

## ORIGINAL ARTICLE

# Quantitative Evidence for Revising the Definition of Primary Graft Dysfunction after Lung Transplant

Edward Cantu<sup>1</sup>, Joshua M. Diamond<sup>2</sup>, Yoshikazu Suzuki<sup>1</sup>, Jared Lasky<sup>1</sup>, Christian Schaufler<sup>1</sup>, Brian Lim<sup>1</sup>, Rupal Shah<sup>2</sup>, Mary Porteous<sup>2</sup>, David J. Lederer<sup>3</sup>, Steven M. Kawut<sup>2,4,5</sup>, Scott M. Palmer<sup>6</sup>, Laurie D. Snyder<sup>6</sup>, Matthew G. Hartwig<sup>7</sup>, Vibha N. Lama<sup>8</sup>, Sangeeta Bhorade<sup>9</sup>, Christian Bermudez<sup>1</sup>, Maria Crespo<sup>2</sup>, John McDyer<sup>10</sup>, Keith Wille<sup>11</sup>, Jonathan Orens<sup>12</sup>, Pali D. Shah<sup>12</sup>, Ann Weinacker<sup>13</sup>, David Weill<sup>14</sup>, David Wilkes<sup>15</sup>, David Roe<sup>15</sup>, Chadi Hage<sup>15</sup>, Lorraine B. Ware<sup>16,17</sup>, Scarlett L. Bellamy<sup>18</sup>, and Jason D. Christie<sup>2,4</sup>; for the Lung Transplant Outcomes Group

<sup>1</sup>Division of Cardiovascular Surgery and <sup>2</sup>Pulmonary, Allergy, and Critical Care Division, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania; <sup>3</sup>Division of Pulmonary, Allergy, and Critical Care Medicine, Columbia University College of Physicians and Surgeons, New York, New York; <sup>4</sup>Center for Clinical Epidemiology and Biostatistics and <sup>5</sup>Penn Cardiovascular Institute, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania; <sup>6</sup>Division of Pulmonary, Allergy, and Critical Care Medicine and <sup>7</sup>Division of Cardiothoracic Surgery, Duke University, Durham, North Carolina; <sup>8</sup>Division of Pulmonary, Allergy, and Critical Care Medicine, University of Michigan, Ann Arbor, Michigan; <sup>9</sup>Division of Pulmonary and Critical Care Medicine, University of Chicago, Chicago, Illinois; <sup>10</sup>Division of Pulmonary, Allergy, and Critical Care, University of Pittsburgh, Pittsburgh, Pennsylvania; <sup>11</sup>Division of Pulmonary and Critical Care Medicine, University of Alabama at Birmingham, Birmingham, Alabama; <sup>12</sup>Division of Pulmonary, Allergy, and Critical Care Medicine, Department of Medicine, Johns Hopkins University Hospital, Baltimore, Maryland; <sup>13</sup>Division of Pulmonary and Critical Care Medicine, Stanford University, Palo Alto, California; <sup>14</sup>Institute for Advanced Organ Disease and Transplantation, University of South Florida, Tampa, Florida; <sup>15</sup>Division of Pulmonary, Allergy, Critical Care, and Occupational Medicine, Indiana University School of Medicine, Indianapolis, Indiana; <sup>16</sup>Department of Medicine and <sup>17</sup>Department of Pathology, Microbiology, and Immunology, Vanderbilt University, Nashville, Tennessee; and <sup>18</sup>Dornsife School of Public Health, Drexel University, Philadelphia, Pennsylvania

ORCID IDs: 0000-0003-1105-0129 (E.C.); 0000-0001-5258-0228 (D.J.L.).

## Abstract

**Rationale:** Primary graft dysfunction (PGD) is a form of acute lung injury that occurs after lung transplantation. The definition of PGD was standardized in 2005. Since that time, clinical practice has evolved, and this definition is increasingly used as a primary endpoint for clinical trials; therefore, validation is warranted.

**Objectives:** We sought to determine whether refinements to the 2005 consensus definition could further improve construct validity.

**Methods:** Data from the Lung Transplant Outcomes Group multicenter cohort were used to compare variations on the PGD definition, including alternate oxygenation thresholds, inclusion of additional severity groups, and effects of procedure type and mechanical ventilation. Convergent and divergent validity were compared for mortality prediction and concurrent lung injury biomarker discrimination.

