Pulmonary Arterial Hypertension

Sample Size and Cost Analysis for Pulmonary Arterial Hypertension Drug Trials Using Various Imaging Modalities to Assess Right Ventricular Size and Function End Points

Karima Addetia, MD; Nicole M. Bhave, MD; Corey E. Tabit, MD, MBA, MPH; Mardi Gomberg-Maitland, MD, MSc; Benjamin H. Freed, MD; Karin E. Dill, MD; Roberto M. Lang, MD; Victor Mor-Avi, PhD*; Amit R. Patel, MD*;

Background—Placebo-controlled trials for pulmonary arterial hypertension are no longer acceptable because new therapies must show clinically significant effects on top of standard treatment. The purpose of this study was to estimate sample sizes and imaging costs for the planning of a hypothetical pulmonary arterial hypertension drug trial using imaging to detect changes in right ventricular size and function in response to combined therapy.

Methods and Results—Same-day cardiovascular MR (CMR) and 2-dimensional (2D) and 3D transthoracic echocardiography (2DTTE and 3DTTE) were performed in 22 patients with pulmonary arterial hypertension (54±13 years of age) twice, 6 months apart. Short-axis CMR cines and full-volume 3DTTE data sets of the right ventricle were used to measure end-diastolic volume and ejection fraction. Fractional area change was obtained from 2DTTE. Sample size calculations used a 2-sample t test model incorporating differences between baseline and 6-month measurements. Cost estimates were made using the Medicare fee schedule. No significant differences were noted between baseline and follow-up measurements. Large SDs reflected variable progression of disease in individual patients on standard therapy and measurement variability. These sources of variability resulted in intertechnique differences in sample sizes: to detect a change of 5% to 15% in 3DTTE-derived right ventricular ejection fraction and fractional area change or change of 15 to 30 mL in 3DTTE right ventricular end-diastolic volume; sample sizes were 2x to 2.5x those required by CMR. As a result, the total cost of a trial using complete TTE was greater than CMR, which was greater than limited TTE.

Conclusions—Because of lower measurement variability, CMR is more cost saving in pulmonary arterial hypertension drug trials than echocardiography, unless limited TTE is used. (Circ Cardiovasc Imaging, 2014;7:115-124.)

Key Words: cardiac imaging techniques ■ hypertension, pulmonary ■ right ventricle ■ sample size

The concept of a placebo-controlled drug trial has evolved considerably during the past few decades. New therapies can no longer be tested against an untreated group. Today, new therapies must show clinically significant results over and above the effects of standard therapy over time.

Clinical Perspective on p 124

Pulmonary arterial hypertension (PAH) is a devastating disease, despite current medical therapy.1 Increased pulmonary vascular resistance, resulting in increased right ventricular (RV) afterload and ultimately right ventricle failure, is the inevitable course of this disease. Prognosis is poor, with mortality in the range of 8% to 15% at 1 year with a mean survival of only 3 years.2-5 There is a degree of uncertainty in the PAH literature as to which end point should be used to assess the clinical response to drug therapy in clinical trials. Primary end points must be sensitive to treatment effect, measurable, and interpretable.6 Consequently, they are generally chosen from prognostic indicators. The 6-minute walk test is the most commonly used primary end point in randomized-controlled trials, essentially because it has prognostic significance in PAH (especially in the shortterm),7,8 is easy to perform, and is inexpensive. Because patients participating in clinical trials these days are less sick, this measure is allegedly less predictive and is thus falling out of favor, and new end points are being sought.

RV function is a major determinant of functional capacity and prognosis in PAH.9-11 In fact, death in PAH usually results from RV failure.12 Thus, imaging techniques such as echocardiography13 and cardiovascular MRI (CMR), which are used to evaluate RV size and function, should theoretically provide good end points for drug trials.14-16 On CMR, it is known that large RV volume and low RV ejection fraction (EF) predict mortality. Accordingly, in this study, we focused on several parameters of RV size and

Received July 17, 2013; accepted October 23, 2013.
From the Section of Cardiology, Department of Medicine (K.A., N.M.B., C.E.T., M.G.-M., B.H.F., R.M.L., V.M.-A., A.R.P.), and Department of Radiology (K.E.D., R.M.L., A.R.P.), University of Chicago, IL.
*Drsv Mor-Avi and Patel contributed equally to this work.
Correspondence to Amit R. Patel, MD, University of Chicago, 5841 South Maryland Ave, MC5084, Chicago, IL 60637. E-mail apatel2@medicine.bsd.uchicago.edu
© 2013 American Heart Association, Inc.
Circ Cardiovasc Imaging is available at http://circimaging.ahajournals.org DOI: 10.1161/CIRCIMAGING.113.000932

115
function obtained using 3 imaging modalities, 2D transthoracic echocardiography (2DTTE), 3D transthoracic echocardiography (3DTTE), and CMR, as potential imaging-based end points in a hypothetical drug trial for patients with PAH. Because patients in the control arm in such a trial would be on personalized standard therapy that could not be withheld and additionally be influenced by variations in volume status related to an individual's daily lifestyle decisions, this study was designed to incorporate into the sample size calculation changes in the measured parameters over time, attributed to both the effects of the standard treatment, the individual lifestyle decisions, and the intrinsic measurement variability of each technique.

