Published in final edited form as:

Clin Chest Med. 2007 March ; 28(1): 75–viii.

Surrogate End Points in Pulmonary Arterial Hypertension: Assessing the Response to Therapy

Jennifer L. Snow, M.D.^a and Steven M. Kawut, M.D., M.S.^{b,c}

^a Department of Medicine, University of Pennsylvania School of Medicine, Philadelphia, PA

^b Department of Medicine, College of Physicians and Surgeons, Columbia University, New York, NY

^c Department of Epidemiology, Joseph L. Mailman School of Public Health, Columbia University, New York, NY

Abstract

Pulmonary arterial hypertension (PAH) is a rare disease which is characterized by increased pulmonary vascular resistance and right heart failure. Recent discoveries in disease pathophysiology have been translated into effective therapies tested in clinical trials. The studies which have led to the regulatory board approval of therapies for PAH have focused on surrogate and intermediate end points, thought to reflect quantity and quality of life, respectively. However, validation of many of these surrogates is incomplete. It is also unknown which indicators of function or long-term survival should be used to formulate decisions regarding addition, discontinuation, or combination of therapies. Identification of suitable end points would therefore not only help investigators design appropriate clinical trials, but also assist clinicians in caring for patients with PAH. Hemodynamic, cardiac imaging, plasma biomarkers, and exercise testing hold some promise as potential surrogate end points for PAH. Functional status and quality of life assessments may also have important roles. Future studies should validate the most promising surrogate markers, so that patients, clinicians, subjects, and investigators may benefit from the advantages they confer on clinical care and on clinical trials.

The past several years have yielded an explosion of randomized clinical trials (RCTs) of several new medical therapies for pulmonary arterial hypertension (PAH). The hopes for these and all treatments for potentially fatal diseases are 1) to improve how a patient feels or functions and/or 2) to prolong survival. While simply stated, the assessment of progress towards these goals may be quite complex [1–6].

New PAH therapies have gained regulatory board approval based on studies with intermediate and surrogate end points as primary outcomes, believed to reflect the impact of these drugs on quality of life (QOL) and survival, respectively. While hemodynamic, plasma, cardiac imaging, exercise, functional, and QOL parameters are promising candidates, the reliability and validity of these measures are not clearly established. In addition, there are no RCTs addressing the use of such end points in guiding management decisions in PAH, such as changing, intensifying, escalating, or combining therapies. Despite this lack of evidence, clinicians who

Address for correspondence: Steven M. Kawut, M.D., M.S., Division of Pulmonary, Allergy, and Critical Care Medicine, Columbia University College of Physicians and Surgeons, PH 8E, Rm. 101, 622 W. 168th Street, New York City, NY 10032. Phone: 212-342-5430, Fax: 212-342-2201, Email: sk2097@columbia.edu.

Supported in part by NIH grants K23 HL67771, RO3 DK064103, RO1 HL086719, and RO1 HL082895.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

care for these patients frequently rely on end points measured in clinical practice and RCTs based on surrogate and intermediate end points to make important treatment decisions, warranting a review of their use.

Definitions

The National Institutes of Health define a surrogate as a characteristic that is objectively measured and evaluated as an indicator of a normal biologic process, pathogenic process, or pharmacologic response to a therapeutic intervention intended to substitute for a clinical end point [7]. The Food and Drug Administration defines a surrogate end point as “a laboratory measurement or physical sign that is used in therapeutic trials as a substitute for a clinically meaningful end point... and is expected to predict the effect of the therapy [8].” Restrictive definitions require that a therapy’s effect on the surrogate encompass, in its entirety, the impact on the clinical end point [9].

Alternatively, an intermediate end point is a parameter which directly reflects how a patient feels [10]. Intermediate end points are important to patients and physicians in and of themselves (unlike a surrogate) even without a validated relationship to survival. For example, a reduction in pulmonary vascular resistance (PVR) does not reflect how much farther a patient can walk without dyspnea but may have long term implications regarding survival, making PVR a potential surrogate end point for PAH. On the other hand, a reported increase in exercise capacity without dyspnea (i.e., a decrease in New York Heart Association (NYHA) class) reflects improved patient function without any assumptions regarding long-term outcome, making this a potential intermediate end point. In addition, an intermediate end point (such as NYHA class) could concomitantly serve as a surrogate end point if treatment-induced changes in the measure predict differences in event-free survival (see below).

What are the “hard” clinical end points in PAH? Prolonging the time until death is the primary goal of treating PAH, however there are other health states which patients and physicians wish to avoid. Lung transplantation is curative for PAH but entails a complicated surgical procedure, peri-operative risk, multiple medications with side-effects, and intensive follow-up. Patients therefore prefer to postpone this and other operative interventions (e.g., atrial septostomy). Hospitalization is another end point which is uncomfortable and inconvenient for the patient and could itself serve as a surrogate end point for worse long-term outcome. Intravenous prostacyclin analogue therapy is an important treatment for PAH, but drug delivery is suboptimal. As a result, both physicians and patients may view the initiation of parenteral prostacyclin analogue therapy as an “undesirable” event (or at least a “trade-off”) despite proven clinical efficacy [11–13]. A measurement made in the short-term which reflected the long-term risk of reaching one or more of these end points would facilitate the assessment of treatment success.

Event-free survival is not the only end point of interest in PAH. Patients complain of dyspnea on exertion, fatigue, lower extremity edema, and limitation in performing daily tasks. A reliable metric incorporating these facets of disease activity (intermediate end point) would therefore be integral to developing drugs to ameliorate such symptoms. A therapy which successfully targeted intermediate end points but did not prolong time to clinical events would still be considered useful, as long as it did not significantly shorten event-free survival. In fact, most current therapies for PAH have been approved on this basis.

Regulatory Issues

The Food and Drug Modernization Act permits approval for drugs for serious diseases without alternatives with a demonstrated “effect on a surrogate end point that is reasonably likely, based on epidemiologic, therapeutic, pathophysiologic, or other evidence, to predict clinical benefit

or on the basis of an effect on a clinical end point other than survival or irreversible morbidity [8].” The existence of multiple effective therapies for PAH may increase the pressure on sponsors to design RCTs which go beyond the “traditional” end points used for approval and focus on “harder” clinical end points to support more definitive claims, increasing confidence that new therapies result in incremental benefit over current drugs. Such Phase III trials based on definitive clinical end points could also serve to validate surrogates, reassuring investigators and clinicians that these end points are indeed appropriate for early-phase clinical testing and managing patients.

Strengths of Surrogate End Points

There are several advantages to using surrogate end points. Surrogates are often based on continuous variables which usually require smaller sample sizes for adequately-powered studies than those required when using dichotomous end points (e.g., dead/alive). Focusing on the mean changes in such measures over time usually reduces sample size requirements even further. For example, most clinical trials in PAH have compared changes in hemodynamics and six-minute walk distance (6MWD) over time between drug and placebo groups, not the absolute values at study end. Usually surrogate end points can be measured sooner than the clinical outcomes they represent, shortening study duration and reducing costs. Measurements such as the six-minute walk test (6MWT) may be serially repeated with minimal discomfort and expense, making them particularly useful for clinical management. Quantitative results from intermediate or surrogate end points may help to rank different therapies in terms of effectiveness, but differences between studies may significantly weaken comparability. Some researchers have proposed using surrogates as “auxiliary end points,” to strengthen the analysis of ultimate clinical end points [14].

