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

H. Chen*, T. De Marco#, E.A. Kobashigawa#, P.P. Katz¶, V.W. Chang# and P.D. Blanc*

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

We prospectively assessed HRQoL in 128 patients with catheterisation-confirmed PAH at baseline and at 6, 12 and post-24 month follow-up visits. Cardiac-specific HRQoL was assessed using the Minnesota Living with Heart Failure Questionnaire (LHFQ); respiratory-specific HRQoL was assessed using the Airways Questionnaire 20 (AQ20); and general health status was assessed 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 patients based on World Health Organization (WHO) functional class, 6-min walking distance (6MWD) and Borg dyspnoea index (BDI), with the exception of a potential floor effect associated with low 6MWD. The LHFQ was more responsive than the AQ20 to changes over time in WHO functional class, 6MWD and BDI. In multivariate analyses, the LHFQ and AQ20 were each longitudinal predictors of general health status, independent of functional class, 6MWD and BDI.

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

KEYWORDS: Pulmonary hypertension, quality of life

Over the past two decades, clinical research in cardiopulmonary disease has broadened from a purely physiological focus to a more comprehensive approach to health-related end-points. In particular, there has been rapid growth of interest in the development and application of patient-centred 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 dyspnoea 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 (HRQoL).

Despite the wealth of data on HRQoL in COPD and heart failure, relatively few studies have addressed HRQoL in PAH. The studies that have been performed have utilised various types of instruments, both generic and disease-specific [6]. Generic HRQoL instruments can be advantageous in that they provide results that are comparable

For editorial comments, see page 512.
This article has supplementary material accessible at www.erj.ersjournals.com

across a heterogeneous mix of conditions; however, such approaches generally provide measures that are less sensitive to the nuances of the specific disease in question. In contrast, disease-specific instruments emphasise aspects of health that are impacted by a particular type of disorder; for example, breathing-related symptoms and their potentially negative impact on a patient's sense of well-being.

As cardiac- and pulmonary-specific measures focus heavily on dyspnoea-related impairment, both have been used to assess HRQoL 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 HRQoL, is of critical importance for the interpretation of results of studies that utilise one or another approach to HRQoL in PAH. To compare the performance characteristics of cardiac-specific, pulmonary-specific and generic HRQoL measures in PAH, we co-administered all three types of instrument longitudinally within the context of a prospective, observational cohort study of individuals with well-defined disease.

METHODS
Study subjects and design
Over a 5-yr period from July 2003 to July 2008, we consecutively screened and enrolled patients with an established diagnosis of PAH from the Pulmonary Hypertension Clinic of the University of California San Francisco (UCSF; San Francisco, CA, USA). Recruitment was performed in parallel with an existing prospective, observational cohort study of second-hand smoke exposure in patients with PAH [9]. Inclusion criteria were: $\geqslant 18$ yrs of age; PAH confirmed by right heart catheterisation; $\leqslant 3$ months of PAH therapy prior to enrolment; and fluent English. Patients unable to complete a 6-min walk test, diagnosed with an unstable psychiatric disorder, or actively smoking (a criterion of the parallel second-hand smoke study) were precluded from enrolment. 22 subjects who were initially enrolled but had missing HRQoL data were excluded from analysis. Human subject study approval was obtained from the Institutional Review Board of the UCSF in 2003 and has been renewed on an annual basis.

Subject characteristics, including demographic information, classification of PAH, and haemodynamic measurements from diagnostic right heart catheterisation, were taken from subjects' medical records using a standardised data collection form. World Health Organization (WHO) functional class, 6-min walking distance (6MWD), Borg dyspnoea index (BDI; post-walk), and HRQoL (see instruments below) were prospectively assessed at baseline and at 6, 12, and post-24 month follow-up visits. The mean time elapsed from baseline to follow-up time-points was: 6.4, 12.9 and 47.7 months, respectively. The median time interval between adjacent visits was 6.2 months (interquartile range (IQR) 5.8–7.8 months). Of the 128 subjects with HRQoL data who were enrolled at baseline, 107 (84%), 65 (51%), and 56 (44%) completed 6, 12 and post-24 month visits, respectively. The overall mean observation time for the entire cohort was 27.6 months.

Study measures
Three HRQoL questionnaires were administered to each subject in random order at each visit: one generic measure and two different disease-specific measures (one cardiac-specific and one pulmonary-specific).