Specifically, we sought to determine the sample sizes necessary to detect the effects of a new drug using these imaging modalities and the costs associated with their use. Accordingly, we first used a group of patients with PAH to study the effects of standard therapy and individual lifestyle decisions on various RV size and function parameters over time. We then used this information to determine the sample sizes required to detect a series of incremental changes in these parameters over time. Finally, we performed a cost analysis to determine what would be the least costly modality to use in a hypothetical standard-therapy–controlled drug trial. As a secondary goal, we aimed to develop a model for sample size determination that could be extrapolated to other disease states needing efficacy evaluation on top of standard therapy.

**Methods**

**Patient Population and Follow-Up**

Patients with PAH were prospectively recruited from the dedicated Pulmonary Hypertension (PH) Clinic at the University of Chicago, one of the largest national referral centers for patients with PAH. Patients were eligible for the study if they: (1) had documented PH on right heart catheterization, based on mean pulmonary artery pressure >25 mm Hg, pulmonary capillary wedge pressure, left atrial pressure or left ventricular end-diastolic pressure ≤15 mm Hg, and a pulmonary vascular resistance >3 Wood units at rest; (2) were categorized as World Health Organization (WHO) group 1 PH (revised WHO Classification);17 (3) were ≥18 years of age; and (4) were able to complete baseline and 6-month follow-up imaging studies. Patients were ineligible if they had standard contraindications to CMR, such as implanted defibrillator or pacemaker, claustrophobia, or inability to perform repeated breath holds.

Each patient had a complete clinical assessment with a PH specialist, including basic demographics, pathogenesis of PAH, WHO functional class, and medication history. Subsequently, each patient underwent 2DTTE, 3DTTE (iE33, Philips), and CMR imaging on the same day. Patients were scheduled for repeat imaging 6 months later. Written informed consent was obtained from all patients. The study was approved by the institutional review board.

**Cardiovascular MRI**

Images were acquired using a 1.5-Tesla scanner (Achieva, Philips; The Netherlands). Retrospectively, ECG-gated cardiac cine images were acquired using a steady-state free-precession pulse sequence (repetition time, 2.9 ms; echo time, 1.5 ms; flip angle, 60°; and temporal resolution, ≈30–40 ms) as a series of short-axis slices (6 mm thickness with 4 mm gap), including both right and left ventricles from base to apex. Images were analyzed (Philips ViewForum; The Netherlands) by a level-III CMR-trained imaging specialist to calculate RV end-diastolic volume (RVEDV), RV end-systolic volume, and RVEF (Figure 1). Delineation of RV boundaries at end diastole and end systole was performed manually for each slice. On the most basal slices, the right ventricle was differentiated from the right atrium by advancing the cine loop frame by frame throughout systole. If the cavity was becoming smaller and the myocardium thicker, it was included in the RV volume, whereas portions of the cavity that were becoming larger and did not show wall thickening were considered part of the atrium. Endocardial trabeculae were included in the RV cavity. Disk summation was used to calculate volumes.

**Two-Dimensional Transthoracic Echocardiography**

Comprehensive 2D and Doppler evaluation was performed by an experienced sonographer using the iE33 imaging system equipped with an S5 transducer (Philips Healthcare; Andover, MA). Digital cine loops of the apical 4-chamber view (including modified views to optimize visualization of the RV free wall) were acquired with careful attention to avoid foreshortening and stored digitally (Xcelera, Philips). Fractional area change (FAC), defined as [end-diastolic area–end-systolic area]/end-diastolic area×100, in the RV-focused apical 4-chamber view, was measured by an echocardiographer with level-III training in accordance with the guidelines18 (Figure 2).

**Three-Dimensional Transthoracic Echocardiography**

3DTTE images of the right ventricle were acquired from a modified apical 4-chamber view using the same imaging system with a matrix array transducer (X5-1) by the same sonographer. Full-volume acquisition was performed using ECG gating for 4 consecutive cardiac cycles during a breath hold.19 Digitally saved 3DTTE data sets were analyzed (4D Right Ventricle-Function 1.1, TomTec Imaging Systems; Untersleisseheim, Germany) by an echocardiographer with level-III training to quantify RVEDV, RV end-systolic volume, and RVEF. This required manual initialization of contours in user-defined end-systolic and end-diastolic frames in the apical 4-chamber, coronal, and sagittal views while including the trabeculae in the RV cavity (Figure 3).

