmonary arterial hypertension. *Proc Am Thorac Soc*. 2008;5:623-30.
7. Cenedese E, Speich R, Dorschner L, Ulrich S, Maggiorini M, Jenni R, et al. Measurement of quality of life in pulmonary hypertension and its significance. *Eur Respir J*. 2006;28:808-15.
8. Taichman DB, Shin J, Hud L, Archer-Chicko C, Kaplan S, Sager JS, et al. Health-related quality of life in patients with pulmonary arterial hypertension. *Respir Res*. 2005;6:92.

9. Chen H, Eisner MD, Teehanke CM, Blanc PD, De Marco T. Environmental tobacco smoke and health status in patients with pulmonary arterial hypertension. *Chest*. 2005;128:364S.
10. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care*. 1992;30:473-83.
11. Rector RS, Kubo SH, Cohn JN. Patients self-assessment of their congestive heart failure: content, reliability and validity of a new measure, the Minnesota Living with Heart Failure questionnaire. *Heart Failure*. 1987;1:198-209.
12. Rector TS, Cohn JN. Assessment of patient outcome with the Minnesota Living with Heart Failure questionnaire: reliability and validity during a randomized, double-blind, placebo-controlled trial of pimobendan. Pimobendan Multicenter Research Group. *Am Heart J*. 1992;124:1017-25.
13. Quirk FH, Jones PW. Repeatability of two new short airways questionnaires. *Thorax*. 1994;49:1075.
14. Alemayehu B, Aubert RE, Feifer RA, Paul LD. Comparative analysis of two quality-of-life instruments for patients with chronic obstructive pulmonary disease. *Value Health*. 2002;5:436-41.
15. Chen H, Eisner MD, Katz PP, Yelin EH, Blanc PD. Measuring disease-specific quality of life in obstructive airway disease: validation of a modified version of the airways questionnaire 20. *Chest*. 2006;129:1644-52.
16. Barley EA, Quirk FH, Jones PW. Asthma health status measurement in clinical practice: validity of a new short and simple instrument. *Respir Med*. 1998;92:1207-14.

17. Hajiro T, Nishimura K, Jones PW, Tsukino M, Ikeda A, Koyama H, et al. A novel, short, and simple questionnaire to measure health-related quality of life in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 1999;159:1874-8.
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;15:103-15.
19. Gomberg-Maitland M, Thenappan T, Rizvi K, Chandra S, Meads DM, McKenna SP. United States validation of the Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR). *J Heart Lung Transplant*. 2008;27:124-30.
20. Davis JA. Elementary survey analysis. Englewood Cliffs, NJ: Prentice-Hall; 1971.
21. Cohen J. Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.; 1988.
22. de Vaus DA. Analyzing social science data: 50 key problems in data analysis. London: Sage Publications Ltd.; 2002.
23. Badesch DB, Raskob GE, Elliott CG, Krichman AM, Farber HW, Frost AE, et al. Pulmonary arterial hypertension: baseline characteristics from the REVEAL Registry. *Chest*. 2010;137:376-87.
24. Humbert M, Sitbon O, Chaouat A, Bertocchi M, Habib G, Gressin V, et al. Pulmonary arterial hypertension in France: results from a national registry. *Am J Respir Crit Care Med*. 2006;173:1023-30.
25. Zlupko M, Harhay MO, Gallop R, Shin J, Archer-Chicko C, Patel R, et al. Evaluation of disease-specific health-related quality of life in patients with pulmonary arterial hypertension. *Respir Med*. 2008;102:1431-8.

26. Chua R, Keogh AM, Byth K, O'Loughlin A. Comparison and validation of three measures of quality of life in patients with pulmonary hypertension. *Intern Med J.* 2006;36:705-10.
27. Bennett SJ, Oldridge NB, Eckert GJ, Embree JL, Browning S, Hou N, et al. Comparison of quality of life measures in heart failure. *Nurs Res.* 2003;52:207-16.
28. Angst F, Aeschlimann A, Stucki G. Smallest detectable and minimal clinically important differences of rehabilitation intervention with their implications for required sample sizes using WOMAC and SF-36 quality of life measurement instruments in patients with osteoarthritis of the lower extremities. *Arthritis Rheum.* 2001;45:384-91.
29. Kosinski M, Zhao SZ, Dedhiya S, Osterhaus JT, Ware JE, Jr. Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. *Arthritis Rheum.* 2000;43:1478-87.
30. Coteur G, Feagan B, Keininger DL, Kosinski M. Evaluation of the meaningfulness of health-related quality of life improvements as assessed by the SF-36 and the EQ-5D VAS in patients with active Crohn's disease. *Aliment Pharmacol Ther.* 2009;29:1032-41.

