Inaccuracy of Doppler Echocardiographic Estimates of Pulmonary Artery Pressures in Patients With Pulmonary Hypertension

Jonathan D. Rich, MDa,*, Sanjiv J. Shah, MD,b Rajiv S. Swamy, MDa Anna Kamp, MDb Stuart Rich, MD, FCCPa

a From the Pritzker School of Medicine Section of Cardiology, Northwestern University Feinberg School of Medicine, Chicago, IL
b Department of Medicine, University of Chicago Medical Center, and the Division of Cardiology, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL
c Department of Pediatrics, University of Michigan Medical Center, Ann Arbor, MI.

* Correspondence to: Jonathan D. Rich, MD, University of Chicago, 5841 S Maryland Ave, MC 6080, Chicago, IL 60637

E-mail address: jonathan.rich@uchospitals.edu

Manuscript received May 17, 2010, accepted August 9, 2010

Background

Recent studies suggest that Doppler echocardiography (DE)-based estimates of pulmonary artery systolic pressure (PASP) may not be as accurate as previously believed. We sought to determine the accuracy of PASP measurements using DE compared with right-sided heart catheterization (RHC) in patients with pulmonary hypertension (PH).

Methods

We compared DE estimates of PASP to invasively measure PASP during RHC in 160 consecutive patients with PH (part one). To account for possible changes in hemodynamics between DE and RHC, we then prospectively determined PASP in an additional 23 consecutive patients undergoing simultaneous RHC and DE (part two). Bland-Altman analyses were performed to evaluate the agreement between RHC and DE measurements of PASP. Accuracy was predefined as 95% limits of agreement within ± 10 mm Hg for PASP estimates.

Results
In part one, there was moderate correlation between DE and RHC measurements of PASP ($r = 0.68$, $P < .001$). However, using Bland-Altman analysis, the bias for DE estimates of PASP was 2.2 mm Hg with 95% limits of agreement ranging from -34.2 to 38.6 mm Hg. DE estimates of PASP were determined to be inaccurate in 50.6% of patients. In part two, there was moderate correlation between DE and RHC measurements of PASP ($r = 0.71$, $P < .01$). However, despite simultaneous DE and RHC measurements, the bias for DE estimates of PASP was 8.0 mm Hg with 95% limits of agreement ranging from -28.4 to 44.4 mm Hg.

**Conclusions**

DE estimates of PASP are inaccurate in patients with PH and should not be relied on to make the diagnosis of PH or to follow the efficacy of therapy.

---

**Abbreviations**

- **DE**: Doppler echocardiography
- **IVC**: inferior vena cava
- **PAP**: pulmonary artery pressure
- **PASP**: pulmonary artery systolic pressure
- **PH**: pulmonary hypertension
- **RAP**: right atrial pressure
- **RHC**: right-sided heart catheterization
- **RV**: right ventricle
- **TR**: tricuspid regurgitant

Pulmonary hypertension (PH) is a condition associated with substantial morbidity and mortality.\(^1\) The prognosis for patients with PH is generally poor and is closely related to several hemodynamic variables.\(^2\) \(^3\) Accurate measurements of hemodynamics, including pulmonary artery pressures (PAP), are critical for both the diagnosis and appropriate management of individual patients with PH. Right-sided heart catheterization (RHC) remains the accepted standard approach in determining PAP.\(^4\) \(^5\) However, because of its widespread availability and ease of use in the outpatient setting, Doppler echocardiography (DE) is implemented with increasing frequency as a tool for determining PAP, both for establishing a diagnosis and for monitoring treatment. In fact, DE has replaced RHC for determining hemodynamics in a National Institutes of Health-sponsored randomized clinical trial in patients with PH.\(^6\) Although DE estimates of PAP have been shown to correlate with RHC measurements of PAP,\(^7\) \(^8\) \(^9\) the reliability of DE to accurately measure PAP in patients with PH is not clear. Studies have suggested that DE estimates of PAP may not be as accurate as previously believed,\(^10\) \(^11\) \(^12\) although these studies did not compare simultaneous measurements of PAP by DE obtained during RHC, and, thus,
fluctuations in PAP at different points in time may have affected their findings.\[13\]

As a large referral center for patients with PH, the University of Chicago has become increasingly concerned about the accuracy of measurements of PAP made solely with DE and, as a consequence, the subsequent inappropriate diagnoses and management decisions that occur. Thus, we conducted the present study to evaluate the accuracy of DE as a noninvasive diagnostic tool in patients with PH.