General HRQoL was assessed using the Medical Outcome Study 36-item Short Form Health Survey (SF-36; version 2). The SF-36 is a generic HRQoL instrument comprised of eight individual domains yielding two summary scores, a physical component summary (PCS) and a mental component summary (MCS) [10]. Summary component scores can also 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 0–100, with a norm-based population mean $\pm$ SD of 50 ± 10. Higher scores indicate better HRQoL.

Cardiac-specific HRQoL 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 utilising a six-point Likert response format. Individual item scores range 0–5, with the overall score ranging 0–105. Higher scores indicate worse HRQoL. Physical (eight items) and emotional (five items) domains have been previously identified [12].

Pulmonary-specific HRQoL 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 with the aim that it would perform in a similar manner but would 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 utilises a ''Yes'', ''No'' or ''Not applicable'' response format. We employed the revised version, which 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 0–20 for the AQ20 and 0–30 for the AQ30. As with the LHFQ, higher scores indicate worse HRQoL. 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 three 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 UK [18] and its validity was recently re-evaluated in PAH patients in the USA [19]. It is comprised of three 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 HRQoL (symptoms and activity scales) and QoL.

Statistical analysis
Mean $\pm$ SD and median scores at baseline were calculated for each HRQoL instrument. To assess for normality and possible floor or ceiling effects, we examined the frequency distribution

of each instrument individually. In addition, the number and percentage of subjects scoring either the minimum or maximum possible score was also calculated. Internal consistency was evaluated for both disease-specific measures using Cronbach's α.

Correlations among the different HRQoL measures were examined using data from all subject visits. To account for repeated measures data, correlations were based on standardised β coefficients obtained using general estimating equations (GEE) and robust variance estimates. The strength of associations between measures was reported using thresholds commonly used in behavioural sciences: trivial ($\leqslant 0.1$), 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 ($r=0.98$, $p<0.0001$; $n=110$); thus, the AQ30 was not further analysed either.

Relationships among the HRQoL measures and other clinical variables (WHO functional class, 6MWD and BDI) 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 HRQoL among levels of each clinical variable were first examined using an overall F-test, followed by a test for trend (i.e. monotonic relationship) across levels using linear contrasts. For 6MWD and the BDI, categories were defined to approximate quartiles.

Relationships between change in HRQoL measures and change in other clinical variables (WHO functional class, 6MWD and BDI) were assessed in a similar fashion using paired visit data. Change was categorised as: better, same or worse for each clinical variable. Change was defined as a difference of: $\geqslant 1$ for WHO functional class; >40 m for 6MWD (equivalent to ~0.5 standard effect size); and $\geqslant 2$ for BDI. 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, ethnicity, aetiology 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 (SAS, Cary, NC, USA) or STATA/IC 11.0 (Stata Corp. LP, College Station, TX, USA).

RESULTS

Demographic and clinical characteristics for the 128 subjects included in the analysis are shown in table 1. Subjects were ~50 yrs of 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 aetiology, haemodynamic parameters, WHO functional class, and 6-min walk test results were not substantially different when compared with 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 effect (fig. 1a and b). The domains most severely affected were: physical functioning, general health and physical role (fig. 2).

TABLE 1	Baseline characteristics of the 128 patients with pulmonary arterial hypertension (PAH)
Age yrs	50 ± 14
Females	95 (74)
Ethnicity	
White, non-Hispanic	80 (63)
Black	9 (7)
Hispanic	19 (15)
Asian	14 (11)
Other	6 (5)
Aetiology of pulmonary hypertension	
Idiopathic PAH	50 (39)
Associated PAH	
Connective tissue disease	36 (28)
Congenital heart disease	3 (2)
Portal hypertension	16 (13)
HIV	6 (5)
Drug- and toxin-induced	11 (9)
Other	6 (5)
Right heart catheterisation	
Mean right atrial pressure[#] mmHg	11 ± 8
Mean pulmonary artery pressure mmHg	50 ± 13
Mean pulmonary capillary wedge pressure mmHg	9.8 ± 4.0
Cardiac index[#] $L \cdot min^{-1} \cdot m^{-2}$	2.4 ± 0.9
Pulmonary vascular resistance[#] $dyne \cdot s \cdot cm^{-5}$	886 ± 464
WHO functional class	
I	8 (6)
II	45 (35)
III	66 (52)
IV	9 (7)
6-minute walk test[¶]	
6MWD m	348 ± 115
BDI post-walk	2.7 ± 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)

Data are presented as mean $\pm$ SD or n (%). WHO: World Health Organization; 6MWD: 6-min walking distance; BDI: Borg dyspnoea index. [#]: data missing for seven subjects; [¶]: data missing for three subjects.