Table 1. Baseline characteristics of 128 patients with pulmonary arterial hypertension

Characteristic	All subjects (n=128)	
	Mean \pm SD	n (%)
Age (years)	50 \pm 14	
Sex		
Female		95 (74)
Race/Ethnicity		
White, non-Hispanic		80 (63)
Black		9 (7)
Hispanic		19 (15)
Asian		14 (11)
Other		6 (5)
Etiology of pulmonary hypertension		
Idiopathic pulmonary arterial hypertension		50 (39)
Associated pulmonary arterial hypertension		
Connective tissue diseases		36 (28)
Congenital heart diseases		3 (2)
Portal hypertension		16 (13)
HIV infection		6 (5)
Drug- and toxin-induced		11 (9)
Other		6 (5)
Right heart catheterization data		
Mean right atrial pressure (mmHg)*	11 \pm 8	
Mean pulmonary artery pressure (mmHg)	50 \pm 13	
Mean pulmonary capillary wedge pressure (mmHg)	9.8 \pm 4.0	
Cardiac index (L/min·m ²)*	2.4 \pm 0.9	
Pulmonary vascular resistance (dyne·sec·cm ⁻⁵)*	886 \pm 464	
WHO functional class		
Class I		8 (6)
Class II		45 (35)
Class III		66 (52)
Class IV		9 (7)
6-minute walk test [†]		
6-minute walk distance (m)	348 \pm 115	
Borg dyspnea index, post-walk	2.7 \pm 2.0	
Medications		
Intravenous or subcutaneous prostacyclin infusion		49 (38)
Inhaled prostacyclin		4 (3)
Oral endothelin receptor antagonist		82 (64)
Oral phosphodiesterase-5 inhibitor		41 (32)

WHO: World Health Organization

* Missing for 7 subjects

† Missing for 3 subjects

Table 2. Summary statistics of health-related quality-of-life measures at baseline

Instrument	Possible range	Actual range	n (%) min. score	n (%) max. score	Mean \pm SD score	Median score	Cronbach's alpha
SF-36							
Physical	0 – 100	13 – 57	0 (0)	0 (0)	32.6 \pm 10.0	32.3	--
Mental	0 – 100	22 – 70	0 (0)	0 (0)	47.2 \pm 11.6	48.6	--
AQ20	0 – 20	0 – 19	8 (6.3)	0 (0)	8.8 \pm 4.9	9.0	0.86
LHFQ*							
Total	0 – 105	0 – 98	2 (1.8)	0 (0)	49.5 \pm 25.4	55.0	0.94
Physical	0 – 40	0 – 40	5 (4.4)	2 (1.8)	21.4 \pm 11.4	23.0	0.94
Emotional	0 – 25	0 – 25	11 (9.7)	7 (6.1)	11.9 \pm 7.9	12.0	0.92

AQ20: Airways Questionnaire 20; LHFQ: Living with Heart Failure Questionnaire; SF-36: Short Form 36
Higher scores indicate better quality-of-life for the SF-36 and worse quality-of-life for the AQ20 and LHFQ

* Missing data for 14 subjects

Table 3. Correlations among different health-related quality-of life measures

	SF-36		AQ20	LHFQ		
	Physical	Mental		Total	Physical	Emotional
SF-36						
Physical	1.00					
Mental	0.20	1.00				
AQ20*	-0.53	-0.53	1.00			
LHFQ*						
Total	-0.63	-0.64	0.75	1.00		
Physical	-0.73	-0.53	0.74	--	1.00	
Emotional	-0.45	-0.69	0.66	--	0.73	1.00

AQ20: Airways Questionnaire 20; LHFQ: Living with Heart Failure Questionnaire; SF-36: Short Form 36

Correlations shown for 275 visits with complete data for all 3 quality-of-life instruments

p<0.01 for all correlations based on robust standard errors used to account for repeated measures