**Statistical Analysis**

Continuous variables were expressed as mean±SD. Baseline and follow-up measures were compared using paired t tests. P <0.05 was considered significant.

**Measurement Reproducibility**

All measurements were performed by 2 independent observers blinded to results from other modalities to assess interobserver variability. Interobserver variability was expressed in terms of intraclass correlation coefficient and percent variability, defined as the mean of the absolute differences between pairs of repeated measurements divided by their mean.

**Sample Size Determination**

Sample size calculations modeling a randomized-controlled drug trial were performed as follows. The study group was used to represent a sample of the standard-therapy–controlled group. For the purposes of this calculation, the treatment group was assumed to have mean baseline characteristics (WHO class, pathogenesis of PAH, severity of disease, etc) and SD of RV parameters similar to the control group. Sample size calculations were performed for the following parameters: FAC on 2DTTE, RVEF and RVEDV on 3DTTE, and RVEF and RVEDV on CMR.

Calculations used the 2-sample t test model, incorporating the difference between baseline and 6-month measurements and the SD of the differences to account for changes over time (PASS version 11). Sample sizes were calculated for 80% power and significance level α<0.05 using the following formula:

\[
 n = 2 \left( \frac{2.8 \times SD}{\mu_1 - \mu_2} \right)^2
\]

where n is the per-group sample size, SD is the standard deviation of the measured mean change over time (averaged over all patients), and \( \mu_1 \) and \( \mu_2 \) are the changes over time in the measured parameter. 
in groups 1 (treatment group) and 2 (control group), such that μ1 − μ2 represented the intergroup difference in the changes over time that investigators in a hypothetical trial may wish to detect.

Sample sizes required to detect [μ1 − μ2] difference equal to 5%, 10%, and 15% improvements in FAC, 3DTTE RVEF, and CMR RVEF and 15, 20, and 30 mL changes in 3DTTE RVEDV and CMR RVEDV were calculated.

Cost Analysis
Comparative cost analysis was performed from the perspective of an independent academic researcher who outsources all research-related diagnostic studies at fair market rates (see Appendix in the online-only Data Supplement for definition of terms). All values are reported in 2013 US dollars. We assumed that all participating centers were financially solvent through clinical activities. To estimate the relative cost difference between CMR and TTE, we estimated the marginal cost of performing each study using modified bottom-up and top-down approaches.

Using the bottom-up approach, US national average salaries for echocardiography and MRI technologists were obtained from the US Bureau of Labor and Statistics. Hourly labor cost was calculated assuming 50 weeks annually and 40 hours weekly. Study durations were estimated as follows: 60 minutes for CMR, 40 minutes for complete TTE, and 20 minutes for limited TTE with a focus on the right ventricle, which are consistent with the durations of these studies at our center and include setup and cleanup activities before and after each examination. The interpreting physician’s opportunity cost per study was estimated using the 2013 Medicare fee schedule for the appropriate current procedural terminology codes for the studies in question: 93306 for a complete echocardiogram, 93308 for limited echocardiogram, and 75557 for CMR without contrast (CMS.gov/}

Figure 1. Example of a series of cardiovascular MRI short-axis slices, spanning both right and left ventricles from base to apex obtained at end systole in a patient with pulmonary arterial hypertension. See Cardiovascular MRI section for details.

Figure 2. Two-dimensional echocardiographic right ventricle–focused apical 4-chamber view obtained in a patient with pulmonary arterial hypertension, shown with manually traced endocardial boundaries at end diastole (left) and end systole (right), which were used to calculate end-systolic and end-diastolic areas and fractional area change.
It was assumed that an independent interpreting physician would fail to earn an amount equal to the standard Medicare professional fee for each research study interpreted. Incidental disposable costs such as syringes and linens were assumed to be similar for both methods and were ignored. In our calculation of marginal cost, fixed costs such as rent and equipment depreciation were ignored as well because the imaging center was assumed to be financially solvent through clinical activities alone. Therefore, using the bottom-up approach, the marginal cost of performing each study was estimated by adding the cost of technologist labor to the professional fee.

Using the top-down approach, Medicare global reimbursements for CMR, complete TTE, and limited TTE were obtained using the CPT codes listed above and used to estimate the opportunity cost to the imaging center performing each research study. The term global reimbursement refers to the sum of the professional and technical components of Medicare reimbursement obtained from the CMS Web site listed above. It was assumed that an otherwise independent and solvent clinical facility would fail to earn an amount equal to the standard Medicare global reimbursement for each research study performed because the clinical facility could not simultaneously perform a paid clinical study. Therefore, using the top-down approach, the estimated marginal cost is equal to the Medicare global reimbursement for a similar clinical study.

Sensitivity analyses were performed around Medicare reimbursement rates using the national average and the 90 different Medicare reimbursement localities. Sensitivity analyses were also performed around technologist labor costs to account for geographic variations.

