Accuracy and precision of echocardiography versus right heart catheterization for the assessment of pulmonary hypertension

Michele D’Altoa,⁎, Emanuele Romee, Paola Argientoa, Antonello D’Andreaa, Rebecca Vanderpoold, Anna Correraa, Eduardo Bossonec, Berardo Sarubbia, Raffaele Calabròa, Maria Giovanna Russoa, Robert Naeijeb

⁎ Corresponding author at: Via Tino di Camaino, 6, 80128 Naples, Italy. Tel.: +39 081 7062501.
E-mail address: mic.dalto@tin.it (M. D’Alto).

a Cardiology, Second University of Naples, Monaldi Hospital, Naples, Italy
b Department of Physiology, Faculty of Medicine, Erasme Academic Hospital, Université Libre de Brussels, Bruxelles, Belgium

c Cardiology and Cardiac Surgery, University Hospital “Scuola Medica Salernitana”, Salerno, Italy

A R T I C L E   I N F O
Article history:
Received 5 March 2013
Received in revised form 23 June 2013
Accepted 1 July 2013
Available online 23 July 2013

Keywords:
Pulmonary hypertension
Hemodynamics
Echocardiography

A B S T R A C T

Background: Echocardiographic studies have contributed to progress in the understanding of the pathophysiology of the pulmonary circulation and have been shown to be useful for screening for and prognostication of pulmonary hypertension, but are considered unreliable for the diagnosis of pulmonary hypertension. We explored this apparent paradox with rigorous Bland and Altman analysis of the accuracy and the precision of measurements collected in a large patient population.

Methods: A total of 161 patients referred for a suspicion of pulmonary hypertension were prospectively evaluated by a Doppler echocardiography performed by dedicated cardiologists within 1 h of an indicated right heart catheterization.

Results: Nine of the patients (6%) were excluded due to an insufficient signal quality. Of the remaining 152 patients, 10 (7%) had no pulmonary hypertension and most others had either pulmonary arterial hypertension (36%) or pulmonary venous hypertension (40%) of variable severities. Mean pulmonary artery pressure, left atrial pressure and cardiac output were nearly identical at echocardiography and catheterization, with no bias and tight confidence intervals, respectively ±3 mm Hg, ±5 mm Hg and ±0.3 L/min. However, the ±2SD limits of agreement were respectively of +19 and −12 mm Hg for left atrial pressure and +1.8 and −1.7 L/min for mean pulmonary artery pressure, ±8 and −12 mm Hg for mean pulmonary artery pressure, ±8 and −12 mm Hg for left atrial pressure and +1.8 and −1.7 L/min for cardiac output.

Conclusions: Doppler echocardiography allows for accurate measurements of the pulmonary circulation, but with moderate precision, which explains why the procedure is valid for population studies but cannot be used for the individual diagnosis of pulmonary hypertension.

© 2013 Elsevier Ireland Ltd. All rights reserved.

1. Background

Doppler echocardiography allows for estimates of pulmonary artery pressure (PAP), left atrial pressure (LAP) and cardiac output (Q), and thus for the calculation of pulmonary vascular resistance (PVR) [1]. Progress in technology and development of portable devices have made it possible to use Doppler echocardiography for the exploration of the effects of environmental stress on the pulmonary circulation, such as exercise [2–4] and hypoxia [5]. Interestingly, in these studies on normal subjects at exercise, which are two conditions of technically more difficult measurements, average slopes of mean PAP (mPAP)-Q plots and derived calculations were strikingly similar to those previously reported by invasive studies [3,4,6–8]. Furthermore, Doppler echocardiography of the pulmonary circulation and the right heart has been repeatedly shown to be of prognostic relevance in patients with pulmonary hypertension, and useful for the screening of patients at risk to develop the disease [1]. However, several studies have concluded that Doppler echocardiography is inaccurate for the diagnosis of pulmonary hypertension [9–13]. The reasons for this apparent paradox are not entirely clear.

We therefore undertook the present prospective study on a larger number of 161 patients referred over a one year period of time for a suspicion of pulmonary hypertension, and who underwent “quasi simultaneous” (within 1 h) a right heart catheterization (RHC) and a trans-thoracic Doppler echocardiography performed by highly trained and dedicated cardiologists. The main question addressed by the study were (1) whether previously reported insufficient diagnostic accuracy of Doppler echocardiography might be improved and, (2) if not, what
are the methodological reasons, including the analysis of the collected results.