**Materials and Methods**

The institutional review board at the University of Chicago (approval number 15271A) approved the conduct of this study, and all patients provided written consent prior to enrollment. This study was conducted in two parts.

**Part One**

We prospectively identified 160 consecutive patients referred for presumed PH and who subsequently underwent both RHC and DE between April 2005 and February 2008. Eligibility criteria included the following: (1) RHC and DE occurred within 30 days of each other, (2) no change in patient clinical status or medications occurred during this 30-day interval, and (3) a measurable tricuspid regurgitant (TR) jet, necessary for DE estimations of pulmonary artery systolic pressure (PASP), was present on the DE study.

**Part Two**

To address the possibility that variations in hemodynamics occurring during the 30-day interval between RHC and DE may affect our findings, we prospectively enrolled an additional 23 consecutive patients referred to our PH specialists for RHC as part of an evaluation of PH. DE imaging was performed on each of these patients in the cardiac catheterization laboratory simultaneous to the invasive hemodynamic measurements made during the RHC.

**Right-Sided Heart Catheterization**

All RHC measurements made during the study were performed by one of two physicians experienced in both cardiac catheterization and PH. Hemodynamic measurements obtained during the RHC, including right atrial pressure (RAP) and PASP, were recorded at end expiration. The physicians performing the RHC were blinded to the DE data obtained during the simultaneous performance of the RHC.

**Echocardiography**

Two-dimensional transthoracic echocardiography and DE were performed using either a Philips Sonos 7500 or IE33 machine with a 3.2 MHz transducer (Philips Medical Systems; Andover, Massachusetts). Echocardiograms were performed in multiple views to obtain the optimal appearing TR jet. Patients were excluded from the study if no tricuspid regurgitation was present on DE or if the TR jet was insufficient to estimate the PASP. After acquiring the necessary images, the TR jet velocities were measured, and estimated PASPs were determined using the modified Bernoulli equation in conjunction with an echocardiographic estimation of RAP. Echocardiographic estimation of RAP was performed based on inferior vena cava (IVC) size and collapsibility, according to the following established criteria\[14\]: RAP was estimated to be 0 to 5 mm Hg when the IVC diameter was $< 17$ mm and $\geq 50\%$ collapsibility; 5 to 10 mm Hg when the IVC diameter was $< 17$ mm and $< 50\%$ collapsibility; 10 to 15 mm Hg when the IVC diameter was $> 17$ mm and $< 50\%$ collapsibility; and 15 to 20 mm Hg when the IVC diameter was $> 17$ mm with absence of any collapsibility. All echocardiographic and DE studies were performed by an experienced, certified sonographer. For the simultaneous DE and RHC part of the study, a single, experienced, certified sonographer performed all the DE studies to exclude the possibility of intersonographer variability. Interpretations of echocardiographic and DE measurements were performed only by physicians experienced with and board certified in echocardiography. The sonographers and echocardiography physicians were all blinded to the RHC data in all parts of the study.

**Statistics**

Descriptive statistics were described using mean (SD) for continuous variables and number (%) for categorical variables. Correlation coefficients between RHC- and DE-derived hemodynamics were calculated using the Pearson correlation method. However, correlation coefficients are often misleading and should not be used to assess the agreement or interchangeability between two different modalities. It is generally expected that when two different modalities (eg, DE and RHC) are used to assess a particular variable (eg, PASP), with each modality designed to measure the same variable, there will be a strong correlation, especially if there is a wide range of values tested. A high correlation, however, does not automatically imply that there is good agreement between the two methods. Therefore, Bland-Altman analyses, which plot the difference between
two modalities against their mean, are necessary to truly evaluate agreement between two methods. For these reasons, we used Bland-Altman analyses to evaluate the agreement between DE estimates of PASP and RAP, respectively, with those same values as determined by RHC. To interpret the results of the Bland-Altman analyses clinically, DE estimates of PASP would be considered accurate if the differences within the 95% limits of agreement between DE and RHC measurements of PASP were small and not considered clinically relevant. Alternatively, it would be unacceptable to consider DE estimates of PASP as accurate compared with RHC measurements of PASP if the 95% limits of agreement were large and thought to be of clinical relevance. Accuracy was predefined as 95% limits of agreement within ± 10 mm Hg for PASP estimates and ± 2.5 mm Hg for RAP estimates. Statistical analyses were performed using Stata 10.1 (StataCorp; College Station, Texas).