TABLE 2	Summary statistics of health-related quality-of-life measures at baseline						
Instrument	Possible range	Actual range	Min. score n (%)	Max. score n (%)	Mean ± SD score	Median score	Cronbach's α
SF-36							
PCS	0–100	13–57	0 (0)	0 (0)	32.6±10.0	32.3	
MCS	0–100	22–70	0 (0)	0 (0)	47.2±11.6	48.6	
AQ20	0–20	0–19	8 (6.3)	0 (0)	8.8±4.9	9.0	0.86
LHFQ[#]							
Total	0–105	0–98	2 (1.8)	0 (0)	49.5±25.4	55.0	0.94
Physical	0–40	0–40	5 (4.4)	2 (1.8)	21.4±11.4	23.0	0.94
Emotional	0–25	0–25	11 (9.7)	7 (6.1)	11.9±7.9	12.0	0.92

Higher scores indicate better quality-of-life for the 36-item Short Form Health Survey (SF-36), and worse quality-of-life for the Airways Questionnaire 20 (AQ20) and the Living with Heart Failure Questionnaire (LHFQ). Min.: minimum; max.: maximum; PCS: physical component summary; MCS: mental component summary. [#]: data missing for 14 subjects.

Pulmonary- (AQ20) and cardiac-specific (LHFQ) measures also demonstrated evidence of impairment, with mean scores falling in the middle of the possible range for each scale (table 2). Both disease-specific measures demonstrated deviations from a normal distribution, particularly at their tails (fig. 1c and d). Minor ceiling effects (best possible (lowest) score) were observed for both measures, particularly in the AQ20, for which 6% of subjects had the best possible score at baseline. Internal consistency, as measured by Cronbach's alpha (table 2), was very high for both pulmonary- and
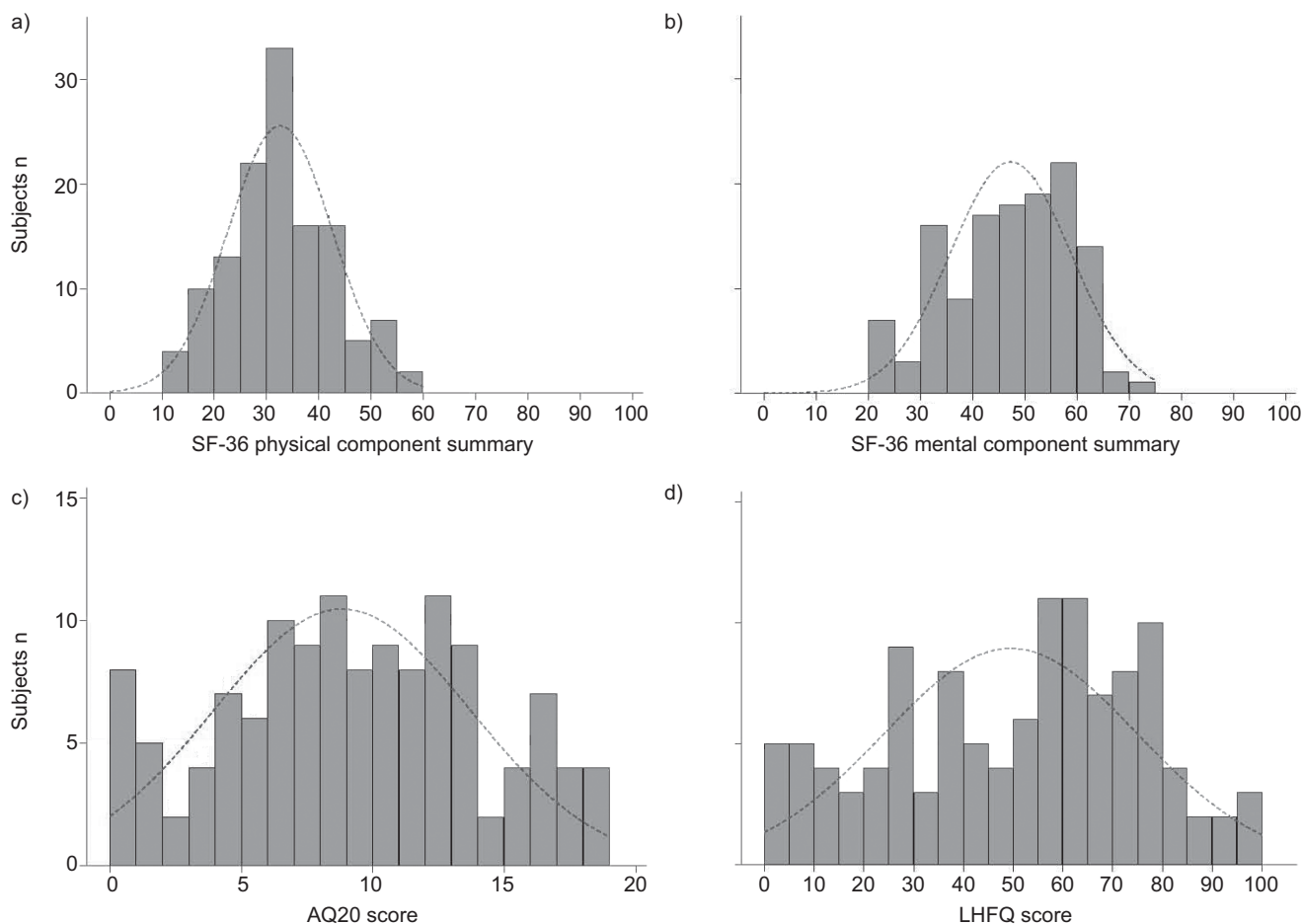