* Instruments are negatively weighted (higher scores represent worse health status)

Table 4. Relationship between health-related quality-of-life measures and functional status, exercise capacity, and dyspnea

		SF-36 PCS		AQ20		LHFQ (total)	
	n	Mean	p	Mean	p	Mean	p
WHO functional class							
I	23	45.7	<0.001	3.3	<0.001	16.8	<0.001
II	99	38.1		7.1		35.5	
III	135	29.9		9.7		56.2	
IV	18	21.3		12.2		69.4	
6-minute walk distance (m)*							
>425	71	38.7	<0.001	6.7	<0.002	34.4	<0.001
>375-425	58	34.7		7.7		41.0	
>275-375	66	32.5		9.6		53.7	
≤275	63	28.3		9.6		54.7	
Borg dyspnea index*							
0-1	77	40.7	<0.001	5.3	<0.001	30.8	<0.001
2	58	36.3		8.1		42.2	
3	71	30.0		9.6		50.6	
≥4	52	25.3		11.5		65.4	

Data shown for 275 visits with complete data for all 3 quality-of-life instruments.

Higher scores indicate better quality-of-life for the SF-36 and worse quality-of-life for the AQ20 and LHFQ.

Generalized estimating equations used to account for repeated measures data.

P values based on F-test for trend.

* 6-minute walk data missing for 17 visits.

Table 5. Responsiveness of health-related quality-of-life measures to changes in functional status, exercise capacity, and dyspnea

		SF-36 PCS		AQ20		LHFQ (total)	
	n	Effect size	p	Effect size	p	Effect size	p
WHO functional class							
Improved	33	0.31	0.09	-0.36	0.10	-0.50	0.01
Same	130	referent		referent		referent	
Worsened	35	-0.15		0.06		0.20	
6-minute walk distance*							
Improved	30	0.31	0.03	-0.34	0.006	-0.39	0.002
Same	110	referent		referent		referent	
Worsened	35	-0.23		0.28		0.43	
Borg dyspnea index*							
Improved	35	0.39	0.002	-0.60	0.03	-0.49	0.03
Same	113	referent		referent		referent	
Worsened	27	-0.29		0.01		0.26	

Data shown for 198 paired visits with complete data for all 3 quality-of-life instruments.

Change defined as ≥ 1 change in functional class, >40 m change in 6-minute walk distance, and ≥ 2 points in Borg score.

Change shown as standardized effect size based on generalized estimating equations to account for repeated measures data.

Higher scores indicate better quality-of-life for the SF-36 and worse quality-of-life for the AQ20 and LHFQ.

P values based on F-test for trend.

* Change in 6-minute walk data missing for 23 paired visits.

Table 6. Predictors of general physical health status (SF-36 PCS) in 101 patients with pulmonary arterial hypertension

Variable	Univariate Model		Multivariate Model 1		Multivariate Model 2	
	$\beta \pm SE$	p	$\beta \pm SE$	p	$\beta \pm SE$	p
WHO functional class						
I	referent		referent		referent	
II	-12.0 \pm 4.9	<0.001	-9.3 \pm 5.0	0.08	-9.0 \pm 4.9	0.08
III	-18.0 \pm 4.8		-12.4 \pm 5.5		-11.6 \pm 5.4	
IV	-32.4 \pm 5.9		-21.5 \pm 8.1		-20.7 \pm 7.9	
6-minute walk distance (per 10 m)	0.3 \pm 0.1	0.002	<0.1 \pm 0.1	0.46	-0.1 \pm 0.1	0.41
Borg dyspnea index (per 1 point)	-2.6 \pm 0.5	<0.001	-1.0 \pm 0.7	0.16	-0.8 \pm 0.7	0.27
AQ20 (per 1 SD)	-4.7 \pm 1.1	<0.001	-3.0 \pm 1.2	0.02	--	
LHFQ (per 1 SD)	-5.4 \pm 1.1	<0.001	--		-3.8 \pm 1.3	0.004

General physical health status measured using physical component summary of Short Form-36 at follow up (median 12.3 mo.).

Adjusted model includes all variables shown in addition to age, sex, race, etiology, and class of therapy.

Adjusted Model 1: $R^2 = 0.34$ ($p < 0.001$)

Adjusted Model 2: $R^2 = 0.36$ ($p < 0.001$)