Weaknesses of Surrogate End Points

Dependence on surrogate end points may also be a liability. The smaller sample size, temporal immediacy, and narrow focus may increase the chances of missing uncommon, late, or extraneous adverse effects in RCTs, which could overshadow the drug benefits in clinical practice. For example, the short-term, apparently beneficial effects of antiarrhythmics on ventricular ectopy and of inotropes and vasodilators on hemodynamics and exercise function in patients with congestive heart failure (CHF) were trumped by an increased risk of death in the long-term [15–17]. Similarly, it is possible (and perhaps even inevitable) to develop a PAH therapy which will decrease PVR, increase cardiac index (CI), and increase the 6MWD over the short-term, yet also increase the long-term risk of adverse events.

The study population from a RCT is virtually always less heterogeneous than the actual clinical patient population; the effect of therapy on the surrogate in a study may therefore not translate to the definitive end point in clinical practice. The validation of a surrogate end point in a clinical trial does not apply to the patients who would have been excluded from that trial. Also, validation of a surrogate for one class of drugs does not guarantee that it will be applicable for other classes. For example, CI may be a valid surrogate for PAH a therapy that targets the pulmonary vasculature, reducing right ventricular afterload. However, an inotrope without pulmonary vascular effects may have the same (or greater) effect on the surrogate marker (CI) without necessarily impacting on the disease course. In this case, CI may be a valid surrogate for drugs with a particular mechanism (i.e., pulmonary vascular remodeling), but may not be appropriate for drugs with different mechanism (i.e., inotropic support).

Patients with PAH may refuse study procedures, drop out, become sicker, or die in RCTs, all resulting in missing end point data which are non-random. Even with complete assessments, technical difficulties in measurement may be related to clinical outcomes, leading to bias (e.g.,

increased variability in plasma brain natriuretic peptide (BNP) levels in a decompensated patient).

Some clinical trials in PAH have used a range of values of an end point for an inclusion criterion, such as 6MWD. The subsequent truncation of the distribution of the measure in the study sample may lead to regression to the mean [18]. The observed mean change in the variable in the study is then not an unbiased estimate of the true change in the actual population of patients. This prevents the extrapolation of the findings of the study in terms of the surrogate to outcomes in clinical practice.

Lastly, there may not be a linear relationship between the value of a surrogate end point and the clinical implications. For example, a placebo-corrected difference in the change in 6MWD of +40 meters in a RCT with Drug A does not necessarily mean that it is more effective than Drug B, which produces a placebo-corrected difference of +30 meters. It is possible that any 6MWD increment more than +25 meters or an absolute 6MWD after treatment > 380 meters results in improved QOL and survival, metrics by which both Drugs A and B may perform identically. Therefore, the differences in changes in 6MWD in the hypothetical trials may not have clinical significance. One recent study in fact supports such a “threshold effect” [19]. The temptation to compare clinical efficacy using surrogate and intermediate end points should therefore be resisted until the implications of such differences are better understood.

Criteria for a Valid Surrogate End Point

The primary requirement for a surrogate or intermediate end point is reliability. Potential surrogates in PAH, such as cardiopulmonary exercise test (CPET) results, may be highly operator-, protocol-, equipment-, and reader-dependent, lessening their usefulness. A recent study implicated differential variability in hampering the use of CPET in a RCT [20]. Secondly, the surrogate should be in the causal pathway to the ultimate outcome or be closely related to a causal factor (Figure). An intervention should target the disease pathway in which the surrogate lies, and alternative pathways to clinical end points should not exist. Additionally, extensive epidemiologic evidence should link the surrogate variable with outcome. While the strength and reproducibility of the relationship between the surrogate marker and the definitive clinical end point are necessary, “a correlate does not a surrogate make [21].” Markers with powerful correlations with clinical outcomes or other end points may themselves be very poor surrogates of clinical efficacy. Therefore, RCTs of similar and different therapies must show that a quantitative modification of the surrogate is associated with a concomitant modification in the target outcome [22]. Such confirmatory trials must incorporate potential surrogate measures yet be sufficiently powered to show clinically-significant effects of the therapy on ultimate end points. Adequate data satisfying this final most important criterion are severely lacking in PAH, ranking most potential PAH surrogates in the lowest category of validation (i.e., Level 4) in a recent hierarchical scheme (Table 1) [23].

Validating a Surrogate-Guided Therapeutic Approach

One recent observational study used a goal-oriented therapeutic strategy in PAH, employing add-on combination therapy until patients met predefined 6MWD and peak exercise criteria [24]. With this approach, patients had a better overall event-free survival compared to a historical control group. The authors qualified their findings by highlighting the nonrandomized study design, the unvalidated nature of the end points used, the unproven efficacy of the combination therapy employed, and the increased side effects and cost which could result from such a strategy. However, if it were true that a marker (or markers) captured all of the current and future effects of therapy, the optimal drug regimen would indeed be the one which reduced or maintained the surrogate at a low (or high) level.

To confirm the efficacy of such an approach requires a RCT of the surrogate measure(s) to guide therapy and escalation with a clearly-effective established treatment algorithm for incomplete response or disease worsening. Patients would be randomized either to serial surrogate measurements with explicit therapeutic algorithms based on the results or to traditional care; such RCTs of surrogate-end point guided therapy are only recently being performed in CHF [25–27]. Without supportive evidence showing the incremental benefit of using goal-directed therapy, a surrogate end point-guided approach could result in significant overuse or underuse of therapies to the detriment of patients with PAH.

Potential Surrogate and Intermediate End Points in PAH

Hemodynamics

Cardiopulmonary hemodynamic abnormalities not only are defining characteristics of PAH but are also the best established indicators of the severity of illness. The earliest studies of PAH showed that increased right atrial (RA) pressure, decreased CI, and increased mean pulmonary artery (PA) pressure were predictors of death or lung transplantation [28–32]. While these findings were published in an era before effective therapy was available, they have generally (but not universally) been validated in more recent studies of patients at diagnosis or initiation of therapy [19,33–50]. Importantly, PAH patients with more severe cardiac dysfunction or history of heart failure at baseline continue to have a higher risk of death than those with less severe heart failure despite treatment with new therapies [33,34,36,38,41,45,46,48,49].

Improvement in these characteristic hemodynamic abnormalities after therapy has been associated with better survival [19,34,49]. McLaughlin et al. have shown that reductions in RA pressure and mean PA pressure and increases in CI after initiation of intravenous epoprostenol predicted a better survival. Similarly, Provencher et al. and Sitbon et al. demonstrated that the persistence of elevated total pulmonary resistance despite a few months of bosentan or epoprostenol therapy, respectively, was associated with an increased risk of death [19,49]. Opitz et al. showed that lower CI and higher RA pressure and PVR after three months of inhaled iloprost predicted an increased risk of death [48].

One study has confirmed that hemodynamic measurements are valid surrogate end points in PAH. A RCT of continuous intravenous epoprostenol showed significant reductions in mean PA pressure and PVR and increases in CI [11]. This trial also showed significantly improved survival in the epoprostenol group. A subsequent RCT of intravenous epoprostenol in patients with PAH and systemic sclerosis demonstrated hemodynamic changes similar to those of the former study but no survival benefit over 12 weeks [51], suggesting that surrogate adequacy may depend on differences in disease states [46,52]. Other RCTs of prostacyclin analogues have shown smaller or no effects on hemodynamics in PAH and no differences in death, need for transplantation, or initiation of intravenous epoprostenol between drug and placebo groups [12,13,53–56]. Clinical trials of endothelin receptor antagonists and sildenafil have shown differences between drug and placebo in mean changes in hemodynamic measures similar to those found in the early epoprostenol study, but no effects on time to reaching ultimate clinical end points, such as initiation of parenteral therapy, septostomy, transplantation, or death [57–59]. This implies that surrogates which appeared promising in past studies need to be re-evaluated in the current era before widespread use.