FIGURE 1. Frequency distribution of health-related quality of life scores by instrument: a) 36-item Short Form Health Survey (SF-36) physical component summary; b) SF-36 mental component summary; c) Airways Questionnaire 20 (AQ20); d) Living with Heart Failure Questionnaire (LHFQ). ---: normal distribution for comparison.
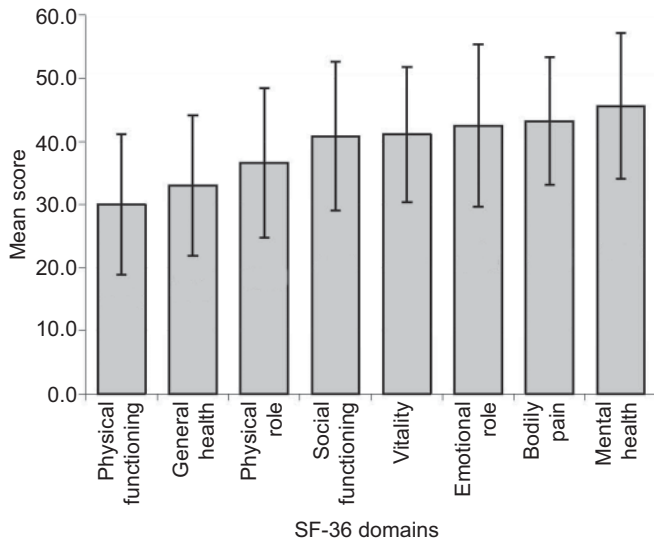
FIGURE 2. Health-related quality of life impairment across eight domains of the 36-item Short Form Health Survey (SF-36). Higher scores indicate better quality of life. A score of 50 represents a norm-based population mean.

cardiac-specific measures, indicating a strong unidimensional construct in each case.

PAH-specific QoL was assessed in 67 subjects who completed the CAMPHOR at baseline (n=17) or at a subsequent visit (n=50). Mean±SD scores were 9.1±6.4 for symptoms (possible range 0–25), 8.3±5.1 for activity (possible range 0–30) and 7.7±6.3 for QoL (possible range 0 –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

TABLE 3	Correlations amongst different health-related quality-of life measures					
	SF-36		**AQ20**	**LHFQ**		
	PCS	**MCS**		**Total**	**Physical**	**Emotional**
SF-36						
PCS	1.00					
MCS	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

Correlations are shown for 275 visits with complete data for all three quality-of-life instruments. SF-36: 36-item Short Form Health Survey; PCS: physical component summary; MCS: mental component summary; AQ20: Airways Questionnaire 20; LHFQ: Living with Heart Failure Questionnaire. 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).

the CAMPHOR and the proprietary nature of the instrument, precluded further psychometric evaluation of its performance relative to other measures.