**Measurements and Main Results:** A total of 1,179 subjects from 10 centers were enrolled from 2007 to 2012. Median length of follow-up was 4 years (interquartile range = 2.4–5.9). No mortality differences were noted between no PGD (grade 0) and mild PGD (grade 1). Significantly better mortality discrimination was evident for all definitions using later time points (48, 72, or 48–72 hours;  $P < 0.001$ ). Biomarker divergent discrimination was superior when collapsing grades 0 and 1. Additional severity grades, use of mechanical ventilation, and transplant procedure type had minimal or no effect on mortality or biomarker discrimination.

**Conclusions:** The PGD consensus definition can be simplified by combining lower PGD grades. Construct validity of grading was present regardless of transplant procedure type or use of mechanical ventilation. Additional severity categories had minimal impact on mortality or biomarker discrimination.

**Keywords:** lung transplant; primary graft dysfunction; lung transplant outcomes

(Received in original form June 9, 2017; accepted in final form September 1, 2017)

Correspondence and requests for reprints should be addressed to Edward Cantu, M.D., Division of Cardiovascular Surgery, University of Pennsylvania Perelman School of Medicine, 3400 Spruce Street, 6 Silverstein Pavilion, Philadelphia, PA 19104. E-mail: [edward.cantu@uphs.upenn.edu](mailto:edward.cantu@uphs.upenn.edu).

This article has an online supplement, which is accessible from this issue's table of contents at [www.atsjournals.org](http://www.atsjournals.org).

Am J Respir Crit Care Med Vol 197, Iss 2, pp 235–243, Jan 15, 2018

Copyright © 2018 by the American Thoracic Society

Originally Published in Press as DOI: 10.1164/rccm.201706-11400C on September 5, 2017

Internet address: [www.atsjournals.org](http://www.atsjournals.org)

**At a Glance Commentary**

**Scientific Knowledge on the Subject:**

The International Society for Heart and Lung Transplantation (ISHLT) consensus definition developed in 2005 has been instrumental in enhancing the transplant community’s ability to study and understand primary graft dysfunction (PGD) after lung transplantation. Since that time, questions have arisen about how best to improve the definition to ensure adequate construct validity with respect to timing of grading, oxygen exchange cutoffs, transplant type, and use of mechanical ventilation.

**What This Study Adds to the Field:**

We tested the construct validity of varying PGD definitions for mortality and biomarker discrimination using a multicenter, prospective cohort study collected and curated by the Lung Transplant Outcomes Group. We determined that the 2005/2017 ISHLT consensus definition can be simplified by combining lower PGD grades. In addition, construct validity of grading was maintained, regardless of transplant procedure type or use of mechanical ventilation. Lastly, additional higher severity categories had minimal impact on mortality or biomarker discrimination.

Primary graft dysfunction (PGD) is a form of acute respiratory distress syndrome (ARDS) developing during the immediate postoperative period after lung transplantation. PGD significantly contributes to increased short-term mortality and morbidity (1–5). Patients who develop PGD also show significantly worse long-term outcomes (1, 3, 6, 7). The

currently employed four-tiered PGD grading system was developed in 2005 by the ISHLT (International Society for Heart and Lung Transplantation) Working Group, and is currently under reconsideration by an updated ISHLT Working Group (8–12). At the time of publication in 2005, the ISHLT Working Group noted the importance of construct validity in a PGD consensus definition, and called for appropriate future definition refinements (13).

The original ISHLT PGD definition paralleled the AECC (American–European Consensus Conference) definition of ARDS (14). The “Berlin” definition, a recent refinement to the AECC ARDS definition, was developed to address the limitations of acuity assessments, effects of ventilator settings on oxygenation, and difficulties with interpretation of chest radiographs that emerged over the subsequent 18 years since its standardization (15). Similar concerns have emerged with the ISHLT PGD definition; over the past 12 years, multiple suggestions for alterations to the PGD grading system have been made,

including modifications to the PaO<sub>2</sub>/FiO<sub>2</sub> (P/F) cutoffs, optimal timing of grading, effects of mechanical ventilation, including positive end-expiratory pressure, as well as differences according to transplant procedure type (13, 16–22). In addition, although the 2005 definition has been used to enhance understanding of PGD risks, mechanisms, and development of new therapies, practice has changed since that time, whereas the updated 2017 ISHLT definition has remained relatively unchanged (23, 24). In recent years, PGD has been considered as the primary outcome by governing bodies as a response indicator variable for new therapies aimed at improving lung transplant outcomes (23, 25).