Integral to the causal pathway of PAH, hemodynamic measurements meet the epidemiologic criteria for surrogates, but the optimal parameter is not established. Only the first-mentioned RCT of epoprostenol found a concurrent effect on hemodynamics and differences in death or transplantation; that similar hemodynamic effects have not translated to survival differences in more recent RCTs is likely attributable to temporal changes in these studies and the participants. Hemodynamic criteria for “clinical improvement” may be applied to patients after

the initiation of medical therapy for PAH. However, there are no data to support the institution of treatment changes based on the failure to meet certain goals. Strengths of hemodynamic measurements include standardization, availability, and strong biologic plausibility. The invasiveness, discomfort, risk, and expense of right heart catheterization pose significant drawbacks to its use in RCTs and in the clinical management of patients with PAH.

Plasma Biomarkers

Biomarkers play an integral role in the diagnosis and management of more common cardiovascular diseases, but their clinical and investigative use has lagged in PAH. Optimal plasma biomarkers would be causal (or at least closely linked) to the extent of pulmonary vascular disease or degree of right heart dysfunction [10]. Circulating von Willebrand factor (vWF) and other substances are released by the vascular endothelium and are increased in PAH [60–62]. Higher vWF levels at baseline and despite therapy are independently associated with an increased risk of death in PAH [60]. Similarly, a variety of biomarkers tied to cardiac function have been investigated in PAH [38,50,63,64]. The best studied, natriuretic peptides are produced by the cardiac atria and ventricles in response to myocyte stretch and may serve as endocrinologic indicators of heart function. Circulating levels of BNP and N-terminal proBNP (NT-proBNP) are increased in PAH and are closely tied to the causal pathway of disease progression [47,65–71]. Baseline and follow-up levels of BNP and NT-proBNP have been shown to be directly associated with the risk of death and other events in PAH [38,47, 71,72].

One small RCT included BNP measurements [73]. This study randomized patients with PAH to either sildenafil or bosentan. Decreases in BNP levels were seen in both groups, however BNP levels did not differ between drug groups, similar to other end points in this trial. Unfortunately, none of the recently published, placebo-controlled RCTs of new, effective therapies have included early assessment of BNP as a secondary end point, marking significant lost opportunities to advance the science of clinical trials and patient management in PAH.

As BNP and other plasma biomarker levels are assessed by phlebotomy, they may be obtained relatively non-invasively and safely, lending to serial evaluations over time. In addition, samples may be stored locally and transported for central analysis, potentially reducing measurement error in RCTs. The weaknesses of some biomarkers lie in their within-subject variability, which may be substantial, and the potential for therapies to directly affect them (endothelin receptor antagonists and endothelin-1, for example). RCTs demonstrating improvement in outcomes with the use of a biomarker-guided strategy is necessary before incorporating routine BNP or other measures into clinical practice in PAH [25–27].

Cardiac Imaging

Considering the integral role of right ventricular function in the morbidity and mortality of PAH, cardiac imaging parameters might not only respond to effective therapies but also associate with outcome. It is well-known that cardiac morphologic changes can long precede hemodynamic changes in left-sided heart disease. Similarly, imaging with echocardiography, radionuclide angiography, or magnetic resonance imaging (MRI) may capture subtle changes in right-sided cardiac function in PAH before hemodynamic decompensation occurs.

Echocardiography

Studies have shown that a variety of atrial, ventricular, and Doppler parameters have been associated with an increased risk of death or transplantation in patients with PAH [37,38,44, 63,74–76]. While pericardial effusion presence and size are also associated with death or transplantation, this parameter is unlikely to be sufficiently discriminating for research

purposes [32,38,44,45,47,77]. It is notable that the estimate of right ventricular systolic pressure from echocardiography has not been found to be predictive of outcome [32,44,45,74].

In the RCT of intravenous epoprostenol for idiopathic PAH, differences in changes in ventricular and septal morphology were found between the epoprostenol and control groups, corresponding with a difference in survival [78]. A substudy of a RCT of bosentan versus placebo showed significant differences between groups in changes in ventricular morphology, the minimum diameter of the inferior vena cava, and Doppler measurements, including right ventricular ejection time and mitral valve peak velocity [37].

While the meaningfulness of echocardiography as a secondary end point in these RCTs shows the promise of this technique, the incremental prognostic value of echocardiography in monitoring patients after treatment has not been studied. Improvement in hemodynamics, exercise performance, and symptoms are often not reflected in routine echocardiographic measurements. In addition, many of the more specialized measurements are not regularly performed by most clinical echocardiography laboratories.

Strengths of echocardiography include noninvasiveness, plausibility, and incorporation in previous RCTs. Unfortunately, echocardiography is expensive and has varying availability. The operator-dependent nature of these highly technical measurements may impact reliability. In addition, the adequacy of echocardiographic measurements of the right ventricle depends on many patient factors (e.g., body habitus), potentially introducing differential measurement error and bias.

Other Cardiac Imaging Modalities

Radionuclide angiography measures the right ventricular ejection fraction (RVEF) [79], as does MRI, considered to be the “gold-standard” for assessing right atrial and ventricular dimensions, ventricular mass, and ventricular wall motion [80–85]. One study has shown an association between lower baseline RVEF and an increased risk of death, independent of other factors [45]. A single short-term RCT of bosentan and sildenafil used right ventricular mass by MRI as the primary end point [73]. Significant differences between the groups were not seen, although there were changes within groups over the study period. The requirement for radioactive agents for nuclear studies and the more complex technical issues and expense which accompany the use of these modalities to manage patients may limit their utility in clinical practice. However, as the most accurate technique to visualize the right ventricle, MRI may play a more prominent role in future RCTs of PAH.

Exercise Testing

Six Minute Walk Test (6MWT)

The 6MWT is thought to reliably reflect the ability to perform activities of daily living in a quantitative manner in a variety of heart and lung diseases, making this an intermediate end point [86]. A recent report substantiated the impact of changes in 6MWD on health status and QOL [87]. However, 6MWD may also be linked to survival in PAH, making it a surrogate end point as well. Initially considered as a substitute for CPET to gauge maximal oxygen consumption, this deceptively simple evaluation is one of the best-established and most often utilized tests in RCTs and clinical management of PAH. Doebeck et al. have shown that the 6MWT was more metabolically demanding than CPET in PAH, whereas Provencher et al. demonstrated that changes in 6MWD after therapy were closely linked to changes in cardiac function [88,89].

It is biologically plausible that interventions which improve exercise tolerance will improve survival in PAH (as it was in CHF before studies showed otherwise) [16]. Miyamoto et al. found a direct association between 6MWD and survival, independent of other noninvasive measures [90]. Almost all other studies of 6MWD in PAH have confirmed this association [19,35,38,40–42,44,49,50,91]. Sitbon et al. found inverse associations between 6MWD at baseline and after three months of epoprostenol therapy with the risk of death [19]. An absolute 6MWD threshold of < 380 meters at three months after initiating therapy increased the risk of death, as opposed to the change in 6MWD after treatment (an often used end point in RCTs) which was not predictive. In another study, the change in and the absolute 6MWD four months after initiation of bosentan therapy predicted survival [49]. This suggests that absolute, incremental, or threshold values in 6MWD after treatment may reflect long-term outcomes depending on the patient severity of illness or type of therapy instituted.

The RCT of epoprostenol showed a significant difference in the mean changes of 6MWD from baseline and follow-up between the drug and control groups (+47 m) with a corresponding difference in survival [11]. RCTs of prostacyclin analogs, bosentan, sitaxsentan, and sildenafil have also shown statistically significant differences in changes in 6MWD, but no differences in survival, need for lung transplantation, or transition to intravenous epoprostenol [12,13, 53,55,57–59,92], showing that the clinically significant increment in 6MWD may need to be redefined in the current era.