2. Methods

The study enrolled prospectively all consecutive patients referred to Pulmonary Hypertension Unit of Monaldi Hospital, Naples, Italy, between 1st June 2011 and 31st May 2012 for a suspicion of pulmonary hypertension who underwent a right heart catheterization. All of them gave an informed consent to the study, which was approved by the Institutional Review Board. The presence of an uncorrected intra- or extra-cardiac shunt, insufficient quality of echo imaging, an estimated systolic pressure at echo < 37 mmHg and absence of additional echocardiographic variable suggestive of pulmonary hypertension (increased velocity of pulmonary valve regurgitation, short right ventricular outflow tract acceleration time, increased dimensions of right heart chambers, abnormal shape and function of the interventricular septum, increased right ventricular (RV) wall thickness, and dilated main pulmonary artery) were considered exclusion criteria.

2.1. Right heart catheterization

Right heart catheterization was performed at rest, without sedation, by two experienced cardiologists (MD and ER), blinded on echo features. Measurements of PAP, right atrial pressure (RAP) and wedged PAP (wPAP) for the LAP were taken at end-expiration. Cardiac output (Q) was measured by thermodilution using an average of at least three measurements. Pulmonary vascular resistance (PVR) was calculated as mPAP minus LAP divided by Q.

2.2. Transthoracic Doppler echocardiography

All the patients had a comprehensive transthoracic Doppler echocardiographic examination within 1 h of the right heart catheterization. The measurements were done by two highly trained cardiologists (PA and AD) using a Philips Sonos 5500 echo machine with a 3.2 MHz transducer (Philips Medical Systems, Andover MA). The procedure was performed following international recommendations [14,15]. Cardiac output (Q) was estimated from left ventricular outflow tract cross sectional area and pulsed Doppler velocity-time integral measurements [16]. Systolic PAP (sPAP) was estimated from a tricuspid regurgitation, as 4 × V2 + 5 mm Hg assigned to right atrial pressure [17]. Mean PAP was calculated as 0.6 × sPAP + 2 [18]. Left atrial pressure (LAP) was estimated from the ratio of Doppler mitral E flow-velocity wave and tissue Doppler mitral annulus flow E′ early diastolic velocity [19]. All data were analyzed off-line by two observers blinded to the patient conditions (EB and AC). Intra-observer and inter-observer variabilities on PAP, LAP and Q were less than 5.0%.

2.3. Statistical analysis

Results are expressed as mean value ± standard deviation (SD). The statistical analysis consisted of least squares linear regression calculations and Student’s t-tests after checking for normality of distributions using the D’Agostino skewness test. Hemodynamic parameters and calculations from echocardiography and catheterization were compared using a Bland and Altman analysis to derive bias, agreement and confidence intervals [20]. Confidence intervals were calculated using the following equation:

\[
\text{mean} \pm 1.96 \times \frac{1}{\sqrt{n}} \sqrt{\frac{1}{n} \sum (x_i - \overline{x})^2}
\]

A p < 0.05 was accepted as statistical significant.

The authors of this manuscript have certified that they comply with the Principles of Ethical Publishing in the International Journal of Cardiology.

3. Results

Nine out of 161 (5.6%) patients were excluded due to insufficient quality of echo imaging. No patient showed indirect echocardiographic signs of pulmonary hypertension in the presence of an estimated sPAP <37 mmHg. The demographics of the other 152 patients are summarized in Table 1. The majority of the patients had pulmonary arterial hypertension (PAH, 36%) or pulmonary venous hypertension (PVH, 40%) on left heart conditions. Ten patients (7%) did not meet the diagnostic criteria for pulmonary hypertension as they had a mPAP of <25 mm Hg at right heart catheterization.

The results of the measurements performed at RHC and Doppler echocardiography are shown in Table 2. There were no significant differences between catheterization and echocardiographic measurements.

The measurements of pulmonary vascular pressures and cardiac output were moderately correlated, with r = 0.67, p < 0.0001 for mPAP and sPAP, r = 0.77, p < 0.0001 for LAP, r = 0.72, p < 0.0001 for Q and r = 0.73, p < 0.0001 for PVR.

The results of the Bland and Altman analysis are shown in Table 3 and illustrated in Figs. 1 to 5. There were no significant biases, between echocardiographic and catheterization measurements of mPAP, LAP and Q, and derived calculation of PVR, and the confidence intervals were on average up to 3 mm Hg for mPAP, 2 mm Hg for LAP, 0.3 L/min for Q and 1 WU for PVR. Thus the echocardiographic measurements compared to catheterization measurements taken as a gold standard were highly accurate. However, the levels of agreement on the difference on the means rather were large, indicating just moderate precision of the echocardiographic measurements compared to gold standard catheterization measurements.

4. Discussion

The present results indicate that trans-thoracic Doppler echocardiography compared to right heart catheterization is accurate and thus allows for valid population studies, but may be insufficiently precise for the diagnosis and estimation of severity of pulmonary hypertension on an individual basis.