The goal of this study was to establish the validity of PGD grading in the post–Lung Allocation Score era, to consider refinements to the definition based on differing thresholds of oxygenation, and to determine the effects of procedure type and use of mechanical ventilation on PGD grade. As there is no “gold standard” for PGD, we used convergent and

P/F ratio	≥ 300 – CXR	≥ 300 + CXR	299–200 + CXR	199–100 + CXR	<100 + CXR
ISHLT	No PGD Grade 0	Mild Grade 1	Moderate Grade 2	Severe Grade 3	
Revised	No PGD Grade 0/1		Moderate Grade 2	Severe Grade 3	
Berlin	No PGD Grade 0/1		Moderate Grade 2	Severe Grade 3	Very Severe Grade 4

**Figure 1.** Definition of primary graft dysfunction (PGD). Depicted are the PaO<sub>2</sub>/FiO<sub>2</sub> (P/F) thresholds for the three variations of the definitions tested. CXR = chest X-ray; ISHLT = International Society for Heart and Lung Transplantation.

Supported by NIH grants HL087115, HL081619, HL096845, HL115354, HL114626, HL121406, HL116656, HL126176, and HL135227 and Robert Wood Johnson grant AMFDP 70640.

Author Contributions: Conception and design—E.C. and J.D.C.; acquisition of data—E.C., J.M.D., Y.S., J.L., C.S., B.L., R.S., M.P., D.J.L., S.M.K., S.M.P., L.D.S., M.G.H., V.N.L., S.B., C.B., M.C., J.M., K.W., J.O., P.D.S., A.W., D. Weill, D. Wilkes, D.R., C.H., L.B.W., S.L.B., and J.D.C.; analysis and interpretation of data—E.C., J.M.D., S.L.B., and J.D.C.; drafting or revising the manuscript for important intellectual content—E.C., J.M.D., Y.S., J.L., C.S., B.L., R.S., M.P., D.J.L., S.M.K., S.M.P., L.D.S., M.G.H., V.N.L., S.B., C.B., M.C., J.M., K.W., J.O., P.D.S., A.W., D. Weill, D. Wilkes, D.R., C.H., L.B.W., S.L.B., and J.D.C.; final approval of the version to be published—E.C., J.M.D., Y.S., J.L., C.S., B.L., R.S., M.P., D.J.L., S.M.K., S.M.P., L.D.S., M.G.H., V.N.L., S.B., C.B., M.C., J.M., K.W., J.O., P.D.S., A.W., D. Weill, D. Wilkes, D.R., C.H., L.B.W., S.L.B., and J.D.C.

divergent validity techniques, focusing on mortality discrimination and established lung injury biomarkers. Some of the results of these studies have been previously reported in the form of an abstract (26).

## Methods

### Study Populations and Design

Subjects were selected from the LTOG (Lung Transplant Outcomes Group) cohort, which is a multicenter prospective cohort study of lung transplant recipients designed to study PGD and its outcomes (NCT00457847) (24). Institutional review board approval and written informed consent were obtained before the recruitment of subjects. We tested the construct that mortality and lung injury biomarkers would diverge or converge for different PGD grades (divergent and convergent validity). Two distinct populations were defined from this cohort to address divergent validity separately for mortality and for lung injury biomarkers, as not all subjects had biomarkers measured. The first population used to assess mortality discrimination was composed of consecutive subjects transplanted between 2007 and 2012 from 10 centers who had PGD grading at each time point (0, 24, 48, and 72 h) with follow-up until March 2017 (see Figure E1 in the online supplement). The second group was composed of consecutive subjects transplanted between 2002 and 2006 from six centers with available biomarker data, which were used to gauge convergent and discriminant validity between PGD definitions (27). Clinical data were collected prospectively, as previously described (24). Multiple organ transplant and retransplant recipients were excluded (Figure E1).

### Definitions

We used the ISHLT consensus definition of PGD (8) and considered alternative definitions based on P/F ratio (Figure 1). Primarily, we tested combining lower grades (P/F > 300) precluding the need for chest radiograph assessments, as well as adding additional severe grades (P/F < 100), and determining the effects of presence or absence of mechanical ventilation, transplant procedure type, and timing of grading (additional detail on definitions provided in the online supplement).