Results

The baseline characteristics of the patients in this study can be seen in Table 1. The majority of patients were women who received a diagnosis of World Health Organization group 1 PAH, although patients with other causes of PH were also included in this study. The hemodynamics obtained during RHC and DE are summarized in Table 2. A wide range of PASP (20 mm Hg to 120 mm Hg) and RAP (1 mm Hg to 35 mm Hg) during RHC was observed. In a small minority of the patients, the RHC assessment of hemodynamics excluded the presence of PH, which was defined as a mean PAP ≥ 25 mm Hg.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Part One (n = 160)</th>
<th>Part Two (n = 23)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean ± SD, y</td>
<td>53.4 ± 14.5</td>
<td>52.8 ± 14.2</td>
</tr>
<tr>
<td>Female</td>
<td>78.1 (125)</td>
<td>78.3 (18)</td>
</tr>
<tr>
<td>PH present</td>
<td>96.7 (155)</td>
<td>91.3 (21)</td>
</tr>
<tr>
<td>PH diagnoses</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WHO group 1 PH</td>
<td>72.3 (112)</td>
<td>76.2 (16)</td>
</tr>
<tr>
<td>IPAH</td>
<td>37.4 (58)</td>
<td>47.6 (10)</td>
</tr>
<tr>
<td>CVD</td>
<td>19.4 (30)</td>
<td>19.0 (4)</td>
</tr>
<tr>
<td>CHD</td>
<td>7.7 (12)</td>
<td>6.3 (1)</td>
</tr>
<tr>
<td>Portal hypertension</td>
<td>7.7 (12)</td>
<td>6.3 (1)</td>
</tr>
<tr>
<td>Non-WHO group 1 PH</td>
<td>27.7 (43)</td>
<td>23.8 (5)</td>
</tr>
<tr>
<td>PVH</td>
<td>14.2 (22)</td>
<td>9.5 (2)</td>
</tr>
<tr>
<td>Lung disease</td>
<td>7.1 (11)</td>
<td>6.3 (1)</td>
</tr>
<tr>
<td>CTEPH</td>
<td>3.2 (5)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Sarcoid</td>
<td>3.2 (5)</td>
<td>9.1 (2)</td>
</tr>
</tbody>
</table>

Data are presented as % (No.) unless otherwise indicated. CHD = congenital heart disease; CTEPH = chronic thromboembolic pulmonary hypertension; CVD = collagen vascular disease; IPAH = idiopathic pulmonary arterial hypertension; PH = pulmonary hypertension; PVH = pulmonary venous hypertension; WHO = World Health Organization.

<table>
<thead>
<tr>
<th>Variable</th>
<th>RHC, mm Hg</th>
<th>DE, mm Hg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part 1 (n = 160)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PASP</td>
<td>79.3 ± 21.0</td>
<td>77.1 ± 24.2</td>
</tr>
<tr>
<td>RAP (n = 109)</td>
<td>8.8 ± 5.7</td>
<td>12.2 ± 5.0</td>
</tr>
<tr>
<td>Part 2 (n = 23)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PASP</td>
<td>65.5 ± 26.1</td>
<td>56.9 ± 20.5</td>
</tr>
<tr>
<td>RAP</td>
<td>7.1 ± 4.7</td>
<td>7.2 ± 3.6</td>
</tr>
</tbody>
</table>

Data are presented as mean ± SD. DE = Doppler echocardiography; PASP = pulmonary artery systolic pressure; RAP =
right atrial pressure; RHC = right-sided heart catheterization.