**Results**

Of the 24 patients recruited for the study, 2 patients were eliminated because of loss to follow-up and poor quality of 3D images (8% dropout). The remaining 22 patients (54±13 years of age; 95% women) had baseline and 6-month follow-up studies (mean, 6.7±0.9 months). Of these patients, 9 (41%) were WHO functional class II at baseline. The majority of patients (14 of 22; 64%) were classified as idiopathic PAH or PAH secondary to connective tissue disease (Table 1). RV measurements for baseline and 6-month follow-up studies are shown in Table 2. No significant differences were noted between 0- and 6-month measurements with CMR, 2DE, or 3DE, but the relatively large SDs reflected variable progression of disease and day-to-day lifestyle decisions in individual patients on standard therapy (Table 2; Figures 4 and 5).

Interobserver variability results are shown in Table 3. Correlation between the repeated measurements depended heavily on the modality and the parameter. All parameters showed good to excellent intraclass correlation (0.8–0.99). Percent variability was worst for 2DTTE-derived FAC.

Sample size results are presented in Tables 4 and 5. For EF parameters (3D RVEF and FAC), sample sizes required for 3DTTE and 2DTTE were identical. These sizes were roughly twice those required for CMR RVEF. For the RVEDV measurements, sample sizes required for 3DTTE were roughly 2.5× those required for CMR.

**Cost Analysis**

Using a bottom-up approach, we estimated the marginal cost to perform 1 CMR as $147.62 and 1 complete/limited TTE as $85.25/$36.51. This reflects component costs of technologist labor (estimated at $70,000/y or $35/h; study durations of 60, 40, and 20 minutes, respectively) and professional fees for interpretation (estimated using national averages of $112.15, $61.92, and $24.84, respectively). Although the marginal cost for each CMR was greater than that for a complete TTE, the far smaller sample size needed for CMR resulted in it being less costly than complete TTE. However, if limited TTE is performed, CMR is more costly than TTE because the far lower marginal cost of limited TTE (driven by the lower cost of interpretation) nullifies the effect of the larger sample size.

Similarly, using a top-down approach, we estimated the marginal cost to perform 1 CMR as $366.77 and 1 complete/
Table 1. Baseline and Follow-Up Patient Characteristics

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Baseline (n=22)</th>
<th>Follow-Up (n=22)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, y, average±SD (range)</td>
<td>54±13 (28–81)</td>
<td></td>
</tr>
<tr>
<td>% Men</td>
<td>(1/22) 5%</td>
<td></td>
</tr>
<tr>
<td>BSA</td>
<td>1.8±0.2</td>
<td></td>
</tr>
<tr>
<td>WHO functional class</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I (%)</td>
<td>6 (27)</td>
<td>4 (18)</td>
</tr>
<tr>
<td>II (%)</td>
<td>9 (41)</td>
<td>8 (36)</td>
</tr>
<tr>
<td>III (%)</td>
<td>6 (27)</td>
<td>9 (41)</td>
</tr>
<tr>
<td>Unknown</td>
<td>1 (5)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>6-minute walk distance</td>
<td>378±120 (n=20)</td>
<td>378±104 (n=18)</td>
</tr>
<tr>
<td>Pathogenesis of PAH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Idiopathic</td>
<td>9 (41%)</td>
<td></td>
</tr>
<tr>
<td>Connective tissue disease/scleroderma</td>
<td>5 (23%)</td>
<td></td>
</tr>
<tr>
<td>Congenital</td>
<td>4 (18%)</td>
<td></td>
</tr>
<tr>
<td>Anorexigen use</td>
<td>3 (14%)</td>
<td></td>
</tr>
<tr>
<td>HIV</td>
<td>1 (5%)</td>
<td></td>
</tr>
<tr>
<td>Noninvasive and invasive measure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sPAP, mm Hg</td>
<td>73±28 (n=18)</td>
<td>75±26 (n=16)</td>
</tr>
<tr>
<td>CMR LVEF, %</td>
<td>59±6</td>
<td>59±8</td>
</tr>
<tr>
<td>Number of PAH Medications at baseline</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No medications</td>
<td>2 (9%)</td>
<td>0</td>
</tr>
<tr>
<td>Single drug therapy</td>
<td>6 (27%)</td>
<td>6 (27%)</td>
</tr>
<tr>
<td>Double drug therapy</td>
<td>11 (50%)</td>
<td>13 (59%)</td>
</tr>
<tr>
<td>Triple drug therapy</td>
<td>3 (14%)</td>
<td>3 (14%)</td>
</tr>
<tr>
<td>Duration of patient follow-up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years since PAH diagnosis</td>
<td>6±5</td>
<td></td>
</tr>
<tr>
<td>Average follow-up, mo</td>
<td>...</td>
<td>6.7±0.9</td>
</tr>
</tbody>
</table>

CMR indicates cardiovascular MRI; LVEF, left ventricular ejection fraction; PAH, pulmonary artery hypertension; sPAP, systolic pulmonary artery pressure; and WHO, World Health Organization.