Advantages of the 6MWT include the ability to assess a global and integrated response of systems required for daily functioning, simplicity of performance, low cost, and evidence of reasonable reliability and validity in other chronic heart and lung diseases under standardized conditions [93,94]. Weaknesses of this test include variation in test conduct, a learning effect after repeated testing, variability based on other activities on the day of the testing, and the effect of musculoskeletal conditions on performance. The 6MWT is not within the causal pathway of PAH, but is likely closely-related to causal factors (Figure). The effort-dependent nature of this test may also be problematic, as it may be difficult to adequately mask investigators and patients in studies of drugs with characteristic side effects, such as prostacyclin analogues. If the subject is able to discern the treatment assignment, test performance may be affected, similar to other volitional exercise assessments.

In addition, the inclusion of a permissible range of 6MWDs in the entry criteria of a RCT may truncate the distribution of this variable in the cohort, weakening the assumptions regarding even a validated surrogate measure in terms of clinical outcome [18]. Some have posited a “ceiling effect” specific to six minute walk testing [95]. As in other diseases, it is likely that healthier PAH patients tend to have smaller improvements after interventions than sicker patients. A final issue with the 6MWT is the prejudice that, because of its simplicity, the results must be less reliable or valid than other more invasive, sophisticated, or technical testing. However, this bias is not supported by the existing evidence.

Cardiopulmonary Exercise Testing (CPET)

CPET quantifies the function of the many systems employed in exercise, and therefore CPET could reflect cardiopulmonary benefits of effective therapies in PAH. Studies in other heart and lung diseases have shown relationships between oxygen consumption at peak exercise (VO_2max) and outcome and have employed CPET measures as surrogate end points [96].

The sophisticated nature and sheer volume of the measurements produced by CPET offer an enticing insight into PAH. Test performance in RCT participants with PAH appears reliable [97]. VO_2max and the ventilatory equivalent for carbon dioxide with exercise are abnormal in patients with PAH, and these and other parameters are associated with hemodynamics [20, 45,97–103].

Despite these findings, it is not clear which (if any) of the many parameters are predictive of long term outcome in PAH. Wensel et al. demonstrated that VO_2max corrected for weight as well as systolic and diastolic blood pressure at peak exercise were discriminating predictors of death in 70 patients undergoing initial evaluation for PAH [33]. This study used treadmill testing for most of the subjects and showed a particularly high risk of death or transplantation at one year (32%), exceeding that of other centers during a similar time period and potentially limiting generalizability. These investigators subsequently published a cohort study of patients treated with inhaled iloprost in which VO_2max again predicted time until events [48]. However, in another study of 72 consecutive PAH patients, there was no association between VO_2max and survival, although lower systolic blood pressure and higher ventilatory equivalent for carbon dioxide at peak exercise did predict worse outcome [45].

A RCT of sitaxsentan utilized the change in VO_2max percent predicted as the primary end point [57]. While there were significant beneficial effects of the drug on hemodynamics and 6MWD (confirmed in a second trial) [92], similar benefits were not seen in CPET parameters. Further analysis of this cohort revealed that correlations of VO_2max with other measures differed between sites and changed over each patient's progress through the trial [20]. Similarly, VO_2max did not track with other more established surrogates in another RCT [54]. The lack of response of CPET to therapy which was otherwise effective in terms of more established end points should elicit caution in using CPET to guide treatment decisions or to study new therapies.

The strength of CPET is its biologic plausibility and its usefulness in other diseases. Weaknesses lie in the paucity and inconsistency of epidemiologic and RCT evidence of validity. CPET requires specialized, calibrated equipment and standardized test performance. The presence of a patent foramen ovale or uncorrected cardiac shunt affects test validity. Patients may be unable or unwilling to give a maximal effort on a cycle ergometer or treadmill. Lastly, CPET may not be safe or practical for severely ill patients.

Functional Classification

The NYHA classification of functional status has been slightly altered for PAH in the form of the World Health Organization (WHO) classification (Table 2). Reliability has not been assessed in patients with PAH, however the validity of the NYHA/WHO classification is supported by the association with survival in a variety of cohort studies [19,29,33,34,40,42,46]. In addition, functional class correlates with QOL, underscoring its role as an intermediate end point [104]. Intravenous epoprostenol, endothelin receptor antagonists, and sildenafil improved functional class, forming an important basis for approval of these drugs for PAH [11,51,56–58,92,105]. The persistence of more limited functional class despite treatment predicted worse survival in some studies [19,46] but not in others [34,49].

Strengths of this end point include the convenience and ease of classification. A treatment-induced change in functional class translates to patient benefit, even without a confirmation of surrogacy for survival. Weaknesses lie in the self-report and interpretive role patients and investigators have in assignment. RCTs in which one or the other party are unmasked by side effects may result in bias. In addition, the blunt nature of this endpoint may mean that this classification is poorly discriminating and that subtle changes in clinical status will not be detected.

Quality of Life (QOL)

The improvement of QOL and health-related QOL are significant goals in the treatment of PAH. The assessment of the success of meeting these goals, however, is not straightforward. Although we look towards functional tests, such as the 6MWT, to provide this information, it

is unlikely that a single physiologic test can capture all of the components that enter into the complex construct of satisfaction of the needs and desires of daily living. Investigators have therefore “asked” PAH patients in epidemiologic studies and RCTs about the components which comprise QOL, using instruments such as the Medical Outcome Study Short Form-36 (SF-36). More specific questionnaires focus on certain characteristics of the disease process under study, such as the Minnesota Living with Heart Failure (MLHF) questionnaire. These instruments are judged not only by content and criterion validity (reasonableness of the questions and correlation of the scale with another “gold standard” measure accepted in the field) but also convergent and discriminant validity (responses track with other assessments of similar constructs and not with unrelated variables). Lastly, change of the measurement in response to effective treatment testifies to the validity of the instrument.

Recent studies have initiated this validation process in PAH [87,104,106]. Taichman et al. examined the SF-36 in reference to other components of disease activity in PAH [104]. Higher WHO functional class and lower 6MWD were associated with a more affected (worse) physical activity summary score (convergent validity). On the other hand, hemodynamics had no association with either the physical or mental component scores (discriminant validity). Investigators have shown similar correlations between the physical subscore of the modified MLHF questionnaire with functional class, 6MWD, RA pressure and CI [87,107]. The MLHF score was also associated with event-free survival [107]. Chua et al. have shown that improvements in 6MWD over time translate to improvements in QOL measured by the SF-36 [87].

One group has taken on the daunting but necessary task of creating a PAH-specific health-related QOL instrument, the Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR) [108]. It appears that this instrument has good test-retest reliability and content, convergent, and discriminant validity. The responsiveness of this instrument to interventions, however, has not yet been assessed.

Strengths of these QOL instruments include content and criterion validity in PAH. Weaknesses include the few studies of convergent and discriminant validity and the limited use in PAH. The symptom constellation and treatment strategies of PAH likely make QOL questionnaires designed for other health conditions less sensitive and specific for the issues which face these patients.

Conclusions

The use of surrogate and intermediate end points is deeply ingrained in the clinical study of PAH, however few definitive, evidence-based statements may be made. New therapies for patients with PAH need to show beneficial hemodynamic effects in early phase trials. There may be a single hemodynamic parameter which is the “best” surrogate or different classes of therapies may warrant different surrogate measures. The literature suggests that an optimal time point for reassessment is at least three to four months after initiation of therapy, but earlier assessment of an end point might also predict treatment success. This could certainly vary based on the type of therapy under study. While beneficial drug effects on hemodynamic surrogates have not sufficed for regulatory approval, the validity of hemodynamics as surrogate end points should be reconsidered based on their merit in PAH, apart from their inadequacy in CHF.