In part one of the study, there was a moderately strong correlation between PASP determined during RHC and estimated PASP using DE (r = 0.68, P < .001). However, using Bland-Altman analyses, the mean bias for DE estimates of PASP was 2.2 mm Hg with 95% limits of agreement ranging from -34.2 to 38.6 mm Hg (Fig 1). Thus, even after excluding the 5% of patients with the most discrepant values between DE and RHC measurement of PASP, DE underestimated PASP as measured by RHC by as much as 34.2 mm Hg and overestimated it by as much as 38.6 mm Hg. When we analyzed the data according to subgroups of patients based on the different causes of their PH (as listed in Table 1), we did not find DE estimates of PASP to be any more or less accurate in any of the individual subgroups. There was also a mild correlation between RAP measured during RHC and the echocardiographic estimate of RAP (r = 0.38, P < .001). However, Bland-Altman analysis revealed a bias of -3.3 mm Hg and 95% limits of agreement ranging from -15.6 to 8.9 mm Hg (Fig 2).

**Figure 1**  Bland-Altman analysis demonstrating a lack of agreement between DE estimates of pulmonary artery systolic pressure (PASP) and PASP determined during RHC (solid line), as highlighted by the 95% limits of agreement, ranging from -34.2 mm Hg to 38.6 mm Hg (dashed lines). Larger circles represent identical observations among multiple patients. The inaccuracy of DE estimates of PASP is particularly apparent at higher PASP. DE = Doppler echocardiography; RHC = right-sided heart catheterization.

**Figure 2**  A wide range of variation and lack of agreement exists between echocardiographic estimates of right atrial pressure (RAP) compared with RAP measured directly during RHC (solid line), as highlighted by the wide 95% limits of agreement, ranging from -15.6 mm Hg to 8.9 mm Hg (dashed lines). Larger circles represent identical observations among multiple patients. The inaccuracy of the echocardiographic estimates of RAP appears to be greater at higher RAPs. Echo = echocardiography. See Figure 1 legend for expansion of other abbreviations.

Using the predefined definition of an inaccurate DE estimation of PASP as ± 10 mm Hg different from PASP measured during...
RHC, we found that DE was inaccurate 50.6% (81/160) of the time. When inaccurate, DE overestimated PASP by > 10 mm Hg in 40.7% (33/81) of cases and underestimated PASP by > 10 mm Hg in 59.3% (48/81) of cases. Using the predefined definition of an inaccurate echocardiographic estimate of RAP as ± 2.5 mm Hg different from RAP measured during RHC, we found that echocardiography was inaccurate in 58.7% (64/109) of the cases. When inaccurate, echocardiography overestimated RAP by > 2.5 mm Hg in 64.1% (41/64) of cases and underestimated RAP by > 2.5 mm Hg in 35.9% (23/64) of the cases.

In part two of the study, there was a moderately strong correlation between PASP determined during RHC and estimated PASP using DE (r = 0.71, P < .01). However, using Bland-Altman analysis of the simultaneous RHC and DE, the mean bias was 8.0 mm Hg and 95% limits of agreement ranged from -28.4 mm Hg to 44.4 mm Hg.

Discussion

In this study of patients with PH, we found that PASP measurements as estimated by DE were inaccurate when compared with PASP determined by RHC. This suggests that DE should not be used to diagnose PH or to determine the response to medical therapy in patients with PH.

The development of quantitative Doppler measurements from TR jets observed on echocardiograms has had a dramatic impact on the clinical diagnosis of PH. In 1984, Yock et al [16] reported that quantitation of the regurgitant tricuspid jet from Doppler measurements to estimate PASP correlated significantly with PASP measured at RHC in 20 patients. Several similar observations were published subsequently, from which it became widely accepted that DE provides an accurate, noninvasive measurement of PASP. [7] . [8] . [9] . [17] As a result, and because of the relative ease and noninvasive nature of echocardiography, DE has now replaced RHC as the method of choice to make hemodynamic measurements in several clinical studies of patients with PH [18] . [19] and even in a National Institutes of Health-sponsored randomized clinical trial. [6] However, it has been shown that a strong correlation between two methods should not be interpreted as indicating an accurate and clinically reliable interchangeability of those same two methods. This common misconception is in part what led Bland and Altman to publish their sentinel paper describing their method of determining the agreement between two methods of clinical measurement. [15]