Correlations amongst 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 HRQoL, as measured by the AQ20, correlated very strongly with cardiac-specific HRQoL, as measured by the total LHFQ (r=0.75) as well as each of its domains.

Relationships between the three main HRQoL measures and functional status, exercise capacity and dyspnoea are shown in table 4. All three HRQoL measures studied showed reasonable discrimination of subjects in terms of categories of severity (i.e. known groups validity) based on WHO functional class, 6MWD and BDI. Although a statistically significant monotonic trend was detected in each case, both disease-specific measures showed evidence of poor discrimination between the two lowest quartiles of 6MWD. No significant associations were observed between any of the haemodynamic data obtained at the time of initial right heart catheterisation and any of the three HRQoL measures assessed at baseline (data not shown).

All three HRQoL measures studied demonstrated responsiveness to change in functional class, exercise capacity and dyspnoea (standardised effect sizes shown in table 5). All measures were less responsive to worsening than improvement, particularly AQ20, suggesting possible floor effects. This phenomenon was least evident for the LHFQ. Of the three 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) to +8.0 (worsened) for the LHFQ. Distribution-based estimates of change derived using the 1 standard error of measurement 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- (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, 6MWD, BDI, the AQ20 and the LHFQ were all predictive of general physical health status. The type of PAH therapy, however, was not, and this included: 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) (data not shown). In our multivariate analyses, both the AQ20 and the LFHQ were independently predictive of general physical health status, even after taking into account functional class, 6MWD, BDI, type of PAH therapy and subject characteristics. WHO functional class, 6MWD and BDI were not independently predictive of general physical health status in multivariate analyses, but a

TABLE 4 Relationship between health-related quality-of-life (QoL) measures and functional status, exercise capacity and dyspnoea

	Subjects n	SF-36 PCS		AQ20		LHFQ (total)	
		Mean	p-value[#]	Mean	p-value[#]	Mean	p-value[#]
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	
6MWD 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	
BDI[¶]							
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 are shown for 275 visits with complete data for all three QoL instruments. Higher scores indicate better QoL for the 36-item Short Form Health Survey (SF-36), and worse QoL for the Airways Questionnaire 20 (AQ20) and the Living with Heart Failure Questionnaire (LHFQ). Generalised estimating equations were used to account for repeated measures data. WHO: World Health Organization; 6MWD: 6-min walking distance; BDI: Borg dyspnoea index. [#]: based on F-test for trend; [¶]: data missing for 17 visits.

non-significant trend was observed for function class. In a model including both the AQ20 and LHFQ together (data not shown), the LHFQ remained a significant predictor of 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 (table S1 of the online supplementary material). Both

TABLE 5 Responsiveness of health-related quality-of-life (QoL) measures to changes in functional status, exercise capacity and dyspnoea

	Subjects n	SF-36 PCS		AQ20		LHFQ (total)	
		Effect size	p-value[#]	Effect size	p-value[#]	Effect size	p-value[#]
WHO functional class							
Improved	33	0.31		-0.36		-0.50	
Same	130	Ref.	0.09	Ref.	0.10	Ref.	0.01
Worsened	35	-0.15		0.06		0.20	
6MWD[¶]							
Improved	30	0.31		-0.34		-0.39	
Same	110	Ref.	0.03	Ref.	0.006	Ref.	0.002
Worsened	35	-0.23		0.28		0.43	
BDI[¶]							
Improved	35	0.39		-0.60		-0.49	
Same	113	Ref.	0.002	Ref.	0.03	Ref.	0.03
Worsened	27	-0.29		0.01		0.26	

Data shown for 198 paired visits with complete data for all three QoL instruments. Change is defined as a change of ⩾1 in functional class, >40 m in 6-minute walking distance (6MWD), and ⩾2 points in Borg dyspnoea index (BDI). Change is shown as the standardised effect size based on generalised estimating equations to account for repeated measures data. Higher scores indicate better QoL for the 36-item Short Form Health Survey (SF-36) and worse QoL for the Airways Questionnaire 20 (AQ20) and the Living with Heart Failure Questionnaire (LHFQ). WHO: World Health Organization; ref: reference. [#]: based on F-test for trend; [¶]: data missing for 23 paired visits.

