Validation of a Method To Screen for Pulmonary Hypertension in Advanced Idiopathic Pulmonary Fibrosis

David A. Zisman, Arun S. Karlamangla, Steven M. Kawut, Oksana A. Shlobin, Rajeev Saggar, David J. Ross, Marvin I. Schwarz, John A. Belperio, Abbas Ardehali, Joseph P. Lynch, III and Steven D. Nathan

Chest 2008;133;640-645; Prepublished online January 15, 2008; DOI 10.1378/chest.07-2488

The online version of this article, along with updated information and services can be found online on the World Wide Web at: http://chestjournal.org/cgi/content/abstract/133/3/640
Validation of a Method To Screen for Pulmonary Hypertension in Advanced Idiopathic Pulmonary Fibrosis*

David A. Zisman, MD, MS, FCCP; Arun S. Karlamangla, PhD, MD; Steven M. Kawut, MD, MS, FCCP; Oksana A. Shlobin, MD; Rajeev Saggar, MD; David J. Ross, MD; Marvin I. Schwarz, MD, FCCP; John A. Belperio, MD; Abbas Ardehali, MD; Joseph P. Lynch, III, MD; and Steven D. Nathan, MD, FCCP

Background: We have developed a method to screen for pulmonary hypertension (PH) in idiopathic pulmonary fibrosis (IPF) patients, based on a formula to predict mean pulmonary artery pressure (MPAP) from standard lung function measurements. The objective of this study was to validate this method in a separate group of IPF patients.

Methods: Cross-sectional study of 60 IPF patients from two institutions. The accuracy of the MPAP estimation was assessed by examining the correlation between the predicted and measured MPAPs and the magnitude of the estimation error. The discriminatory ability of the method for PH was assessed using the area under the receiver operating characteristic curve (AUC).

Results: There was strong correlation in the expected direction between the predicted and measured MPAPs ($r = 0.72$; $p < 0.0001$). The estimated MPAP was within 5 mm Hg of the measured MPAP 72% of the time. The AUC for predicting PH was 0.85, and did not differ by institution. A formula-predicted MPAP $> 21$ mm Hg was associated with a sensitivity, specificity, positive predictive value, and negative predictive value of 95%, 58%, 51%, and 96%, respectively, for PH defined as MPAP from right-heart catheterization $> 25$ mm Hg.

Conclusions: A prediction formula for MPAP using standard lung function measurements can be used to screen for PH in IPF patients.

Key words: idiopathic pulmonary fibrosis; prediction; pulmonary fibrosis; pulmonary hypertension

Abbreviations: AUC = area under the receiver operating characteristic curve; CI = confidence interval; DLCO = diffusing capacity of the lung for carbon monoxide; IPF = idiopathic pulmonary fibrosis; MPAP = mean pulmonary artery pressure; NPV = negative predictive value; PFT = pulmonary function test; PH = pulmonary hypertension; PPV = positive predictive value; RHC = right-heart catheterization; $\text{SpO}_2$ = resting room air pulse oximetry; UCLA = University of California, Los Angeles

Pulmonary hypertension (PH) frequently complicates advanced idiopathic pulmonary fibrosis (IPF) and is associated with poor outcome.1–7 Currently, right-heart catheterization (RHC) is the only accepted method for the diagnosis of PH in patients with IPF. However, RHC is invasive and expensive. Although echocardiography and CT-determined main pulmonary artery diameter are commonly used tests to screen for PH in patients with IPF, they are not reliable.4,5,6 Reliable, noninvasive approaches to the diagnosis of PH in patients with IPF would improve patient safety, reduce costs, and enable the appropriate timing of RHC.

We recently demonstrated that the ratio of the FVC percentage of predicted to DLCO percentage of predicted and room air resting pulse oximetry ($\text{SpO}_2$) data can be combined in a linear regression formula to screen for PH in patients with IPF.4 It was shown that a cutoff of 25 mm Hg for the formula-estimated mean pulmonary artery pressure (MPAP) had sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for PH (defined as mean pulmonary artery pressure [MPAP] from RHC $> 25$ mm Hg) of 71%, 81%, 71%, and 81%, respectively. By selecting a lower cutoff of 21 mm Hg for the
formula-estimated MPAP, we maximized sensitivity (100%) for PH (defined as MPAP from RHC > 25 mm Hg) with the least compromise in specificity (40%). The performance of the formula was assessed by bootstrap techniques; however, internal validation does not guarantee adequate performance in other populations. Hence, independent external validation is essential before recommendations can be made for adoption in clinical practice. Accordingly, the aim of this study was to validate the PH screening formula in an external population of IPF patients.