### PGD Biomarkers

In previous studies, several biomarkers have been reported to be associated with PGD (28–32). Of the reported biomarkers, plasminogen activator inhibitor (PAI)-1 was found to be the most strongly associated with severe PGD, and was thus chosen as our primary marker for discrimination (27). Plasma samples were prospectively collected 24 hours after transplantation in a subset of LTOG participants between 2002 and

2006. Samples were centrifuged within 60 minutes and then stored at  $-80^{\circ}\text{C}$  for subsequent analysis, and plasma levels of PAI-1 were determined using commercially available ELISAs, as previously described (27, 33).

### Statistical Analysis

The predictive (discriminant) validity for 30-day mortality of each definition was compared with receiver operating curve analysis in logistic regression models using

**Table 1.** Demographics

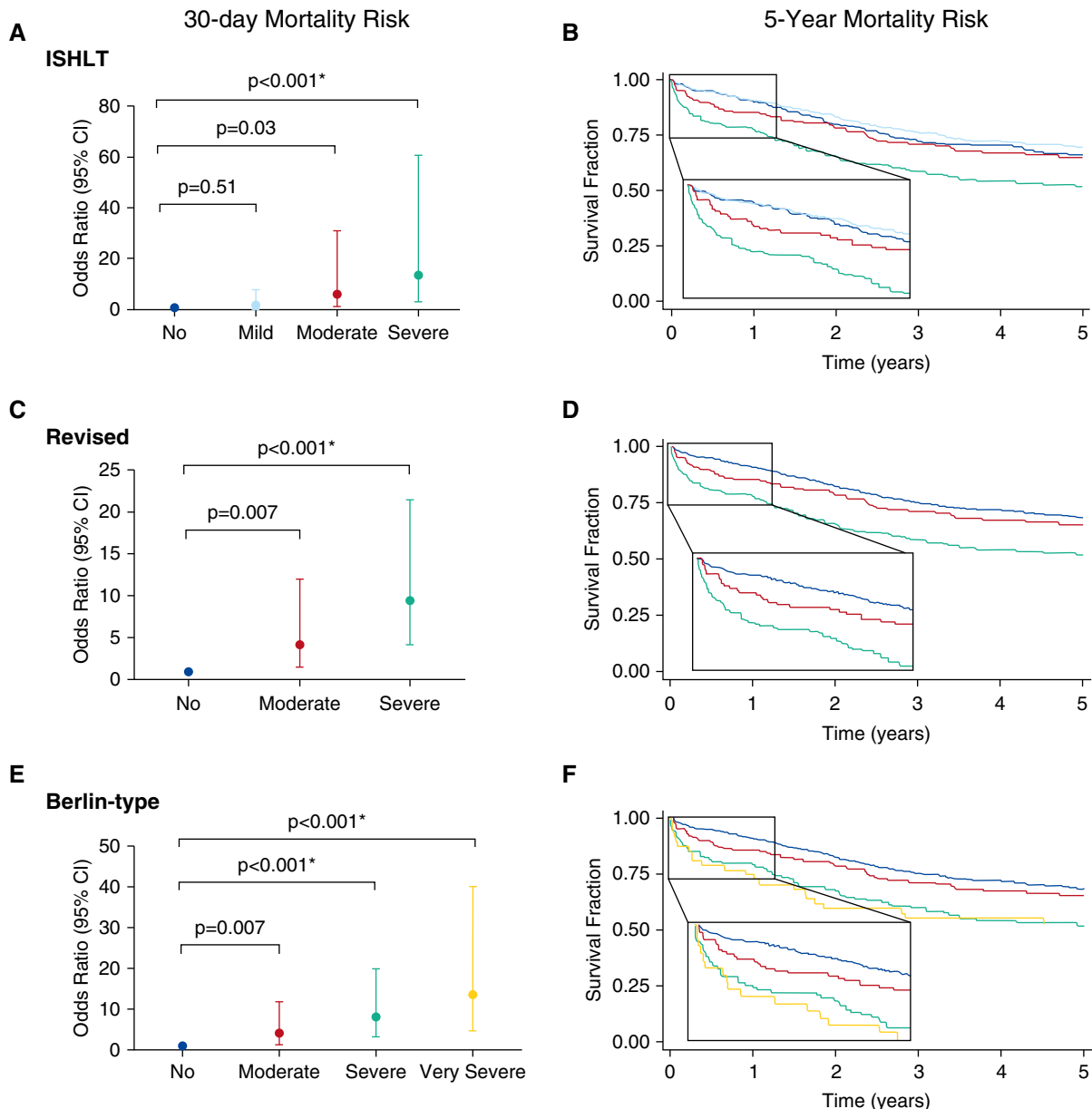
	Overall Population (n = 1,179)
Donor	
Age, mean (SD), yr	34.8 (14.1)
Female, %	36.6
Ethnicity, %	
Caucasian	60.3
African American	23.5
Hispanic	13.1
Other	3.1
Ventilatory support, mean (SD), d	3.1 (2.3)
Pa <sub>O<sub>2</sub></sub> , mean (SD), mm Hg	492 (77.2)
Smoking history, %	30.2
BMI, mean (SD), kg/m <sup>2</sup>	26.1 (5.8)
Mode of death, %	
Blunt trauma	4.3
Head trauma	37.8
Stroke	36.6
Anoxia	13.2
Other	8.0
Recipient	
Age, mean (SD), yr	54.7 (13.3)
Female, %	42.0
Ethnicity, %	
Caucasian	86.3
African American	8.5
Hispanic	3.1
Other	2.2
BMI, mean (SD), kg/m <sup>2</sup>	25.0 (4.8)
Diagnosis, %	
COPD	34.1
Idiopathic pulmonary fibrosis	40.0
Cystic fibrosis	14.3
Pulmonary arterial hypertension	3.2
Other	8.4
Lung allocation score	44.8 (14.9)
Operative	
Transplant type, %	
Single	29.6
Bilateral	70.4
CPB used, %	36.2
CPB time, mean (SD), min	220.9 (81.4)
Intraoperative nitric oxide, %	44.7
Total ischemic time, mean (SD), min	333.4 (103.8)
Outcomes	
Mechanical ventilation, mean (SD), d	4.1 (18.0)
Ventilator-free days, mean (SD)	25.7 (18.1)
Hospital length of stay, mean (SD), d	23.3 (24.9)
Total survival, mean (SD), d	1,098.8 (619.9)