TABLE 6	Predictors of general physical health status[#] in 101 patients with pulmonary arterial hypertension					
Variable	**Univariate model**		**Multivariate model 1**		**Multivariate model 2**	
	β±SE	p-value	β±SE	p-value	β±SE	p-value
WHO functional class						
I	Ref.	<0.001	Ref.	0.08	Ref.	0.08
II	-12.0±4.9		-9.3±5.0		-9.0±4.9	
III	-18.0±4.8		-12.4±5.5		-11.6±5.4	
IV	-32.4±5.9		-21.5±8.1		-20.7±7.9	
6MWD per 10 m	0.3±0.1	0.002	<0.1±0.1	0.46	-0.1±0.1	0.41
BDI per 1 point	-2.6±0.5	<0.001	-1.0±0.7	0.16	-0.8±0.7	0.27
AQ20 per 1 SD	-4.7±1.1	<0.001	-3.0±1.2	0.02		
LHFQ per 1 SD	-5.4±1.1	<0.001			-3.8±1.3	0.004

[#]: measured using the physical component summary of the 36-item Short Form Health Survey at follow-up (median 12.3 m). The adjusted models included all variables shown in addition to age, sex, race, aetiology and class of therapy. Multivariate model 1: R^2=0.34 (p<0.001); multivariate model 2: R^2=0.36 (p<0.001). WHO: World Health Organization; 6MWD: 6-min walking distance; BDI: Borg dyspnoea index; AQ20: Airways Questionnaire 20; LHFQ: Living with Heart Failure Questionnaire; ref.: reference.

the AQ20 and the LFHQ were independently predictive of PAH-specific QoL, even after taking into account functional class, 6MWD, BDI and subject characteristics. In a model including both the AQ20 and LHFQ, the LHFQ remained a significant predictor of PAH-specific QoL (p<0.02), whereas the AQ20 did not.

DISCUSSION

In the present study, we compared the performance characteristics of cardiac- (LHFQ) and pulmonary-specific (AQ20) HRQoL measures in patients with PAH with respect to traditional end-points. Our results largely support the validity of both disease-specific instruments, but also indicate the potential limitations of each. Both the LHFQ and AQ20 demonstrated difficulty in discriminating between patients in the next to lowest and lowest quartiles of 6MWDs. Although both instruments were sensitive to improvement in functional class, 6MWD and BDI 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 general physical health and PAH-specific QoL measured longitudinally, even after taking functional class, exercise capacity and perceived dyspnoea into account. Overall, the LHFQ performed slightly better 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 was suggested by the lack of discrimination among patients with shorter walk distances and decreased sensitivity to worsening over time, but alternative explanations exist. The lack of a difference in disease-specific HRQoL scores between the two lower quartiles of 6MWD could be due to non-dyspnoea-related impairment (*e.g.* pain, fatigue). This explanation is supported by the observation that the SF-36 PCS was able to discriminate between patients with shorter walk distances, and that similar floor effects were not observed

when measures based predominantly on dyspnoea (functional class and BDI score) were used as the reference. Decreased sensitivity of the HRQoL 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 HRQoL have died, thus attenuating the observed decrease in HRQoL 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 dyspnoea-related impairment. It is likely that its strong performance in this cohort is due to the fact that patients had few comorbidities and therefore, the major driver of their physical health status was 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 HRQoL published to date have included either a cardiac- [7, 25] or a 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 HRQoL measures, CHUA *et al.* [26] studied 83 patients using pooled data from three different clinical trials. 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 modelling of asymmetric bidirectional change, and did not standardise parameter estimates, thereby precluding any comparison of effect size.