While echocardiography and BNP hold promise as non-invasive surrogates, they are not ready for use as primary end points. Multiple long-term RCTs should incorporate BNP to validate this biomarker against definitive clinical end points with various drug classes. A RCT of BNP-

guided medical management would assist clinicians in deciding if monitoring this biomarker benefits patients in clinical practice.

Differences in 6MWD between therapy and placebo groups over the short term may or may not be clinically meaningful. Future studies should better define the parameter(s) and increment of the 6MWD which reflect patient-perceived improved function. It would also be vitally important to extrapolate such findings to differences in event-free survival. The efficacy of therapies with prominent side effects and potential unmasking should be substantiated by non-effort-dependent tests. CPET measures should not be used as primary outcomes in RCTs until further examination of these metrics has been completed.

Functional status and QOL are important end points to include in RCTs of PAH therapy, as they reflect patient well-being. Longer-term trials with these end points would prove that such beneficial effects were durable. Validation of commonly used instruments and the further development of instruments specific for PAH (such as CAMPHOR) are necessary to achieve the goal of designing therapeutic approaches which definitively improve the QOL of patients with PAH.

The scientific discoveries and therapeutic advances in PAH require similar developments in clinical trial design. Many innovative technologies are changing how patients can be profiled and respond to therapy, including genetics, genomics, proteomics, metabolomics, and imaging. While these new techniques offer the opportunity to describe patients in more detail than ever, validation of the clinical utility of these technologies is less “high-tech” but no less difficult or important than the process of invention. Seemingly established surrogate and intermediate end points, such as hemodynamics and exercise testing, remain incompletely validated and poorly described in PAH; these priority candidates should be vetted before others are considered.

Therefore, future RCTs in PAH should abide by the recommendations and goals of the National Institutes of Health workshop on surrogate end points [109]. Progress in this area can only be made through incorporation of these measures into clinical trials powered to detect differences in time to clinically-important end points. This makes for more expensive trials of longer duration, but this is necessary to gain confidence in these tests. Prospective planning for meta-analyses across RCTs of new and combination therapies would provide additional opportunities for sponsors and investigators to advance the science of clinical trials in PAH, while not distracting from the main hypotheses under study [110,111]. Such investments in the conduct of RCTs in PAH will pay great dividends in the future study and care of patients with PAH.

Acknowledgements

The authors are grateful for the administrative assistance of Tracey Huger and for the thoughtful review of the manuscript by Norman L. Stockbridge, M.D., Ph.D.

References

1. Hoeper MM, Oudiz RJ, Peacock A, et al. End points and clinical trial designs in pulmonary arterial hypertension: clinical and regulatory perspectives. *J Am Coll Cardiol* 2004;43:48S–55S. [PubMed: 15194178]
2. Peacock A, Naeije R, Galie N, Reeves JT. End points in pulmonary arterial hypertension: the way forward. *Eur Respir J* 2004;23:947–953. [PubMed: 15219011]
3. Kawut SM, Palevsky HI. Surrogate end points for pulmonary arterial hypertension. *Am Heart J* 2004;148:559–565. [PubMed: 15459583]
4. Kawut SM, Palevsky HI. New answers raise new questions in pulmonary arterial hypertension. *Eur Respir J* 2004;23:799–801. [PubMed: 15218988]

5. Rich S. The current treatment of pulmonary arterial hypertension: time to redefine success. *Chest* 2006;130:1198–1202. [PubMed: 17035456]
6. Roberts K, Preston I, Hill NS. Pulmonary hypertension trials: current end points are flawed, but what are the alternatives? *Chest* 2006;130:934–936. [PubMed: 17035418]
7. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* 2001;69:89–95. [PubMed: 11240971]
8. 57 FR 58958 as amended at 64 FR 402. 1992. New drug, antibiotic, and biological drug product regulations; accelerated approval; final rule.
9. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. *Stat Med* 1989;8:431–440. [PubMed: 2727467]
10. Temple R. Are surrogate markers adequate to assess cardiovascular disease drugs? *JAMA* 1999;282:790–795. [PubMed: 10463719]
11. Barst RJ, Rubin LJ, Long WA, et al. A comparison of continuous intravenous epoprostenol with conventional therapy for primary pulmonary hypertension. *N Engl J Med* 1996;334:296–302. [PubMed: 8532025]
12. Olschewski H, Simonneau G, Galie N, et al. Inhaled iloprost for severe pulmonary hypertension. *N Engl J Med* 2002;347:322–329. [PubMed: 12151469]
13. Simonneau G, Barst RJ, Galie N, et al. Continuous subcutaneous infusion of treprostinil, a prostacyclin analogue, in patients with pulmonary arterial hypertension: a double-blind, randomized, placebo-controlled trial. *Am J Respir Crit Care Med* 2002;165:800–804. [PubMed: 11897647]
14. Fleming TR, Prentice RL, Pepe MS, Glidden D. Surrogate and auxiliary endpoints in clinical trials, with potential applications in cancer and AIDS research. *Stat Med* 1994;13:955–968. [PubMed: 8047747]
15. Echt DS, Liebson PR, Mitchell LB, et al. Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial. *N Engl J Med* 1991;324:781–788. [PubMed: 1900101]
16. Cohn JN, Goldstein SO, Greenberg BH, et al. A dose-dependent increase in mortality with vesnarinone among patients with severe heart failure. Vesnarinone Trial Investigators. *N Engl J Med* 1998;339:1810–1816. [PubMed: 9854116]
17. Califf RM, Adams KF, McKenna WJ, et al. A randomized controlled trial of epoprostenol therapy for severe congestive heart failure: The Flolan International Randomized Survival Trial (FIRST). *Am Heart J* 1997;134:44–54. [PubMed: 9266782]
18. Wittes J, Lakatos E, Probstfield J. Surrogate endpoints in clinical trials: cardiovascular diseases. *Stat Med* 1989;8:415–425. [PubMed: 2727465]
19. 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. [PubMed: 12204511]
20. Oudiz RJ, Barst RJ, Hansen JE, et al. Cardiopulmonary exercise testing and six-minute walk correlations in pulmonary arterial hypertension. *Am J Cardiol* 2006;97:123–126. [PubMed: 16377296]
21. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med* 1996;125:605–613. [PubMed: 8815760]
22. Bucher HC, Guyatt GH, Cook DJ, Holbrook A, McAlister FA. Users' guides to the medical literature: XIX. Applying clinical trial results. A. How to use an article measuring the effect of an intervention on surrogate end points. Evidence-Based Medicine Working Group. *JAMA* 1999;282:771–778. [PubMed: 10463714]
23. Fleming TR. Surrogate endpoints and FDA's accelerated approval process. *Health Aff (Millwood)* 2005;24:67–78. [PubMed: 15647217]
24. Hoeper MM, Markevych I, Spiekerkoetter E, Welte T, Niedermeyer J. Goal-oriented treatment and combination therapy for pulmonary arterial hypertension. *Eur Respir J* 2005;26:858–863. [PubMed: 16264047]
25. Lainchbury JG, Troughton RW, Frampton CM, et al. NTproBNP-guided drug treatment for chronic heart failure: design and methods in the "BATTLESCARRED" trial. *Eur J Heart Fail* 2006;8:532–538. [PubMed: 16829189]