Our findings in this study confirm those found in previous studies that a moderate correlation does exist between DE and RHC measurements of PASP. [7] . [8] . [17] However, when Bland-Altman analyses are performed to assess the agreement between DE and RHC, the accepted reference standard by which to measure hemodynamics, we find that DE is inaccurate > 50% of the time. Underestimations and overestimations of PASP were both common occurrences using DE, although DE was somewhat more likely to underestimates PASP, consistent with other studies using DE estimates of hemodynamics. [10] Performance of simultaneous DE and RHC measurements of PASP did not alter the findings that DE estimates of PASP in patients with PH are inaccurate. The extent of the imprecision of DE estimates of PASP compared with RHC-based estimates appeared greater at higher PASPs (Fig 1). This confusion over the interchangeability between DE and RHC can lead to inappropriate diagnoses and mismanagement of patients with PH. [10]

Considered a rare disease, the first published incidence of primary PH confirmed by RHC was reported in 1996 [20] and prevalence was estimated at approximately one per 1 million, although the prevalence of PH from all causes was felt to be considerably higher. [21] Recently, hospital discharge data from the Centers for Medicare and Medicaid Services reported 301,400 discharges of patients with a presumed diagnosis of PH in 1997, with a dramatic increase in this figure to 456,500 in 2005. [22] The importance of this remarkable finding is magnified by the fact that physicians commonly use DE in place of RHC in the diagnosis and assessment of PH, which is likely contributing to this increase in PH diagnoses. [23] In addition, we have found that DE is too unreliable to detect changes in PASP of 10 mm Hg, which is the change that occurs from pulmonary vasodilator therapies. [24]

The increase in the number of patients carrying a presumed diagnosis of PH has important implications. The management of the patient with PH is extremely expensive. For example, the annual cost of therapy with the endothelin receptor antagonist, bosentan, a commonly used first-line therapy to treat PH, is $36,208. [4] . [5] . [25] When coupled with the additional expenses of frequent office visits and other testing modalities, the diagnosis of PH becomes very costly. The problem of using DE as an efficacy measure in treating PH patients is underscored by a recent registry on the practices of physicians treating patients with PH, which reported that one-third of patients were receiving combination therapies that have not been shown to be effective. [26] Unfortunately, physicians are frequently unaware of the costs of such medical tests and therapies. [27] It is thus of paramount importance that accurate measurements of hemodynamics be performed in patients when making diagnostic and therapeutic decisions.

There appear to be several identifiable potential pitfalls that can explain the inaccuracy of using DE to determine PASP. Whereas PASP is measured directly during RHC, estimation of PASP using DE requires a two-step process, with each step having its own inherent imprecision. The pressure gradient between the right ventricle (RV) and right atrium during systole must be determined, and this requires the patient to have tricuspid regurgitation. If tricuspid regurgitation is present, a TR
Doppler signal suitable for performing a precise measurement of the peak TR velocity must also be present. Unfortunately, there is substantial variability in the quality of the TR jet in patients with PH, which contributes to imprecise determinations of the peak TR velocity.\(^{[10]}\) Also, the modified Bernoulli equation assumes a perfect, parallel alignment between the Doppler beam and the TR jet. Failure to correct for the angle between the Doppler beam and TR jet can result in an inaccurate determination of the peak TR velocity used to determine the pressure gradient. Even when the TR jet is of sufficient quality and the Doppler beam is aligned optimally, the presence of severe tricuspid regurgitation, a common finding in patients with PH, may result in an underestimation of this pressure gradient.\(^{[10]}\) Determination of this pressure gradient between the RV and right atrium is followed by the addition of an echocardiographic estimate of RAP, using the parameters of IVC size and collapsibility as a surrogate for the directly measured RAP. However, a recent study has shown that RAP estimates using echocardiography are relatively imprecise.\(^{[28]}\) We, too, have confirmed in the present study that echocardiographic estimates of RAP are inaccurate when compared with RAP measured during RHC (Fig 2). Because of the many steps involved in the noninvasive estimation of PASP, each with its own inherent limitations, it is not particularly surprising that we have found using DE estimates of PASP to be an unreliable method in patients with PH.