The particular strengths of our study were not only the inclusion of different disease-specific measures but also the

prospective evaluation of a sizeable 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 provided distributional and anchor-based estimates of meaningful change for each measure. Our estimates, which approximated 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 was the lack of direct psychometric comparison between the CAMPHOR and other disease-specific measures. The CAMPHOR was released for use in the USA 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 the remaining subjects. Our preliminary data support the predictive validity of existing measures in relation to the CAMPHOR, but further study is required. Despite its comparatively large size in PAH, another limitation of our study was 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 limited 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 HRQoL measures perform reasonably well in most situations. Use of less strictly defined dyspnoea-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 dyspnoea-based impairment in patients with pulmonary hypertension due to left ventricular dysfunction (WHO group II). Likewise, the AQ20 has been used extensively in patients with mixed obstructive disease [14, 15] and therefore might be better suited to those patients with pulmonary hypertension in the setting of obstruction (WHO group III). In contrast, instruments exclusively focused on PAH cannot be applied in other disease states and thus have limited value when comparing dyspnoea-related impairment across conditions or when evaluating populations in which diagnostic heterogeneity may exist.

In conclusion, both cardiac- and respiratory-specific measures can be used to assess HRQoL 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 if we are to understand how best to utilise these important investigative tools.

SUPPORT STATEMENT

STATEMENT OF INTEREST

ACKNOWLEDGEMENTS

REFERENCES

1 Garratt A, Schmidt L, Mackintosh A, *et al.* 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: 1477–1482.

3 Sitbon O, Humbert M, Nunes H, *et al.* Long-term intravenous epoprostenol infusion in primary pulmonary hypertension: prognostic factors and survival. *J Am Coll Cardiol* 2002; 40: 780–788.

4 McLaughlin VV, Sitbon O, Badesch DB, *et al.* Survival with first-line bosentan in patients with primary pulmonary hypertension. *Eur Respir J* 2005; 25: 244–249.

5 Galie N, Manes A, Negro L, *et al.* 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: 623–630.

7 Cenedese E, Speich R, Dorschner L, *et al.* Measurement of quality of life in pulmonary hypertension and its significance. *Eur Respir J* 2006; 28: 808–815.

8 Taichman DB, Shin J, Hud L, *et al.* Health-related quality of life in patients with pulmonary arterial hypertension. *Respir Res* 2005; 6: 92.

9 Chen H, Eisner MD, Teehankee CM, *et al.* 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: 473–483.

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: 1017–1025.

13 Quirk FH, Jones PW. Repeatability of two new short airways questionnaires. *Thorax* 1994; 49: 1075.

14 Alemayehu B, Aubert RE, Feifer RA, *et al.* Comparative analysis of two quality-of-life instruments for patients with chronic obstructive pulmonary disease. *Value Health* 2002; 5: 436–441.

15 Chen H, Eisner MD, Katz PP, *et al.* Measuring disease-specific quality of life in obstructive airway disease: validation of a modified version of the airways questionnaire 20. *Chest* 2006; 129: 1644–1652.

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: 1207–1214.

17 Hajiro T, Nishimura K, Jones PW, *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: 1874–1878.

18 McKenna SP, Doughty N, Meads DM, *et al.* 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.

19 Gomberg-Maitland M, Thenappan T, Rizvi K, *et al.* United States validation of the Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR). *J Heart Lung Transplant* 2008; 27: 124–130.

20 Davis JA. Elementary Survey Analysis. Englewood Cliffs, Prentice-Hall; 1971.

21 Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd Edn. Hillsdale, 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, *et al.* Pulmonary arterial hypertension: baseline characteristics from the REVEAL Registry. *Chest* 2010; 137: 376–387.

24 Humbert M, Sitbon O, Chaouat A, *et al.* Pulmonary arterial hypertension in France: results from a national registry. *Am J Respir Crit Care Med* 2006; 173: 1023–1030.

25 Zlupko M, Harhay MO, Gallop R, *et al.* Evaluation of disease-specific health-related quality of life in patients with pulmonary arterial hypertension. *Respir Med* 2008; 102: 1431–1438.

26 Chua R, Keogh AM, Byth K, *et al.* Comparison and validation of three measures of quality of life in patients with pulmonary hypertension. *Intern Med J* 2006; 36: 705–710.

27 Bennett SJ, Oldridge NB, Eckert GJ, *et al.* Comparison of quality of life measures in heart failure. *Nurs Res* 2003; 52: 207–216.

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: 384–391.

29 Kosinski M, Zhao SZ, Dedhiya S, *et al.* 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: 1478–1487.

30 Coteur G, Feagan B, Keininger DL, *et al.* 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: 1032–1041.