

EDITORIAL

New answers raise new questions in pulmonary arterial hypertension

S.M. Kawut*, H.I. Palevsky[#]

Research and treatment of pulmonary vascular diseases pose distinct challenges. Despite these hurdles (or perhaps because of them), a cadre of basic scientists, geneticists, physiologists, clinical researchers and clinicians have dedicated themselves to the study and care of patients with pulmonary vascular disease. The *European Respiratory Journal* (ERJ) series "Advances in Pathobiology, Diagnosis, and Treatment of Pulmonary Hypertension" [1–10] comprehensively categorises the achievements of these individuals and concludes this month with a fitting finale. PEACOCK *et al.* [11] (this issue of the ERJ) and HOEPER *et al.* [12] have written statements from recent conferences focused on the methodology of randomised controlled trials (RCTs) in pulmonary arterial hypertension (PAH); until recently, a seemingly quixotic discussion. As with all important research, these documents potentially raise more questions than they answer.

The availability of bosentan (an endothelin-1 receptor antagonist) and other safe, effective therapies for PAH led the working groups to re-evaluate the ethical and scientific merit of using untreated placebo-control arms in RCTs [2, 13–15]. This is a very controversial area faced by researchers in other conditions with pre-existing effective therapies, such as congestive heart failure (CHF) [16–19]. In general, it is considered to be ethical to perform such trials if: 1) the patient is fully informed about the alternatives; and 2) there is minimal risk of irreversible or severe morbidity, or mortality, posed by withholding effective therapy.

While the former concern is relatively straightforward, the working groups recognised the difficulty in defining the boundaries of the latter. As the success of future placebo-controlled trials in this rare disease will depend on patients' and physicians' willingness to enrol in such trials, future study designs will have to be acceptable to both groups. Such acceptability may be important to investigate before initiating new studies [20]. Researchers will also have to satisfy local and central regulatory requirements if they are to delay clinically-proven therapy during the conduct of RCTs with placebo.

Could other study designs suffice? The working-group statements considered performing future placebo-controlled studies in PAH with concurrent clinically-proven therapy, a standard approach in CHF which minimises ethical dilemmas. Scientifically, it makes sense to test new therapies with distinct mechanisms in the setting of established treatment, as clinicians and patients demand new therapies that provide incremental improvement over the current standard of care.

A significant problem with this approach is that a therapy which may be effective when used singularly may not appear effective in combination with other therapies because of an interaction. This could potentially shelve a useful drug for patients who are intolerant of established therapies. Co-administration with other therapy may also require increased sample size and longer trial duration in order to achieve the same precision in results. Lastly, this approach is not appropriate for drugs with mechanisms similar to those of current therapies.

The working groups point out the significant difficulties with active-control equivalency (or noninferiority) studies in PAH. While a noninferiority study without a placebo arm can answer whether the effect of a new drug is "not worse" than that of an established drug, there is no way to prove that the active control (and subsequently the new drug) had a benefit in that particular study (assay sensitivity) [16]. We depend entirely on evidence from other studies in identical populations that the active drug is effective. This assumption is very difficult to meet in common diseases with large clinical trials, much less in the rare disease of PAH. Other limitations, including increased sample size, confinement to the use of previous endpoints and inclusion criteria, establishment of a non-inferiority margin, problems with "per protocol" analysis and the stance of regulatory bodies, made the working groups hesitant to endorse this study design for PAH trials [19, 21]. Clearly, the design of future RCTs in PAH will have to meet a combination of ethical, scientific and feasibility requirements.

The statements provide a much-needed definition for endpoints in RCTs for PAH. The ultimate clinical endpoint for this fatal disease is time-until-death. Certain PAH therapies, such as intravenous epoprostenol and lung transplantation, entail risk and some compromise in quality of life as a trade-off for the benefit received, so that prolonging the time until resorting to these modalities would also be desirable.