Table 2. Baseline and Follow-Up Measurements for Right Ventricle Parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Total Patient Population (n=22)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
</tr>
<tr>
<td>CMR</td>
<td></td>
</tr>
<tr>
<td>RVEDV, %</td>
<td>35±15</td>
</tr>
<tr>
<td>RVEDV, mL</td>
<td>243±57</td>
</tr>
<tr>
<td>2D echocardiography</td>
<td></td>
</tr>
<tr>
<td>FAC</td>
<td>29±13</td>
</tr>
<tr>
<td>3D echocardiography</td>
<td></td>
</tr>
<tr>
<td>RVEDV, %</td>
<td>33±11</td>
</tr>
<tr>
<td>RVEDV, mL</td>
<td>191±60</td>
</tr>
</tbody>
</table>

2D indicates 2-dimensional; 3D, 3-dimensional; CMR, cardiovascular MRI; RVEDV, right ventricle end-diastolic volume; RVEF, right ventricle ejection fraction; and FAC, fractional area change.

Discussion

PAH is thought to be a result of endothelial dysfunction and adverse vascular remodeling secondary to a mixture of excessive cell proliferation, reduced apoptosis, thrombosis, and vasoconstriction. Three main pathways play a role in this imbalance: the prostacyclin pathway, the NO pathway, and the endothelin pathway.20 Currently available therapies, although designed specifically to target these pathways, are not associated with a clinically significant reduction in mortality and only a small improvement in exercise capacity21 so that much still needs to be performed in the development of therapies with higher clinical effect. In meeting this goal, testing of therapies must also be performed using outcome measures that truly parallel survival, especially because the population of patients being targeted for drug therapy today is less symptomatic.

In this study, we evaluated several RV parameters measured by 2DTTE, 3DTTE, and CMR as potential end points to study the effects of new therapies in PAH. Sample sizes necessary to detect specific changes in RV function (RVEF and FAC) were ≈2× higher for 2DTTE and 3DTTE than CMR. With respect to volume, RVEDV is not a parameter of RV function, and accordingly, a 15-mL change in volume cannot be directly equated with a 5% change in RVEF. Nevertheless, our results showed that sample sizes required to detect 15, 20, and 30 mL changes in RVEDV were considerably larger than those required to detect 5%, 10%, and 15% changes in RVEF. In both top-down and bottom-up analyses, the marginal cost of CMR is greater than that of TTE. However, because of the higher sensitivity of CMR and smaller sample sizes needed to detect changes in RV size and function, CMR is less expensive than complete TTE in its total cost. However, if a limited TTE is performed instead of complete TTE, CMR is again more costly because of the far lower marginal cost.

Because Medicare reimbursement rates vary across the United States (Tables 6 and 7), we performed similar analyses using rates from each of the 90 different localities (Medicare Fee Schedule). Although reimbursement rates varied considerably (global fees: $272.81–$472.62 for CMR, $142.00–$243.56 for complete TTE, and $73.56–$130.73 for limited TTE; professional fees: $100.64–$152.40 for CMR, $55.64–$84.67 for complete TTE, and $22.50–$34.17 for limited TTE), the cost savings of CMR compared with complete TTE and the cost savings of limited TTE compared with CMR were preserved in all 90 localities because of strong correlation between the reimbursement rates for CMR and TTE. Geographic variation in technologist labor costs was trivial, and technologist labor was not a major driver of total cost.

RV Parameters as End Points in PAH Drug Trials

It is widely agreed that end points in PAH drug trials should be reliable, valid, reasonably easy to repeat, interpretable, and
should reflect the effect of an intervention on clinical outcome. RV function is a major determinant of functional capacity and prognosis in PAH. On CMR, large RV volume, poor RV function, and reduced left ventricular volume are strong independent predictors of mortality and treatment failure. On echocardiography, significant RV enlargement/dysfunction and elevated myocardial performance (Tei) index are predictive of poor prognosis. Death from PH usually results from RV failure; thus, such changes in RV size and function as detected by imaging techniques should theoretically provide useful end points for trials. In the case of left heart disease, it is known that cardiac morphology changes can long precede hemodynamic changes. Similarly, imaging may capture changes in the right heart before hemodynamic decompensation occurs.

**Accounting for Changes in Standard Therapy**

Placebo-controlled clinical trials in principle can no longer be performed in PAH because withholding current therapy, even for as short a time period as 12 to 16 weeks, worsens prognosis. Previous studies have only taken into account the intermeasurement variability in the calculation of sample size. We chose to study the effects of current therapy on RV function parameters over time and then to account for this variation in our sample size determination. We discovered that individual patients on standard therapy had widely variable changes in RV parameters over the chosen follow-up period, resulting in relatively large SDs and consequently no significant change in the group as a whole. To incorporate the changes in RV parameters over time, we used in the sample size calculation the average SD for the measured change over time, which takes into account both the change in the given parameter over time and the interobserver variability.