Many previous studies aiming at the validation of Doppler echocardiographic measurements of pulmonary vascular pressures and flows relied on correlation calculations [9–11,16,17,19]. However, correlation coefficients largely reflect the variability of the subjects being measured. If one measurement is always twice as big as the other, they are highly correlated but do not agree. Bland and Altman addressed this problem by designing difference versus average plots. This analysis has since become gold standard to compare methods of measurements.
Two crucial information are provided: 1) the bias, or the difference between the means and whether it is constant over the range of measurements, and 2) the limits of agreement, or the range of possible errors. Bias informs about accuracy, and agreement informs about precision. Two previous studies concluded about insufficient accuracy of echocardiography compared to catheterization for the assessment of pulmonary hypertension [12,13]. However, the Bland and Altman plots showed almost no bias but large limits of agreement, rather indicating insufficient precision. The present results obtained on a large number of patients with a wide range of measurements confirm nearly identical means and almost no bias, confirming very good accuracy, both in parameters used in daily practice such as sPAP and in more complex and derived ones such as mPAP, LAP, CO and PVR. However, the limits of agreement appeared to be large, confirming potentially insufficient precision.

Bland and Altman recently deployed that the users of their analysis often quote limits of agreement without their confidence intervals [21]. In the present study we calculated these confidence intervals and found them to vary somewhat over the entire range of measurements, but to be very small, in the range of 2–3 mm Hg on vascular pressures and up to 0.3 L/min on cardiac output, confirming excellent accuracy. However, insufficient precision may be a problem when a single number is used for decision making, such as for example a wPAP > 15 mm Hg for left heart disease. In the present study, we tested the sensitivity of echocardiographic determination of LAP to diagnose PVH defined by a wPAP > 15 mm Hg. The measurement was diagnostic in 46/61 of the patients (75%) using a previously reported equation LAP = 1.24 E/E′ + 1.9 mm Hg [19] and 52/61 (85%) using the linear regression on present data LAP = 1.22 E/E′ + 4.1 mm Hg. Whether such single number misdiagnosis is of real clinical relevance is not certain. Results of a single test are usually integrated in a clinical context, and echocardiography allows for a lot of internal controls [1,15]. For example, an increased mPAP that would not fit into a clinical context of preserved exercise capacity and a normal echocardiography of the right heart can be counterchecked by the analysis of pulmonary flow waves or pulmonary regurgitant jets [1,15].

Agreed statistics are straightforward when one of the two methods of measurement being assessed is a recognized reference, or “gold standard”. There is currently a consensus that right heart catheterization is the method of reference for the assessment of pulmonary hypertension [22,23]. However, routine right heart catheterization relies on the use of fluid-filled catheters, which have an insufficient frequency response [24]. The frequency responses of standard Swan–Ganz catheter–tubing–manometer systems used in clinical practice does not exceed 12 Hz, while more than 50 Hz would be a minimum requirement for instantaneous pressures [24]. Dynamic calibration of catheter–tubing–transducer–flush device systems is often omitted, which exposes to important differences in pressures compared to the true gold standard high-fidelity micromanometer-tipped catheters [25]. Cardiac output in most studies is measured by thermodilution. When compared to the gold standard direct Fick method, thermodilution measurements show little bias, with a mean difference of 0.1 L/min and a confidence interval of 0.2 L/min, indicating excellent accuracy even in the presence

### Table 3

Bias and limits of agreement of echocardiographic versus right heart catheterization estimates of pulmonary vascular pressures and flows in 152 patients.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Bias mean ± SD (95% CI)</th>
<th>Lower limit, mean − SD (95% CI)</th>
<th>Upper limit, mean + SD (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sPAP (mm Hg)</td>
<td>−0.5 ± 9 (−2/−1)</td>
<td>−19 (−22/−17)</td>
<td>18 (15/21)</td>
</tr>
<tr>
<td>mPAP (mm Hg)</td>
<td>−0.5 ± 9 (−2/−1)</td>
<td>−19 (−22/−17)</td>
<td>18 (15/21)</td>
</tr>
<tr>
<td>wPAP (mm Hg)</td>
<td>2 ± 5 (1/3)</td>
<td>−8 (−10/−7)</td>
<td>12 (11/14)</td>
</tr>
<tr>
<td>Q (L/min)</td>
<td>−0.06 ± 0.89</td>
<td>1.84</td>
<td>3.72</td>
</tr>
<tr>
<td>PVR (mm Hg/(L/min))</td>
<td>−0.3 ± 3.2</td>
<td>−6.8 (−7.7/−5.9)</td>
<td>6.2 (5.3/7.1)</td>
</tr>
</tbody>
</table>