The working groups endorsed the use of combining these events in composite endpoints. The advantages of composites include increased precision (from an increased number of events), a less restrictive definition of "failure of therapy", and the inclusion of multiple endpoints without increasing Type I error [22]. A potential drawback of this approach is a less certain answer. Therapeutic decisions and "clinical-worsening" endpoints included in the composite endpoint are often subjectively determined. Significant effects on such endpoints may not be generalised without objective protocol guidelines. Also, the components of a composite endpoint are equally weighted. Clearly, hospitalisation or initiation of epoprostenol is preferable to death; however, these three events would be considered equivalent in a combined endpoint.

Substantial improvement or maintenance of daily function and quality of life (intermediate endpoints) in this progressive disease would also be quite valuable, even if the time until ultimate clinical endpoints was unchanged in the end. Intermediate endpoints are distinct from surrogate endpoints, defined by the Food and Drug Administration as "a laboratory measurement or physical sign that is used in therapeutic

*Lung Transplant Program, Division of Pulmonary, Allergy and Critical Care Medicine, College of Physicians and Surgeons, and Dept of Epidemiology, Joseph L. Mailman School of Public Health, Columbia University, NY and [#]Pulmonary Vascular Disease Program, Pulmonary, Allergy, and Critical Care Division, University of Pennsylvania School of Medicine, PA, USA.

Correspondence: S.M. Kawut, Division of Pulmonary, Allergy and Critical Care Medicine, PH 8E, Rm. 101, 622 W. 168th Street, New York City, NY 10032, USA. Fax: 1 2123422201. E-mail: sk2097@columbia.edu

trials as a substitute for a clinically meaningful endpoint... and is expected to predict the effect of the therapy" [23]. Valid surrogate endpoints are measures or quantities which, although not necessarily clinically important in themselves, reflect the effect of a therapy on the ultimate clinical endpoints.

A single measure may be both a surrogate and an intermediate endpoint. For example, it appears that therapy which affects changes in distance walked in six minutes (6MWD) also reduces mortality [24], so that this measure may be a surrogate endpoint in PAH. However, the 6MWD may also be an intermediate endpoint, in that some absolute change in distance walked may actually improve a patient's quality of life. In this latter case, while we may not draw inferences about beneficial effects on survival, we would welcome a drug which improved function in a clinically significant way (providing it did not shorten survival).

PEACOCK *et al.* [11] and HOEPER *et al.* [12] highlight the unclear clinical relevance of certain parameters; these may then not be optimal intermediate endpoints. On the other hand, if we believe that a measure, such as cardiac index, is a valid surrogate endpoint in PAH, even a small effect on this measure may still have dramatic implications regarding long-term outcomes.

PEACOCK *et al.* [11] and other authors [25, 26] touch on the criteria for valid surrogate endpoints, which are helpful to review, as follows. 1) Surrogate measures must be reliable. The parameter must not vary within an individual subject and the measurement must have intra- and intercentre reproducibility for multicentre trials. 2) Biological plausibility is a consideration. The surrogate should be integral (or closely related) to a pathophysiological factor which leads to morbidity or mortality in PAH. 3) There must be a strong prognostic relationship between the marker and clinical outcomes, preferably documented in multiple studies. 4) There must be evidence from other RCTs in PAH that a quantitative modification of the surrogate measure reliably leads to a similar modification in the target outcome.

Investigators must incorporate the measurement of potential surrogates as secondary outcomes into RCTs, which are adequately powered to show an effect of an intervention on established endpoints. These trials must show that the effect of an intervention on clinical outcomes is preceded by some quantifiable effect on the surrogate, in order to validate the measure in PAH. While such validation studies may be expensive, these investigations prevent the use of unvalidated surrogate endpoints and uninterpretable results. Inconclusive studies are more costly to sponsors and can possibly delay approval of effective therapies.

Which potential surrogates considered by the working groups fulfil these criteria? Haemodynamics fulfil all of the criteria for valid surrogates [24], although the "right" parameter for PAH studies is unclear and could differ depending on the intervention being tested. For example, cardiac index may serve well as a surrogate endpoint for a clinical trial of a pulmonary vasodilator, whereas a pure inotrope may have a short-term effect on this endpoint without having a significant impact on long-term survival.