**Sample Size Results**

Because PAH is a rare disease, keeping the required sample size small is important. Given that it took >1 year to enroll the

![Figure 4. Individual patients’ responses to standard therapy as reflected by cardiovascular MRI measures of right ventricular ejection fraction (RVEF) and right ventricular end-diastolic volume (RVEDV) during 6-month follow-up. Note that group averages did not show significant differences because of a wide range of individual responses.](http://circimaging.ahajournals.org/)

![Figure 5. Individual patients’ responses to standard therapy as reflected by 2-dimensional (2D; left) and 3D (middle and right) echocardiographic indices of right ventricular function during 6-month follow-up: right ventricular fractional area change (FAC), right ventricular ejection fraction (RVEF), and right ventricular end-diastolic volume (RVEDV).](http://circimaging.ahajournals.org/)
patients in this study, it is also true that the number of patients obtainable for any PAH drug study will likely be small. The per-group sample size (see Equation) is directly proportional to the square of SD and inversely proportional to the square of the difference between the change over time in the 2 groups that one wishes to detect. Thus, the SD of the mean change over time in the population sample is the driving force for the calculated sample size. In our study, patients’ responses to the standard therapy varied widely, resulting in large SDs (Table 2), and as a result, the calculated sample sizes were relatively large. The SD of the changes over time is also affected by intrinsic measurement variability, which differed among parameters and modalities (Table 3). The overall variability in 3DTTE and 2D FAC was greater than for CMR RVEF so that the sample size for RVEF using CMR was smaller than that for 2D or 3DTTE.

Of note, the calculated sample size needs to be put in perspective of intrinsic measurement variability. To detect a given change in the treatment group compared with the control group, that change must be larger than the intrinsic measurement variability. In this study, the desired 5% change for CMR RVEF and 3DTTE RVEF constituted 14% and 15% of the baseline values, respectively (Table 2), whereas the corresponding interobserver variability values were 11.5% and 16% (Table 3). Accordingly, for CMR, a 5% expected change is reasonable to detect because 14% is greater than 11.5%. However, a 5% expected change with 3DTTE may be more difficult to detect because 15% is similar to 16%. When the same principle is applied to volumes (CMR RVEDV and 3DTTE RVEDV) and FAC, the smallest expected changes in these parameters tested in this study (15 mL for volumes and 5% for FAC) were larger than the intrinsic variability of the respective technique. This means that these small changes would be difficult to detect with sample sizes calculated in this study, and larger samples would be required to achieve this goal.

Interobserver variability depends on the specific group of patients, as well as on the measurement methodology, which may vary between sites and among observers. Accordingly, it is important to evaluate measurement variability in the specific patient population that is to be studied in the hands of the observers who will be making the measurements. Measurement variability can be minimized by unifying measurement strategy between observers and sites.

### Cost Analysis

The magnitude of cost difference between TTE and CMR may be relatively large or small depending on the degree of measurement sensitivity required (Tables 6 and 7). These variations in relative cost are preserved across the 90 localities, with only small fluctuations in the relative cost between the imaging modalities. Opportunity cost was overwhelmingly the primary cost driver in our analysis. Although technologist labor cost varies by the duration of each study, this cost is far outweighed by opportunity cost and is therefore not a major driver. Geographic variation in opportunity cost (ie, Medicare reimbursement rates) was substantial (Figure 6).

### Limitations

Our choice to use RV imaging parameters as surrogate end points in PAH drug trials might be criticized because there is not enough evidence in the literature to support it and because of the large SDs in these parameters. However, RV size and function are known to be highly prognostic in these patients and are likely relevant in the assessment of improvement in response to new therapy. Accordingly, we think that RV imaging holds promise in this capacity. Another approach would be to use a composite outcome, such as a combination of
RV function with the 6-minute walk test. The effect of this strategy would require further study but could still be accomplished within the framework we proposed.

Our sample size calculations involved several assumptions that may not necessarily be true in every trial. Accordingly, our results may not be automatically extrapolated into every future PAH drug trial. Nevertheless, the reasoning used here to calculate sample size can be valuable for planning drug trials wherein standard therapy cannot be withheld.

One might question the size of our study cohort. It is important to remember that this cohort was used to obtain preliminary data to perform the sample size calculations, specifically the SDs of the relevant RV parameters measured in patients with PAH. There is no standard procedure to calculate the size of this preliminary cohort. We think that 22 is a reasonable number for this purpose, given the rarity of patients with PAH.