Materials and Methods

Validation Sample

We reviewed the medical records of all IPF patients from the Inova Fairfax Hospital between July 1997 and February 2007 and from University of California, Los Angeles (UCLA) Medical Center between July 2006 (following the close of the derivation study) and June 2007. Hence, the group of patients in this study is totally separate from that used to develop the formula. The Inova Fairfax Hospital and UCLA institutional review boards approved the study. All patients met accepted diagnostic criteria for IPF, and the majority (71%) had histopathologic evidence of usual interstitial pneumonia. One hundred thirty-two IPF patients were candidates for inclusion in this study. To be included in the study, participants had to have had RHC and have pulmonary function test (PFT) and SpO2 data while breathing room air within 3 months of the RHC. All RHCs were performed as part of standard lung transplant evaluation. Patients were excluded for the following reasons: (1) missing data (46 patients), and (2) PFT or SpO2 not done within 3 months of RHC (26 patients). Sixty patients met the entry criteria and comprised the validation cohort (Inova Fairfax Hospital, 35 patients; UCLA, 25 patients).

Measurements

We defined PH as resting MPAP from RHC > 25 mm Hg. Pulmonary artery occlusion pressure < 15 mm Hg and pulmonary vascular resistance > 3 Wood units were not required to define PH because these measurements do not provide prognostic information above and beyond MPAP in IPF patients. Hence, we selected MPAP > 25 mm Hg from RHC as the outcome to screen for with our method and later confirm with RHC (when the rest of the hemodynamic variables would become available). Standard methodology was used for PFTs and pulse oximetry. After at least 5 min of rest, SpO2 was measured on room air. The equations of Crapo et al were used to calculate predicted FVC values. The equations of Crapo and Morris were used to calculate predicted DLCO values. To assess the impact of alternate predicted values on the discriminatory capacity of the method, we tested additional equations for predicted FVC and DLCO. The following equation, derived by our group, was used to calculate predicted MPAP (in millimeters of mercury):

$$\text{MPAP} = 11.9 + 0.272 \times \text{SpO2} + 0.00659 \times (100 - \text{SpO2})^2 + 3.06 \times (\text{percentage of predicted FVC/percentage of predicted DLCO}).$$

Statistical Analysis

Our objectives were to test the utility of the MPAP estimation formula by assessing both the accuracy of the MPAP prediction and the reliability of the PH prediction. First, the quality of the MPAP prediction was assessed by examining the percentage of MPAP estimates that fell within 5 and 10 mm Hg of the MPAP measured by directly RHC. This was also examined separately in the UCLA and Inova Fairfax Hospital samples. The Pearson correlation coefficient between the predicted MPAP and RHC-measured MPAP was examined. Second, the discriminatory ability of the PH-prediction method was assessed using the area under the receiver operating characteristic curve (AUC). The AUC was also examined separately in the UCLA and Inova Fairfax Hospital samples, and we tested for a difference in the two AUC figures using the method published by Hanley and McNeil. All tests were two-tailed, and p values < 0.05 were required for statistical significance. All statistical analysis was performed using statistical software (SAS version 9.1; SAS Institute; Cary, NC; and MedCalc for Windows, version 9.2.0.0; MedCalc Software; Mariakerke, Belgium).

Results

Comparisons of Patients in the Validation and Excluded Cohorts

Patients in the validation sample (n = 60) were younger and had more advanced pulmonary disease than those excluded from the study (n = 72) but were similar to the rest of the cohort with respect to gender, race, and SpO2 (Table 1). This is consistent
with the fact that RHC is offered primarily to patients with advanced disease undergoing lung transplant evaluation.

**Accuracy of Predicted MPAP**

There was a strong correlation in the expected direction between the predicted MPAP and the RHC-measured MPAP ($r = 0.72$, $p < 0.0001$; Fig 1). For the entire cohort, MPAP prediction was accurate (within 10 mm Hg of RHC measurement) in 93% (95% confidence interval [CI], 84 to 97%) of the patients. Moreover, 72% (95% CI, 60 to 82%) of the predicted MPAPs were within 5 mm Hg of RHC-measured MPAP. The accuracy of the estimated MPAP in the validation sets from Inova Fairfax Hospital and UCLA were similar.