26. Shah MR, Claise KA, Bowers MT, et al. Testing new targets of therapy in advanced heart failure: the design and rationale of the Strategies for Tailoring Advanced Heart Failure Regimens in the Outpatient Setting: BRAIN Natriuretic Peptide Versus the Clinical Congestion Score (STARBRITE) trial. *Am Heart J* 2005;150:893–898. [PubMed: 16290955]
27. Troughton RW, Frampton CM, Yandle TG, Espiner EA, Nicholls MG, Richards AM. Treatment of heart failure guided by plasma aminoterminal brain natriuretic peptide (N-BNP) concentrations. *Lancet* 2000;355:1126–1130. [PubMed: 10791374]
28. Sandoval J, Bauerle O, Palomar A, et al. Survival in primary pulmonary hypertension. *Circulation* 1994;89:1733–1744. [PubMed: 8149539]
29. D'Alonzo GE, Barst RJ, Ayres SM, et al. Survival in patients with primary pulmonary hypertension. Results from a national prospective registry. *Ann Intern Med* 1991;115:343–349. [PubMed: 1863023]
30. Kanemoto N. Natural history of pulmonary hemodynamics in primary pulmonary hypertension. *Am Heart J* 1987;114:407–413. [PubMed: 3604899]
31. Glanville AR, Burke CM, Theodore J, Robin ED. Primary pulmonary hypertension: length of survival in patients referred for heart-lung transplantation. *Chest* 1987;91:675–681. [PubMed: 3105966]
32. Eysmann SB, Palevsky HI, Reichel N, Hackney K, Douglas PS. Two-dimensional and Doppler-echocardiographic and cardiac catheterization correlates of survival in primary pulmonary hypertension. *Circulation* 1989;80:353–360. [PubMed: 2752562]
33. Wensel R, Opitz CF, Anker SD, et al. Assessment of survival in patients with primary pulmonary hypertension: importance of cardiopulmonary exercise testing. *Circulation* 2002;106:319–324. [PubMed: 12119247]
34. McLaughlin VV, Shillington A, Rich S. Survival in primary pulmonary hypertension: the impact of epoprostenol therapy. *Circulation* 2002;106:1477–1482. [PubMed: 12234951]
35. Ewert R, Wensel R, Opitz C, et al. Prognosis in patients with primary pulmonary hypertension awaiting lung transplantation. *Transplant Proc* 2001;33:3574–3575. [PubMed: 11750520]
36. Bossone E, Paciocco G, Iarussi D, et al. The prognostic role of the ECG in primary pulmonary hypertension. *Chest* 2002;121:513–518. [PubMed: 11834666]
37. Galie N, Hinderliter AL, Torbicki A, et al. Effects of the oral endothelin-receptor antagonist bosentan on echocardiographic and doppler measures in patients with pulmonary arterial hypertension. *J Am Coll Cardiol* 2003;41:1380–1386. [PubMed: 12706935]
38. Fijalkowska A, Kurzyna M, Torbicki A, et al. Serum N-terminal brain natriuretic peptide as a prognostic parameter in patients with pulmonary hypertension. *Chest* 2006;129:1313–1321. [PubMed: 16685024]
39. Nunes H, Humbert M, Sitbon O, et al. Prognostic factors for survival in human immunodeficiency virus-associated pulmonary arterial hypertension. *Am J Respir Crit Care Med* 2003;167:1433–1439. [PubMed: 12615632]
40. Williams MH, Das C, Handler CE, et al. Systemic sclerosis associated pulmonary hypertension: improved survival in the current era. *Heart* 2006;92:926–932. [PubMed: 16339813]
41. Mahapatra S, Nishimura RA, Sorajja P, Cha S, McGoon MD. Relationship of pulmonary arterial capacitance and mortality in idiopathic pulmonary arterial hypertension. *J Am Coll Cardiol* 2006;47:799–803. [PubMed: 16487848]
42. 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. [PubMed: 15684287]
43. Frank H, Mlczech J, Huber K, Schuster E, Gurtner HP, Kneussl M. The effect of anticoagulant therapy in primary and anorectic drug-induced pulmonary hypertension. *Chest* 1997;112:714–721. [PubMed: 9315805]
44. Raymond RJ, Hinderliter AL, Willis PW, et al. Echocardiographic predictors of adverse outcomes in primary pulmonary hypertension. *J Am Coll Cardiol* 2002;39:1214–1219. [PubMed: 11923049]
45. Kawut SM, Horn EM, Berekashvili KK, et al. New predictors of outcome in idiopathic pulmonary arterial hypertension. *Am J Cardiol* 2005;95:199–203. [PubMed: 15642552]
46. Kuhn KP, Byrne DW, Arbogast PG, Doyle TP, Loyd JE, Robbins IM. Outcome in 91 consecutive patients with pulmonary arterial hypertension receiving epoprostenol. *Am J Respir Crit Care Med* 2003;167:580–586. [PubMed: 12446266]

47. Nagaya N, Nishikimi T, Uematsu M, et al. Plasma brain natriuretic peptide as a prognostic indicator in patients with primary pulmonary hypertension. *Circulation* 2000;102:865–870. [PubMed: 10952954]
48. Opitz CF, Wensel R, Winkler J, et al. Clinical efficacy and survival with first-line inhaled iloprost therapy in patients with idiopathic pulmonary arterial hypertension. *Eur Heart J* 2005;26:1895–1902. [PubMed: 15888496]
49. Provencher S, Sitbon O, Humbert M, Cabrol S, Jais X, Simonneau G. Long-term outcome with first-line bosentan therapy in idiopathic pulmonary arterial hypertension. *Eur Heart J* 2006;27:589–595. [PubMed: 16431875]
50. Torbicki A, Kurzyna M, Kuca P, et al. Detectable serum cardiac troponin T as a marker of poor prognosis among patients with chronic precapillary pulmonary hypertension. *Circulation* 2003;108:844–848. [PubMed: 12900346]
51. Badesch DB, Tapon VF, McGoon MD, et al. Continuous intravenous epoprostenol for pulmonary hypertension due to the scleroderma spectrum of disease. A randomized, controlled trial. *Ann Intern Med* 2000;132:425–434. [PubMed: 10733441]
52. Kawut SM, Taichman DB, Archer-Chicko CL, Palevsky HI, Kimmel SE. Hemodynamics and survival in patients with pulmonary arterial hypertension related to systemic sclerosis. *Chest* 2003;123:344–350. [PubMed: 12576350]
53. Galie N, Humbert M, Vachiery JL, et al. Effects of beraprost sodium, an oral prostacyclin analogue, in patients with pulmonary arterial hypertension: a randomized, double-blind, placebo-controlled trial. *J Am Coll Cardiol* 2002;39:1496–1502. [PubMed: 11985913]
54. Barst RJ, McGoon M, McLaughlin V, et al. Beraprost therapy for pulmonary arterial hypertension. *J Am Coll Cardiol* 2003;41:2119–2125. [PubMed: 12821234]
55. McLaughlin VV, Gaine SP, Barst RJ, et al. Efficacy and safety of treprostinil: an epoprostenol analog for primary pulmonary hypertension. *J Cardiovasc Pharmacol* 2003;41:293–299. [PubMed: 12548091]
56. Rubin LJ, Mendoza J, Hood M, et al. Treatment of primary pulmonary hypertension with continuous intravenous prostacyclin (epoprostenol). Results of a randomized trial. *Ann Intern Med* 1990;112:485–491. [PubMed: 2107780]
57. Barst RJ, Langleben D, Frost A, et al. Sitaxsentan therapy for pulmonary arterial hypertension. *Am J Respir Crit Care Med* 2004;169:441–447. [PubMed: 14630619]
58. Galie N, Ghofrani HA, Torbicki A, et al. Sildenafil citrate therapy for pulmonary arterial hypertension. *N Engl J Med* 2005;353:2148–2157. [PubMed: 16291984]
59. Channick RN, Simonneau G, Sitbon O, et al. Effects of the dual endothelin-receptor antagonist bosentan in patients with pulmonary hypertension: a randomised placebo-controlled study. *Lancet* 2001;358:1119–1123. [PubMed: 11597664]
60. Kawut SM, Horn EM, Berekashvili KK, Widlitz AC, Rosenzweig EB, Barst RJ. von Willebrand factor independently predicts long-term survival in patients with pulmonary arterial hypertension. *Chest* 2005;128:2355–2362. [PubMed: 16236894]
61. Lopes AA, Maeda NY. Circulating von Willebrand factor antigen as a predictor of short-term prognosis in pulmonary hypertension. *Chest* 1998;114:1276–1282. [PubMed: 9824001]
62. Veyradier A, Nishikubo T, Humbert M, et al. Improvement of von Willebrand factor proteolysis after prostacyclin infusion in severe pulmonary arterial hypertension. *Circulation* 2000;102:2460–2462. [PubMed: 11076816]
63. Nagaya N, Uematsu M, Satoh T, et al. Serum uric acid levels correlate with the severity and the mortality of primary pulmonary hypertension. *Am J Respir Crit Care Med* 1999;160:487–492. [PubMed: 10430718]
64. Bendayan D, Shitrit D, Ygla M, Huerta M, Fink G, Kramer MR. Hyperuricemia as a prognostic factor in pulmonary arterial hypertension. *Respir Med* 2003;97:130–133. [PubMed: 12587962]
65. Nagaya N, Nishikimi T, Okano Y, et al. Plasma brain natriuretic peptide levels increase in proportion to the extent of right ventricular dysfunction in pulmonary hypertension. *J Am Coll Cardiol* 1998;31:202–208. [PubMed: 9426041]