In performing the cost analyses, several key assumptions were made. First, it was assumed that the researcher would pay fair market rates for all services, a situation that often does not exist within universities or health systems, for which costs are frequently negotiated. Second, we assumed that the Medicare Fee Schedule reflected the true opportunity cost for each study. This assumption in turn assumes that the imaging center and interpreting physician would have performed a paid clinical study during the time of the research study. Any deviation from this assumption would alter the cost model. Third, we assumed the perspective of an independent researcher who outsources all imaging and is therefore not responsible for fixed costs, such as rent, maintenance, and insurance. However, in many circumstances, researchers purchase their own equipment and hire their own staff, which would also alter the cost model. Finally, there are several other costs that likely have an economic effect on an independent academic researcher. Namely, in a population of uncommon patients, doubling the sample size is associated with the cost of more staff, investigative pharmaceutical agents, patient incentives, and other costs that may overshadow the differences in imaging cost. These costs are difficult to estimate but must be anticipated in future clinical trials.

### Conclusions

Echocardiography shows higher intrinsic measurement variability over time compared with CMR. The use of CMR not only allows for the reduction of sample size in PAH drug trials but also reduces overall cost, unless limited TTE is used. Information gleaned from this study could potentially help in the construction of future PAH drug trials, which choose to use imaging as an end point. Further, the design of this study

### Table 6. Comparative Cost of Performing a Trial Using CMR, Complete TTE, or Limited TTE* Using Top-Down Analyses

<table>
<thead>
<tr>
<th>Change</th>
<th>CMR, n</th>
<th>CMR Global Fee, $</th>
<th>TTE, n</th>
<th>Complete TTE Global Fee, $</th>
<th>TTE vs CMR, $</th>
<th>Limited TTE Global Fee, $</th>
<th>Limited TTE vs CMR, $</th>
</tr>
</thead>
<tbody>
<tr>
<td>RVEF 5%</td>
<td>23</td>
<td>8435</td>
<td>51</td>
<td>9665</td>
<td>1229</td>
<td>5118</td>
<td>(3316)</td>
</tr>
<tr>
<td>RVEF 10%</td>
<td>6</td>
<td>2200</td>
<td>13</td>
<td>2463</td>
<td>263</td>
<td>1304</td>
<td>(895)</td>
</tr>
<tr>
<td>RVEF 15%</td>
<td>3</td>
<td>1100</td>
<td>6</td>
<td>1137</td>
<td>367</td>
<td>602</td>
<td>(498)</td>
</tr>
<tr>
<td>RVEDV 15 mL</td>
<td>68</td>
<td>24940</td>
<td>161</td>
<td>30511</td>
<td>5570</td>
<td>16159</td>
<td>(8780)</td>
</tr>
<tr>
<td>RVEDV 20 mL</td>
<td>38</td>
<td>13937</td>
<td>91</td>
<td>17245</td>
<td>3308</td>
<td>9133</td>
<td>(4803)</td>
</tr>
<tr>
<td>RVEDV 30 mL</td>
<td>41</td>
<td>6235</td>
<td>41</td>
<td>7769</td>
<td>1534</td>
<td>4115</td>
<td>(2119)</td>
</tr>
<tr>
<td>FAC 5%</td>
<td>51</td>
<td>9665</td>
<td>51</td>
<td>9665</td>
<td>5118</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAC 10%</td>
<td>13</td>
<td>2463</td>
<td>13</td>
<td>2463</td>
<td>1304</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAC 15%</td>
<td>6</td>
<td>1137</td>
<td>6</td>
<td>1137</td>
<td>602</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Numbers in parentheses represent negative values. CMR indicates cardiovascular MRI; RV, right ventricle; TTE, transthoracic echocardiography.

*Using national average Medicare reimbursement rates.

### Table 7. Comparative Cost of Performing a Trial Using CMR, Complete TTE, or Limited TTE* Using Bottom-Up Analyses

<table>
<thead>
<tr>
<th>Change</th>
<th>CMR, n</th>
<th>CMR Pro + Labor, $</th>
<th>TTE, n</th>
<th>Complete TTE Pro + Labor, $</th>
<th>TTE vs CMR, $</th>
<th>Limited TTE Pro + Labor, $</th>
<th>Limited TTE vs CMR, $</th>
</tr>
</thead>
<tbody>
<tr>
<td>RVEF 5%</td>
<td>23</td>
<td>3395</td>
<td>51</td>
<td>4348</td>
<td>953</td>
<td>1862</td>
<td>(1533)</td>
</tr>
<tr>
<td>RVEF 10%</td>
<td>6</td>
<td>885</td>
<td>13</td>
<td>1108</td>
<td>223</td>
<td>475</td>
<td>(411)</td>
</tr>
<tr>
<td>RVEF 15%</td>
<td>3</td>
<td>442</td>
<td>6</td>
<td>512</td>
<td>69</td>
<td>219</td>
<td>(224)</td>
</tr>
<tr>
<td>RVEDV 15 mL</td>
<td>68</td>
<td>10038</td>
<td>161</td>
<td>13726</td>
<td>3688</td>
<td>5878</td>
<td>(4161)</td>
</tr>
<tr>
<td>RVEDV 20 mL</td>
<td>38</td>
<td>5609</td>
<td>91</td>
<td>7758</td>
<td>2148</td>
<td>3322</td>
<td>(2287)</td>
</tr>
<tr>
<td>RVEDV 30 mL</td>
<td>17</td>
<td>2509</td>
<td>41</td>
<td>3495</td>
<td>986</td>
<td>1497</td>
<td>(1013)</td>
</tr>
<tr>
<td>FAC 5%</td>
<td>51</td>
<td>4348</td>
<td>51</td>
<td>4348</td>
<td>1862</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAC 10%</td>
<td>13</td>
<td>1108</td>
<td>13</td>
<td>1108</td>
<td>475</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAC 15%</td>
<td>6</td>
<td>512</td>
<td>6</td>
<td>512</td>
<td>219</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Numbers in parentheses represent negative values. CMR indicates cardiovascular MRI; EDV, end-diastolic volume; EF, ejection fraction; FAC, fractional area change; Pro, professional fee; and TTE, transthoracic echocardiography.