**Performance Characteristics of the PH Predictor**

The performance of the PH-prediction formula was excellent and very similar to that in the original report (Table 2). The NPV of the PH predictor is high: 83% of patients with a predicted MPAP $\leq$ 25 mm Hg did not have PH (defined as MPAP from RHC $> 25$ mm Hg), while 96% of patients with predicted MPAP $\leq$ 21 mm Hg did not have PH (defined as MPAP from RHC $> 25$ mm Hg). Varying the cutoff point of the predicted MPAP changed the sensitivity, specificity, PPV, and NPV of the prediction formula. At one extreme, for a formula-predicted MPAP $\leq$ 19 mm Hg, the NPV was 100% and the PPV was 90% (for PH defined as MPAP from RHC $> 25$ mm Hg); at the other extreme, for a formula-predicted MPAP $\leq$ 35 mm Hg, NPV was 73% and the PPV was 100% (for PH defined as MPAP from RHC $> 25$ mm Hg). The NPV remained uniformly high ($\geq$ 73%) for various scenarios.

Discrimination is a measurement of the ability to distinguish between patients who do and do not have the outcome of interest, in this case PH. The discriminatory ability of the PH-prediction method was assessed using the AUC (Fig 2). A value of 1.0 indicates that the model perfectly discriminates between patients with different outcomes, while a value of 0.5 indicates that the model contains no predictive information. The discriminatory ability of the PH-prediction method was excellent reflected by an AUC of 0.85 (95% CI, 0.73 to 0.93), which was similar to the AUC in the derivation cohort (0.83). Furthermore, the discriminatory properties of the PH predictor in the validation sets from Inova Fairfax Hospital and UCLA patients were similar (AUC = 0.82 and AUC = 0.80, respectively; $p = 0.4$). The AUCs were virtually identical when alternate prediction equations were used to calculate predicted FVC and DLCO values (data not shown).

**Discussion**

We validated a recently developed clinical prediction formula for MPAP in advanced IPF patients. This instrument could serve as a screening method for PH in these patients. The prediction method uses a parsimonious set of readily available physiologic variables ($\text{SpO}_2$, percentage of predicted FVC, and percentage of predicted DLCO) to identify individu-

![Figure 1. Relationship between predicted MPAP and RHC-measured MPAP ($r = 0.72$, $p < 0.0001$).](image-url)
als with a low likelihood of PH, thereby reducing the need for RHC. A screening test will of necessity have an increased ability to avoid missing a true case (sensitivity) at the expense of an increase in the number of individuals without the disease who will erroneously be picked up by the screening program (specificity). Thus, altering the criterion of positivity will influence both the sensitivity and specificity of the test. Using a cutoff of 21 mm Hg from the formula (as proposed in the derivation study), there was high sensitivity and high NPV (95% and 96%, respectively) for PH (defined as MPAP from RHC > 25 mm Hg). Hence, if the formula generates a value ≤ 21 mm Hg, only 5% of IPF patients with PH (defined as MPAP from RHC > 25 mm Hg) will be “missed.” The lower the cutoff of the formula-predicted MPAP, the less the likelihood that the patient will have PH. For instance, if the formula generates a value ≤ 19 mm Hg, no IPF patient with PH (defined as MPAP from RHC > 25 mm Hg) will be missed. Importantly, the high NPV of the method is not because of low prevalence of PH in our study sample. In fact, PH was present in 32% of the sample, a figure comparable to that reported by others.1,4,6,7,25

Using a cutoff of 21 mm Hg from the formula, however, the specificity and PPV of the method for PH are low: only 51% of patients with a formula-predicted MPAP > 21 mm Hg will have PH (defined as MPAP from RHC > 25 mm Hg). However, it is generally accepted in clinical medicine that a screening test meant to be used on all comers needs to have high sensitivity and NPV, and is typically followed by a more specific confirmatory test with high PPV (eg, following a positive antinuclear antibody test result for systemic lupus erythematosus with the anti-DNA antibody test). If the new PH predictor is used as a screening test, then confirmatory testing (RHC) will be required in patients with a formula-predicted MPAP > 21 mm Hg in order to confirm and quantify the presence of PH and determine other important hemodynamic variables (pulmonary artery occlusion pressure, cardiac output, pulmonary vascular resistance). However, given that confirmatory RHC offers risk and there are no proven therapies for PH in IPF, clinicians may prefer to miss some PH cases rather than offer more RHC; hence, a higher cutoff (eg, formula-predicted MPAP > 25 mm Hg) could be selected.