Abbreviations: See previous Table 2.
of tricuspid regurgitation, but limits of agreement are ±1 L/min indicating moderate precision [26]. Furthermore, patients with pulmonary hypertension may present with spontaneous variations in up to 22% for mPAP and 36% for PVR within only few hours [27]. In the present study like in previous ones, we sought to limit this cause of decreased agreement by performing the echocardiographic and catheterization procedures almost immediately one after the other (<1 h), but one never knows how much spontaneous variability remains. Thus a routine right heart catheterization may not really offer stable gold standard reference measurements, and actually contributes to the large limits of agreement suggestive of insufficient precision of echocardiographic measurements.

Wedged PAP for the estimation of left ventricular end-diastolic pressure (LVEDP) was recently assessed in a large patient population [28]. The Bland and Altman analysis of the results showed an expected mean bias of −3 mm Hg, which is on average expected [29]. This confirmed previously reported good accuracy of wedged PAP to estimate LAP [30]. However the limits of agreement ranged from −10 to +15 mm Hg, led the authors to warn against an excessive risk misdiagnosis of PVH on the basis of an isolated wPAP measurement [28]. How much this uncertainty accounts for lack of precision of echocardiographic estimates of LAP in the present study is unclear. However these results altogether again underscore the uncertainties of single number cut-off values for decision-making.

While routine right heart catheterization measurements are thus fraught with more than commonly assumed lack of precision, there is of course the variability of the echocardiographic estimates. In the present study the echocardiographic examination were performed by dedicated trained cardiologists. This allowed for a high, 91% recovery rate of a complete set measurement needed to assess the functional state of the pulmonary circulation, and limited the variability inherent to routine examinations. However, each echocardiographic measurement relies of assumptions. The simplified Bernoulli equation to calculate trans-tricuspid pressure gradients may lead to an underestimation of high systolic pressures, as shown in a study which used high-fidelity micromanometer-tipped catheters as a true gold standard for vascular pressure measurements [31]. We did not attempt to estimate RAP from inferior vena cava dimensions and collapsibility [15], which is reputedly imprecise [13], but added a fixed pressure of 5 mm Hg to the calculated trans-tricuspid pressure gradient. In spite of these limitations, it is remarkable that there was no significant pressure related bias on mPAP in the present study. It may be speculated that over-zealous flushing of fluid-filled catheters could have caused an overestimation of systolic pressures due to under-damping, thereby spuriously correcting underestimation by echocardiography. We calculated mPAP from sPAP using a formula derived from high-fidelity catheter measurements [18], and this apparently did not introduce a bias either. The E/E′ ratio to estimate LAP could be criticized on the basis of previously reported too loose correlations [19]. The linear adjustment of our measurements produced a prediction equation of LAP as 1.22 × E/E′ + 4.13 which was similar to previously reported by Nague et al. as 1.24 × E/E′ + 1.9 [19], and the bias in the present study was quite limited. Finally, aortic flow-derived cardiac output is sensitive to errors on left ventricular outflow tract dimensions often resulting in underestimations [16].

Doppler echocardiographic studies of the pulmonary circulation have been recently reported in normal subjects at exercise, which is technically more demanding than in PH patients at rest [15]. The results showed average slopes of multipoint mPAP-Q relationships and derived resistive vessel distensibility calculations that were identical to those previously reported in limited size invasive studies [2–5]. The approach allowed disclosing subtle sex and age-related differences in pulmonary vascular distensibility [3,4] quantified the effects of chronic hypoxic exposure [5], and helped to the understanding of pulmonary vascular function as a limiting factor to exercise capacity [4]. The present results confirm the high accuracy of Doppler echocardiographic measurements, which can therefore be recommended for further population studies. However, the results also confirm a relative lack of precision of Doppler echocardiography on the estimates of each of the components of the PVR equation, so that the procedure cannot be recommended for individual diagnostic decisions based on cut-off values.

References


Fig. 4. Bland–Altman analysis demonstrating near-absence of bias but moderate agreement between echocardiographic and right heart catheterization estimates of cardiac output in 152 patients. The shaded area represent the confidence interval on the difference between the means, and the dashed lines: mean ± 2SD.

Fig. 5. Bland–Altman analysis demonstrating near-absence of bias but moderate agreement between echocardiographic and right heart catheterization estimates of pulmonary vascular resistance in 152 patients. The shaded area represents the confidence interval on the difference between the means, and the dashed lines: mean ± 2SD.


[14] Lang RM, Bierig M, Devereux RB, 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. J Am Soc Echocardiogr 2005;18:1440–63.