*Definition of abbreviations:* BMI = body mass index; COPD = chronic obstructive pulmonary disease; CPB = cardiopulmonary bypass.

dummy variables for defined PGD categories based on P/F variation (34). Discrimination of overall survival between the different grades was assessed using Kaplan-Meier and multivariable Cox proportional hazard methods, also employing grades as dummy variables. Models were fit using potential confounding variables previously associated

with PGD and/or mortality (donor smoking history, recipient diagnosis, recipient body mass index, recipient pulmonary artery pressure, transplant type, cardiopulmonary bypass use, and recipient reperfusion  $Fi_{O_2}$ ) (24, 35, 36). Analyses of effects of interval timing, transplant type, and mechanical ventilator status on discrimination were similarly compared

using logistic regression and receiver operating curve comparison using the “roccomp” and “rocgold” commands in STATA v12.1 software (STATA Corp.) (37). Tests for interaction between transplant type, ventilator status, and PGD grade on mortality were performed using multiplicative interaction terms in the regression framework. Center effects were



**Figure 2.** Risk of 30-day and 5-year mortality. Summarized in the left-most panels are the odds ratio point estimates and 95% confidence intervals (CIs) for 30-day mortality obtained from logistic regression for each definition at the 48- to 72-hour intervals for the (A) International Society for Heart and Lung Transplantation (ISHLT), (C) revised, and (E) Berlin-type definitions. The asterisk denotes significant increase of overall mortality in Cox models adjusted for donor smoking history, recipient diagnosis, body mass index, pulmonary artery pressure, transplant type, cardiopulmonary bypass use, and reperfusion  $Fi_{O_2}$ . In the right-most panels, Kaplan-Meier survival curves demonstrate unadjusted 5-year survival for the (B) ISHLT, (D) revised, and (F) Berlin-type definitions.

evaluated using conditional logistic regression. One-way ANOVA was used for comparison of biomarker concentrations at 24 hours between each grade. In all analyses, a two-sided  $P$  value of 0.05 was considered statistically significant.