*Using national average Medicare reimbursement rates.
could be used as a model for other disease states, wherein new therapy needs to be tested in addition to standard-of-care treatment.

Acknowledgments
We thank Kristen Wroblewski, MS, senior biostatistician, Department of Health Studies, University of Chicago, for her invaluable help with the statistical section of this study.

Sources of Funding
This study was funded by research grants from Philips, Ventripoint, and the American Society of Echocardiography.

Disclosures
None.

References


Placebo-controlled drug trials are no longer acceptable in pulmonary arterial hypertension (PAH), primarily because withholding current therapy, even for a short period of time, worsens prognosis in these patients. Therefore, any newly developed therapy must be tested on top of standard treatment. For this to work, the effect of standard therapy on the parameter to be used to assess treatment effect must be known and incorporated in the sample size estimation for testing the new drug. In PAH, right ventricular size and function are known to be prognostic indicators; however, imaging, despite its proven value in the assessment of the right ventricle, is not used consistently as a parameter to assess the effects of treatment in these patients. This study suggests a model for designing PAH drug trials, which use imaging-based end points (2D and 3D echocardiography and cardiovascular MRI) to assess treatment efficacy in patients on standard therapy by providing a framework for sample size determination and a breakdown cost analysis. Because of lower measurement variability, cardiovascular MRI allows for significant reductions in sample size and is more cost saving in PAH drug trials than echocardiography, unless a limited echocardiogram is used. Use of imaging techniques that allow for the reduction of total sample size may be of particular importance in rare disorders such as PAH.
Sample Size and Cost Analysis for Pulmonary Arterial Hypertension Drug Trials Using Various Imaging Modalities to Assess Right Ventricular Size and Function End Points
Karima Addetia, Nicole M. Bhave, Corey E. Tabit, Mardi Gomberg-Maitland, Benjamin H. Freed, Karin E. Dill, Roberto M. Lang, Victor Mor-Avi and Amit R. Patel

Circ Cardiovasc Imaging. 2014;7:115-124; originally published online November 5, 2013; doi: 10.1161/CIRCIMAGING.113.000932
Circulation: Cardiovascular Imaging is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2013 American Heart Association, Inc. All rights reserved.
Print ISSN: 1941-9651. Online ISSN: 1942-0080

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circimaging.ahajournals.org/content/7/1/115

Data Supplement (unedited) at:
http://circimaging.ahajournals.org/content/suppl/2013/11/05/CIRCIMAGING.113.000932.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation: Cardiovascular Imaging can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation: Cardiovascular Imaging is online at:
http://circimaging.ahajournals.org//subscriptions/
SUPPLEMENTAL MATERIAL

Definition of Terms for Cost Analysis

Bottom-up approach – refers to the synthesis of component costs to estimate total cost.

Top-down approach – refers to the analysis of total cost to estimate component costs.

Marginal cost
The cost associated with producing one additional unit of product or service. For our model, we assumed that the marginal cost contains only direct medical costs and that overhead costs are covered by the imaging center’s clinical operations.

Direct medical costs
Costs directly associated with producing a product or providing a service. Examples in our model include technologist labor, physician time to interpret studies, ultrasound gel, linens, etc.

Overhead costs
Costs incurred but not directly associated with producing a product or providing a service. Examples in our model include rent for the imaging center, heat, equipment maintenance, management, etc.

Opportunity cost
The cost of an activity incurred from having foregone the next best alternative option. Examples in our model include the revenue a physician would have received from interpreting a paying clinical study instead of a research study, or the revenue an imaging center would have received from performing a paying clinical study instead of a research study.

Sensitivity analysis
The study of how variation in model inputs relates to variation in model outputs. For example, as technologist labor cost varies (up or down), how does this affect total cost?