The invasiveness of these measures is a drawback. Also, regulatory agencies are hesitant to accept haemodynamics as surrogates due to their inadequacy in CHF, where therapies which improved the haemodynamic profile subsequently had adverse effects on overall outcomes [27]. It is difficult to extrapolate the success or failure of potential surrogate endpoints, or therapies, from one disease to the other. Decisions regarding acceptability of endpoints must be individualised according to disease.

Echocardiographic measurements also show promise as surrogates, since they fulfil all of the criteria [28, 29]. One

benefit over haemodynamics is the noninvasiveness of the measure. Intercentre variability and study quality, based on subject characteristics, are potential weaknesses.

Plasma biomarkers are an attractive option as assessment is relatively noninvasive, some have low within-subject variability and assays may be performed centrally. While certain biomarkers meet the first three criteria (*e.g.* brain natriuretic peptide [30]), RCTs of effective therapies with biomarker data have not been published, so the last criterion remains to be fulfilled.

6MWD appears to satisfy the requirements for validity and is the primary endpoint in virtually all recent RCTs in PAH, although some issues remain [2]. This effort-dependent endpoint can be subtly and unintentionally biased by clinicians or subjects in clinical trials of therapies with prominent side-effects or tell-tale laboratory abnormalities. For example, a patient in a study of a prostacyclin analogue with prominent flushing, jaw pain and gastrointestinal side-effects may put forth a different effort during exercise than a patient with no side-effects. In addition, conduct of the test may affect subject performance; methods must be standardised to ensure consistency in studies.

Not only is it important to decide which measure is the best surrogate, but which form of that measure should be used. A difference in mean 6MWD between two groups at the conclusion of a study is a different result than a difference between the changes in 6MWD between two groups. One of these functional forms may be a valid surrogate endpoint and the other may not. Future research should better define optimal parameters.

How can we maximise what we learn from each clinical investigation? The working group participants proposed sharing individual-patient data. There has been a relative explosion of RCTs in PAH. Each trial assembles a well-characterised cohort of patients with an exceedingly rare disease, offering the opportunity to: 1) better define the epidemiology of the disease and predictors of outcome; 2) merge datasets to perform meta-analyses; and 3) collect samples of plasma and genetic material for high-throughput analyses. While additional costs are a concern to sponsors, the first two options mainly require prospective planning in terms of informed consent and appropriate safeguards for data-sharing and cooperation among study investigators, which are ostensibly achievable goals.

A two-by-two factorial trial allows investigators to study two therapies "for the price of one" in the right setting. In this design, all patients are randomised to either active drug 1 or placebo 1 and either active drug 2 or placebo 2. Such trials are most efficient when the effect of each therapy does not depend on the presence or the absence of the other therapy. While these studies may be slightly more complex, they present the opportunity to test two drugs with a single trial in this rare disease.

These working-group statements initiate the discussion of how we should best study potential therapies for pulmonary arterial hypertension. We are fortunate to have the opportunity to address these issues. While many unknowns regarding the longevity of effects of new medical therapies and the timing of lung transplantation remain, discoveries may lie in novel methodological approaches. Future working groups such as these will surely be fertile ground for innovation, as we re-evaluate the answers we already have and formulate the next questions we want to ask.