We expect that this instrument will be used in clinical practice to identify IPF patients who need not undergo invasive assessment for PH. Clinicians can be confident that patients with low values of formula-predicted MPAP are unlikely to have PH by RHC. Furthermore, since outside of lung transplant evaluation, RHCs are not routinely performed in IPF patients, our method could be utilized as an exclusion criterion (formula-predicted MPAP ≤ 21

Table 2—Performance Characteristics of the Method in Establishing or Excluding a Diagnosis of PH Defined as MPAP From RHC > 25 mm Hg

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Sensitivity (95% CI), %*</th>
<th>Specificity (95% CI), %*</th>
<th>PPV (95% CI), %*</th>
<th>NPV (95% CI), %*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formula-predicted MPAP &gt; 19 mm Hg</td>
<td>100 (66–100)</td>
<td>25 (13–41)</td>
<td>39 (13–41)</td>
<td>100 (66–100)</td>
</tr>
<tr>
<td>Formula-predicted MPAP &gt; 21 mm Hg</td>
<td>95 (74–99)</td>
<td>58 (41–73)</td>
<td>51 (35–70)</td>
<td>96 (78–100)</td>
</tr>
<tr>
<td>Validation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Derivation, %†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formula-predicted MPAP &gt; 25 mm Hg</td>
<td>63 (38–84)</td>
<td>85 (70–94)</td>
<td>67 (36–86)</td>
<td>83 (63–89)</td>
</tr>
<tr>
<td>Validation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Derivation, %†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formula-predicted MPAP &gt; 35 mm Hg</td>
<td>21 (6–46)</td>
<td>100 (91–100)</td>
<td>100 (22–99)</td>
<td>73 (60–82)</td>
</tr>
</tbody>
</table>

*PH was considered to be present when MPAP by RHC was > 25 mm Hg.
†For comparison purposes, the performance of the predictor in the original derivation cohort is also listed.4
PH is more critical. Younger patients may have more occult collagen vascular disease with autoimmune vasculopathy rather than the hypoxic vasoconstriction or destruction of vessels that may play a larger role in older IPF patients. However, given that the mean age difference between our study and excluded cohorts was relatively small, we do not believe that age differences biased our findings. Since the RHC in this study were not recorded with exercise, we could not assess for exercise-induced PH. However, the implications and clinical importance of exercise-induced PH are unknown. The cross-sectional nature of the study should also be noted. A low formula-predicted MPAP during a single evaluation does not rule out the possibility of PH developing in the future. However, the formula-predicted MPAP can be followed on a serial basis because it is computed using clinical variables that are routinely measured.

This study is important for three reasons. First, it establishes the empirical validity of a new, easy-to-use, clinical screening method for PH in advanced IPF patients. Second, this study shows that the formula can be applied in advanced IPF populations from different geographic areas/medical centers, that is, the method is transportable. Third, a cutoff of formula-predicted of 21 mm Hg has high sensitivity and NPV for PH (defined as MPAP from RHC > 25 mm Hg), so as to justify withholding RHC from patients predicted least likely to have PH (formula-predicted MPAP = 21 mm Hg). This would prevent unnecessary risk, discomfort, and cost to this group of patients.

In conclusion, the recently developed PH screening method in IPF is valid and transportable with a high NPV and good discriminatory ability in patients with IPF. The use of this screening method will prevent unnecessary RHC.

REFERENCES

7 Nathan SD, Noble PW, Tudor RM. Idiopathic pulmonary fibrosis and pulmonary hypertension: connecting the dots. Am J Respir Crit Care Med 2007; 175:875–880
9 Zisman DA, Karlamangla AS, Ross DJ, et al. High-resolution chest computed tomography findings do not predict the...
21 Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1985; 161:189–196
22 Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982; 143:29–36
Validation of a Method To Screen for Pulmonary Hypertension in Advanced Idiopathic Pulmonary Fibrosis
David A. Zisman, Arun S. Karlamangla, Steven M. Kawut, Oksana A. Shlobin, Rajeev Saggar, David J. Ross, Marvin I. Schwarz, John A. Belperio, Abbas Ardehali, Joseph P. Lynch, III and Steven D. Nathan
Chest 2008;133:640-645; Prepublished online January 15, 2008; DOI 10.1378/chest.07-2488

This information is current as of May 5, 2008

Updated Information & Services
Updated information and services, including high-resolution figures, can be found at:
http://chestjournal.org/cgi/content/full/133/3/640

References
This article cites 25 articles, 18 of which you can access for free at:
http://chestjournal.org/cgi/content/full/133/3/640#BIBL

Citations
This article has been cited by 1 HighWire-hosted articles:
http://chestjournal.org/cgi/content/full/133/3/640

Permissions & Licensing
Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at:
http://chestjournal.org/misc/reprints.shtml

Reprints
Information about ordering reprints can be found online:
http://chestjournal.org/misc/reprints.shtml

Email alerting service
Receive free email alerts when new articles cite this article sign up in the box at the top right corner of the online article.

Images in PowerPoint format
Figures that appear in CHEST articles can be downloaded for teaching purposes in PowerPoint slide format. See any online article figure for directions.