66. Williams MH, Handler CE, Akram R, et al. Role of N-terminal brain natriuretic peptide (N-TproBNP) in scleroderma-associated pulmonary arterial hypertension. *Eur Heart J* 2006;27:1485–1494. [PubMed: 16682379]
67. Leuchte HH, Holzapfel M, Baumgartner RA, et al. Clinical significance of brain natriuretic peptide in primary pulmonary hypertension. *J Am Coll Cardiol* 2004;43:764–770. [PubMed: 14998614]
68. Souza R, Bogossian HB, Humbert M, et al. N-terminal-pro-brain natriuretic peptide as a haemodynamic marker in idiopathic pulmonary arterial hypertension. *Eur Respir J* 2005;25:509–513. [PubMed: 15738296]
69. Souza R, Jardim C, Julio Cesar Fernandes C, Silveira Lapa M, Rabelo R, Humbert M. NT-proBNP as a tool to stratify disease severity in pulmonary arterial hypertension. *Respir Med*. 2006
70. Mukerjee D, Yap LB, Holmes AM, et al. Significance of plasma N-terminal pro-brain natriuretic peptide in patients with systemic sclerosis-related pulmonary arterial hypertension. *Respir Med* 2003;97:1230–1236. [PubMed: 14635979]
71. Andreassen AK, Wergeland R, Simonsen S, Geiran O, Guevara C, Ueland T. N-terminal pro-B-type natriuretic peptide as an indicator of disease severity in a heterogeneous group of patients with chronic precapillary pulmonary hypertension. *Am J Cardiol* 2006;98:525–529. [PubMed: 16893710]
72. Park MH, Scott RL, Uber PA, Ventura HO, Mehra MR. Usefulness of B-type natriuretic peptide as a predictor of treatment outcome in pulmonary arterial hypertension. *Congest Heart Fail* 2004;10:221–225. [PubMed: 15470298]
73. Wilkins MR, Paul GA, Strange JW, et al. Sildenafil versus endothelin receptor antagonist for pulmonary hypertension (SERAPH) study. *Am J Respir Crit Care Med* 2005;171:1292–1297. [PubMed: 15750042]
74. Yeo TC, Dujardin KS, Tei C, Mahoney DW, McGoon MD, Seward JB. Value of a Doppler-derived index combining systolic and diastolic time intervals in predicting outcome in primary pulmonary hypertension. *Am J Cardiol* 1998;81:1157–1161. [PubMed: 9605059]
75. Bustamante-Labarta M, Perrone S, De La Fuente RL, et al. Right atrial size and tricuspid regurgitation severity predict mortality or transplantation in primary pulmonary hypertension. *J Am Soc Echocardiogr* 2002;15:1160–1164. [PubMed: 12411899]
76. Forfia PR, Fisher MR, Mathai SC, et al. Tricuspid annular displacement predicts survival in pulmonary hypertension. *Am J Respir Crit Care Med* 2006;174:1034–1041. [PubMed: 16888289]
77. Hinderliter AL, Willis PW, Long W, et al. Frequency and prognostic significance of pericardial effusion in primary pulmonary hypertension. *Am J Cardiol* 1999;84:481–484. [PubMed: 10468096]
78. Hinderliter AL, Willis PW, Barst RJ, et al. Effects of long-term infusion of prostacyclin (epoprostenol) on echocardiographic measures of right ventricular structure and function in primary pulmonary hypertension. Primary Pulmonary Hypertension Study Group. *Circulation* 1997;95:1479–1486. [PubMed: 9118516]
79. Nichols K, Saouaf R, Ababneh AA, et al. Validation of SPECT equilibrium radionuclide angiographic right ventricular parameters by cardiac magnetic resonance imaging. *J Nucl Cardiol* 2002;9:153–160. [PubMed: 11986559]
80. Saba TS, Foster J, Cockburn M, Cowan M, Peacock AJ. Ventricular mass index using magnetic resonance imaging accurately estimates pulmonary artery pressure. *Eur Respir J* 2002;20:1519–1524. [PubMed: 12503713]
81. Katz J, Whang J, Boxt LM, Barst RJ. Estimation of right ventricular mass in normal subjects and in patients with primary pulmonary hypertension by nuclear magnetic resonance imaging. *J Am Coll Cardiol* 1993;21:1475–1481. [PubMed: 8473659]
82. Roeleveld RJ, Vonk-Noordegraaf A, Marcus JT, et al. Effects of epoprostenol on right ventricular hypertrophy and dilatation in pulmonary hypertension. *Chest* 2004;125:572–579. [PubMed: 14769740]
83. Blyth KG, Groenning BA, Martin TN, et al. Contrast enhanced-cardiovascular magnetic resonance imaging in patients with pulmonary hypertension. *Eur Heart J* 2005;26:1993–1999. [PubMed: 15899927]
84. Boxt LM, Katz J, Kolb T, Czegledy FP, Barst RJ. Direct quantitation of right and left ventricular volumes with nuclear magnetic resonance imaging in patients with primary pulmonary hypertension. *J Am Coll Cardiol* 1992;19:1508–1515. [PubMed: 1593046]