Despite the unreliability of DE to accurately determine PASP, echocardiography remains a critically important tool in patients with PH by providing an assessment of various indices of RV size and function, many of which have proven to be independent predictors of outcomes in patients with PH.\(^{[29]}\) In addition, echocardiography is invaluable for identifying many of the conditions that can cause secondary PH. Because DE estimates of PASP do correlate moderately well with PASP determined by RHC, the use of echocardiography combined with DE as an initial screening test is warranted.\(^{[30]}\)

Echocardiography should thus be considered complementary to RHC in the evaluation of patients with PH.\(^{[4]} \)\(^{[5]}\)

**Conclusions**

In conclusion, we have confirmed that DE estimates of PASP in patients with PH are inaccurate compared with RHC-based measurements. Echocardiography is a useful screening tool for suspected PH, but the definitive diagnosis and follow-up hemodynamic assessments should be performed with RHC. In particular, DE estimates of PASP should not be relied on when making management decisions.

**Acknowledgments**

**Author contributions:** Dr J. D. Rich: contributed to the design of the study, data collection and analyses, and drafting of the manuscript.

Dr Shah: contributed to data collection and analyses and revisions to the manuscript.

Dr Swamy: contributed to data collection and revisions to the manuscript.

Dr Kamp: contributed to data collection and revisions to the manuscript.

Dr S. Rich: contributed to the design of the study, data collection and analyses, and revisions to the manuscript.

**Financial/nonfinancial disclosures:** The authors have reported to CHEST that no potential conflicts of interest exist with any companies/organizations whose products or services may be discussed in this article.

**REFERENCES:**


Respiratory Society (ERS), endorsed by the International Society of Heart and Lung Transplantation (ISHLT).  
*Eur Heart J* 30. (20): 2493-2537.2009;  Citation


Updated March 5, 2011.


8 Hecht SR, Berger M, Berdoff RL, Van Tosh A, Stimola JM: Use of continuous-wave Doppler ultrasound to evaluate and manage primary pulmonary hypertension.  
*Chest* 90. (5): 781-783.1986;  Abstract

*J Heart Lung Transplant* 26. (9): 927-934.2007;  Abstract

*Am J Respir Crit Care Med* 179. (7): 615-621.2009;  Abstract


*Hepatology* 44. (6): 1502-1510.2006;  Abstract

*Am J Cardiol* 55. (1): 159-163.1985;  Abstract

14 Lang RM, Bierig M, Devereux RB, Chamber Quantification Writing Group, American Society of Echocardiography’s Guidelines and Standards Committee, European Association of Echocardiography, et al: Recommendations for chamber quantification: a report from the American Society of Echocardiography’s Guidelines and Standards Committee and the Chamber Quantification Writing Group, developed in conjunction with the European Association of Echocardiography, a branch of the European Society of Cardiology.  


16 Yock PG, Popp RL: Noninvasive estimation of right ventricular systolic pressure by Doppler ultrasound in patients with tricuspid regurgitation.  
*Circulation* 70. (4): 657-662.1984;  Abstract


18 Sastry BK, Narasimhan C, Reddy NK, Raju BS: Clinical efficacy of sildenafil in primary pulmonary hypertension: a randomized, placebo-controlled, double-blind, crossover study.  
*J Am Coll Cardiol* 43. (7): 1149-1153.2004;  Abstract


*Am J Respir Crit Care Med* 173. (9): 1023-1030.2006;  Abstract


*Chest* 130. (4): 1198-1202.2006;  Full Text


*Chest* 137. (2): 376-387.2010;  Full Text


*J Am Coll Cardiol* 39. (7): 1214-1219.2002;  Abstract

*J Am Coll Cardiol* 43. (12 Suppl S): 40S-47S.2004;  Abstract