References

1. Humbert M, Trembath RC. Genetics of pulmonary hypertension: from bench to bedside. *Eur Respir J* 2002; 20: 741–749.

2. Galiè N, Manes A, Branzi A. The new clinical trials on pharmacological treatment in pulmonary arterial hypertension. *Eur Respir J* 2002; 20: 1037–1049.
3. Chemla D, Castelain V, Hervé P, Lecarpentier Y, Brimiouille S. Haemodynamic evaluation of pulmonary hypertension. *Eur Respir J* 2002; 20: 1314–1331.
4. Eddahibi S, Morrell N, d'Ortho M-P, Naeije R, Adnot S. Pathobiology of pulmonary arterial hypertension. *Eur Respir J* 2002; 20: 1559–1572.
5. Widlitz A, Barst RJ. Pulmonary arterial hypertension in children. *Eur Respir J* 2003; 21: 155–176.
6. Moloney ED, Evans TW. Pathophysiology and pharmacological treatment of pulmonary hypertension in acute respiratory distress syndrome. *Eur Respir J* 2003; 21: 720–727.
7. Barberà JA, Peinado VI, Santos S. Pulmonary hypertension in chronic obstructive pulmonary disease. *Eur Respir J* 2003; 21: 892–905.
8. Dorfmüller P, Perros F, Balabanian K, Humbert M. Inflammation in pulmonary arterial hypertension. *Eur Respir J* 2003; 22: 358–363.
9. Maggiorini M, Velarde F-L. High-altitude pulmonary hypertension: a pathophysiological entity to different diseases. *Eur Respir J* 2003; 22: 1019–1025.
10. Darteville P, Fadel E, Mussot S, *et al.* Chronic thromboembolic pulmonary hypertension. *Eur Respir J* 2004; 23: 637–648.
11. Peacock A, Naeije R, Galiè N, Reeves JT. End points in pulmonary arterial hypertension: the way forward. *Eur Respir J* 2004; 23: 947–953.
12. Hooper MM, Oudiz RJ, Peacock A, *et al.* Endpoints and clinical trial designs in pulmonary arterial hypertension: clinical and regulatory perspectives. *J Am Coll Cardiol* 2004 (in press).
13. Rubin LJ, Badesch DB, Barst RJ, *et al.* Bosentan therapy for pulmonary arterial hypertension. *N Engl J Med* 2002; 346: 896–903.
14. Olschewski H, Simonneau G, Galie N, *et al.* Inhaled iloprost for severe pulmonary hypertension. *N Engl J Med* 2002; 347: 322–329.
15. 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.
16. Temple R, Ellenberg SS. Placebo-controlled trials and active-control trials in the evaluation of new treatments. Part 1: ethical and scientific issues. *Ann Intern Med* 2000; 133: 455–463.
17. Ellenberg SS, Temple R. Placebo-controlled trials and active-control trials in the evaluation of new treatments. Part 2: practical issues and specific cases. *Ann Intern Med* 2000; 133: 464–470.
18. Rothman KJ, Michels KB. The continuing unethical use of placebo controls. *N Engl J Med* 1994; 331: 394–398.
19. International Conference on Harmonisation. Topic E 10 Choice of control group in clinical trials 2000. www.emea.eu.int/pdfs/human/ich/036496en.pdf. Date accessed: March 1 2004.
20. Halpern SD, Karlawish JH, Casarett D, Berlin JA, Townsend RR, Asch DA. Hypertensive patients' willingness to participate in placebo-controlled trials: implications for recruitment efficiency. *Am Heart J* 2003; 146: 985–992.
21. Adequate and well-controlled studies. 21 CFR 314.126. Washington, DC, US Government Printing Office, GPO Access, 2003.
22. Freemantle N, Calvert M, Wood J, Eastaugh J, Griffin C. Composite outcomes in randomized trials: greater precision but with greater uncertainty? *JAMA* 2003; 289: 2554–2559.
23. New drug, antibiotic, and biological drug product regulations; accelerated approval- FDA. Final rule. *Fed Regist* 1992; 57: 58942–58960.
24. 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.
25. International Conference on Harmonisation. Topic E 9 Statistical principles for clinical trials 1998. www.emea.eu.int/pdfs/human/ich/036396en.pdf. Date accessed: March 1 2004.
26. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med* 1996; 125: 605–613.
27. 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.
28. Hinderliter AL, Willis PW 4th, 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. *Circulation* 1997; 95: 1479–1486.
29. 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.
30. 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.