85. Hoepfer MM, Tongers J, Leppert A, Baus S, Maier R, Lotz J. Evaluation of right ventricular performance with a right ventricular ejection fraction thermodilution catheter and MRI in patients with pulmonary hypertension. *Chest* 2001;120:502–507. [PubMed: 11502650]
86. Sciruba F, Criner GJ, Lee SM, et al. Six-minute walk distance in chronic obstructive pulmonary disease: reproducibility and effect of walking course layout and length. *Am J Respir Crit Care Med* 2003;167:1522–1527. [PubMed: 12615634]
87. 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;36:705–710. [PubMed: 17040356]
88. Deboeck G, Niset G, Vachier JL, Moraine JJ, Naeije R. Physiological response to the six-minute walk test in pulmonary arterial hypertension. *Eur Respir J* 2005;26:667–672. [PubMed: 16204599]
89. Provencher S, Chemla D, Herve P, Sitbon O, Humbert M, Simonneau G. Heart rate responses during the 6-minute walk test in pulmonary arterial hypertension. *Eur Respir J* 2006;27:114–120. [PubMed: 16387943]
90. Miyamoto S, Nagaya N, Satoh T, et al. Clinical correlates and prognostic significance of six-minute walk test in patients with primary pulmonary hypertension. Comparison with cardiopulmonary exercise testing. *Am J Respir Crit Care Med* 2000;161:487–492. [PubMed: 10673190]
91. Paciocco G, Martinez FJ, Bossone E, Pielsticker E, Gillespie B, Rubenfire M. Oxygen desaturation on the six-minute walk test and mortality in untreated primary pulmonary hypertension. *Eur Respir J* 2001;17:647–652. [PubMed: 11401059]
92. Barst RJ, Langleben D, Badesch D, et al. Treatment of pulmonary arterial hypertension with the selective endothelin-A receptor antagonist sitaxsentan. *J Am Coll Cardiol* 2006;47:2049–2056. [PubMed: 16697324]
93. ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med* 2002;166:111–117. [PubMed: 12091180]
94. Solway S, Brooks D, Lacasse Y, Thomas S. A qualitative systematic overview of the measurement properties of functional walk tests used in the cardiorespiratory domain. *Chest* 2001;119:256–270. [PubMed: 11157613]
95. Frost AE, Langleben D, Oudiz R, et al. The 6-min walk test (6MW) as an efficacy endpoint in pulmonary arterial hypertension clinical trials: demonstration of a ceiling effect. *Vascul Pharmacol* 2005;43:36–39. [PubMed: 15890561]
96. ATS/ACCP statement on cardiopulmonary exercise testing. *Am J Respir Crit Care Med* 2003;167:211–277. [PubMed: 12524257]
97. Hansen JE, Sun XG, Yasunobu Y, et al. Reproducibility of cardiopulmonary exercise measurements in patients with pulmonary arterial hypertension. *Chest* 2004;126:816–824. [PubMed: 15364761]
98. Sun XG, Hansen JE, Oudiz RJ, Wasserman K. Gas exchange detection of exercise-induced right-to-left shunt in patients with primary pulmonary hypertension. *Circulation* 2002;105:54–60. [PubMed: 11772876]
99. Sun XG, Hansen JE, Oudiz RJ, Wasserman K. Exercise pathophysiology in patients with primary pulmonary hypertension. *Circulation* 2001;104:429–435. [PubMed: 11468205]
100. Rhodes J, Barst RJ, Garofano RP, Thoele DG, Gersony WM. Hemodynamic correlates of exercise function in patients with primary pulmonary hypertension. *J Am Coll Cardiol* 1991;18:1738–1744. [PubMed: 1960322]
101. Oudiz RJ. The role of exercise testing in the management of pulmonary arterial hypertension. *Semin Respir Crit Care Med* 2005;26:379–384. [PubMed: 16121314]
102. Yasunobu Y, Oudiz RJ, Sun XG, Hansen JE, Wasserman K. End-tidal PCO₂ abnormality and exercise limitation in patients with primary pulmonary hypertension. *Chest* 2005;127:1637–1646. [PubMed: 15888840]
103. Yetman AT, Taylor AL, Doran A, Ivy DD. Utility of cardiopulmonary stress testing in assessing disease severity in children with pulmonary arterial hypertension. *Am J Cardiol* 2005;95:697–699. [PubMed: 15721127]
104. Taichman DB, Shin J, Hud L, et al. Health-related quality of life in patients with pulmonary arterial hypertension. *Respir Res* 2005;6:92. [PubMed: 16092961]
105. Rubin LJ, Badesch DB, Barst RJ, et al. Bosentan therapy for pulmonary arterial hypertension. *N Engl J Med* 2002;346:896–903. [PubMed: 11907289]

106. Shafazand S, Goldstein MK, Doyle RL, Hlatky MA, Gould MK. Health-related quality of life in patients with pulmonary arterial hypertension. *Chest* 2004;126:1452–1459. [PubMed: 15539712]
107. 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. [PubMed: 16707511]
108. McKenna SP, Doughty N, Meads DM, Doward LC, Pepke-Zaba J. The Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR): a measure of health-related quality of life and quality of life for patients with pulmonary hypertension. *Qual Life Res* 2006;15:103–115. [PubMed: 16411035]
109. De Gruttola VG, Clax P, DeMets DL, et al. Considerations in the evaluation of surrogate endpoints in clinical trials. summary of a National Institutes of Health workshop. *Control Clin Trials* 2001;22:485–502. [PubMed: 11578783]
110. Buyse M, Molenberghs G, Burzykowski T, Renard D, Geys H. The validation of surrogate endpoints in meta-analyses of randomized experiments. *Biostatistics* 2000;1:49–67. [PubMed: 12933525]
111. Daniels MJ, Hughes MD. Meta-analysis for the evaluation of potential surrogate markers. *Stat Med* 1997;16:1965–1982. [PubMed: 9304767]

Figure A.

The intervention and surrogate act in the single causal pathway of PAH to morbidity or mortality. This setting has the greatest potential for the surrogate end point to be valid. Situations in which surrogates may fail include: B. The surrogate is not in the causal pathway of PAH. C. Of several causal PAH disease pathways to morbidity and mortality, the intervention affects only the pathway mediated through the surrogate. D. The surrogate is not in the pathway of the intervention's effect or is insensitive to its effect. E. The intervention has mechanisms of action (either beneficial or adverse) independent of the disease process of PAH. Dotted lines = mechanisms of action that might exist. Adapted from Fleming TR and DeMets

DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med* 1996;125:605–613.

Table 1

Hierarchy for Outcome Measures in Pulmonary Arterial Hypertension. Adapted from Fleming TR. Surrogate endpoints and FDA's accelerated approval process. *Health Affairs* 2005;24:67–78

Level	Definition	Details	Examples
I	True clinical efficacy end point	Outcome directly reflects tangible benefit	Hospitalization
II	Validated surrogate end point	Outcome not reflecting tangible benefit but which reliably predicts the level of benefit	
III	Nonvalidated surrogate end point	1) Considerable evidence that intervention effect on the measure will accurately represent the intervention's effect on the predominant mechanism through which the disease process induces clinical events; 2) Considerable clinical evidence that the experimental intervention does not have important adverse effects on the clinical end point not captured by the surrogate end point; 3) Statistical analyses suggest that the net effect of the intervention on the true clinical-efficacy end point is consistent with what would be predicted by the level of effect on the measure; and 4) Targeted effect on the outcome measure is sufficiently strong and durable that, based on #1–3, this is reasonably likely to predict <u>meaningful clinical benefit</u>	Cardiac index, Pulmonary vascular resistance, Distance walked in six minutes
IV	Measure of biologic activity	Treatment effects on measure shows that intervention is biologically active, but without sufficient evidence of translation to clinical benefit	Brain natriuretic peptide

Table 2
World Health Organization Functional Classification in Pulmonary Arterial Hypertension

Class	
I	No limitation of physical activity. Ordinary physical activity does not cause undue dyspnea or fatigue, chest pain, or near syncope.
II	Slight limitation of physical activity. Patient is comfortable at rest. Ordinary physical activity causes undue dyspnea or fatigue, chest pain, or near syncope.
III	Marked limitation of physical activity. Patient is comfortable at rest. Less than ordinary activity causes undue dyspnea or fatigue, chest pain, or near syncope.
IV	Inability to carry out any physical activity without symptoms. Patient manifests signs of right-heart failure. Dyspnea and/or fatigue may be present at rest. Discomfort is increased by any physical activity.