## Results

### Characteristics of Mortality Discrimination Cohort

A total of 1,179 subjects was included from the LTOG cohort, had PGD grading at each time point (0, 24, 48, and 72 h), and were evaluated for mortality discrimination. Demographics of this cohort are presented in Table 1. Follow-up was completed in 1,179 subjects (100%) with a median length of follow-up of 4 years (interquartile range = 2.4–5.9). Survival at 90 days, 6 months, and 1, 3, and 5 years was 94%, 92%, 88%, 72%, and 66%, respectively (Figure E2).

The incidence of PGD severity by P/F ratio at discrete time points and during specified intervals is displayed in Figure E3. There was decreasing incidence of higher-grade PGD with grading at later time points. As a consequence, the proportion of transplanted patients with a P/F ratio greater than 300 steadily increased with time after transplant, and was greater than 0.8 by the 72-hour time point. Specific incidences of PGD using the 72-hour, 0- to 72-hour, or 48- to 72-hour time points are available in Table E2.

### Mortality Discrimination by Definition

The mortalities by grade using the standard ISHLT, revised, and Berlin-type modifications based on P/F ratio thresholds were compared. Under the ISHLT definition, mild PGD (P/F > 300 with abnormal chest radiograph) had no significant 30-day mortality difference from no PGD (P/F > 300 with normal chest radiograph,  $P = 0.51$ ; Figure 2a) using the most common interval grading (0–72 h) or at any other individual time point or other interval ( $P$  values ranged between 0.14 at time (T) 0 to 0.91 at T72; see Tables E3–E7). Conditional logistic regression demonstrated nearly identical point estimates and significance, indicating minimal center effects (Table E8). Kaplan-Meier analysis displayed no differences between no and mild PGD with overlapping survival curves (log-rank  $P = 0.70$ ; Figure 2B).

At almost all discrete time points or combined intervals, the revised definition with combined grades of P/F greater than 300 demonstrated better stratified mortality risk based on PGD severity (Figure 2C). Longer-term mortality risk (90 d and 1 yr) also demonstrated good risk stratification using this definition (log-rank  $P < 0.0001$ ; Figure 2D). The Berlin-type definition significantly improved stratified mortality risk based on addition of a very severe category (Figure 2E). However, there was overlap without differential longer-term mortality risk between the severe and very severe groups (log-rank  $P = 0.93$ ; Figure 2F), indicating that any benefit from addition of a very severe category was limited to early mortality.

Comparisons of the ISHLT, revised, and Berlin-type definitions also demonstrated significant improvement in discriminant validity using later discrete (48 or 72 h; Figures E4–E6) or interval (48–72 h; Figures E4–E6) time points (Table 2). Despite superior performance with later grading, early risk stratification (0 h) was possible among all definitions. However, earlier grading time points attenuated discriminant validity for ISHLT (area under the curve [AUC] = 0.65), revised (AUC = 0.68), and Berlin-type (AUC = 0.68) definitions.

A multivariable Cox proportional hazards model was fit using potentially confounding variables previously associated with PGD and/or mortality (24, 35). Common to all definitions was an approximately twofold increase in risk of mortality for severe PGD grades after adjustment. Of note, the Berlin-type very

severe stratum did not appreciably differ from the severe category, with a similar twofold increased risk of mortality (Table E9), indicating little effect of addition of this severity group on long-term mortality.

### Divergent Discrimination of Known Lung Injury Biomarkers

Based on prior publications (27–29), we chose a known lung injury biomarker with the strongest association with PGD to test divergent discrimination with the different PGD definitions. PAI-1 plasma levels for each grade, by definition, from 315 recipients are depicted in Figure 3. There was no significant difference in PAI-1 plasma level between no PGD and mild PGD ( $P = 0.99$ ), concordant with our demonstrated lack of mortality risk difference between these grades. Plasma PAI-1 levels demonstrated significant incremental elevation with moderate (ISHLT,  $P = 0.04$ ; revised,  $P = 0.03$ ; Berlin-type,  $P = 0.055$ ) and severe grades of PGD (ISHLT,  $P = 0.001$ ; revised,  $P < 0.001$ ; Berlin-type,  $P < 0.001$ ), but not for very severe grades ( $P = 0.14$ ). Sensitivity analyses using nonparametric methods also confirmed these observations.

### Effects of Ventilator Status and Transplant Procedure Type on Definitions

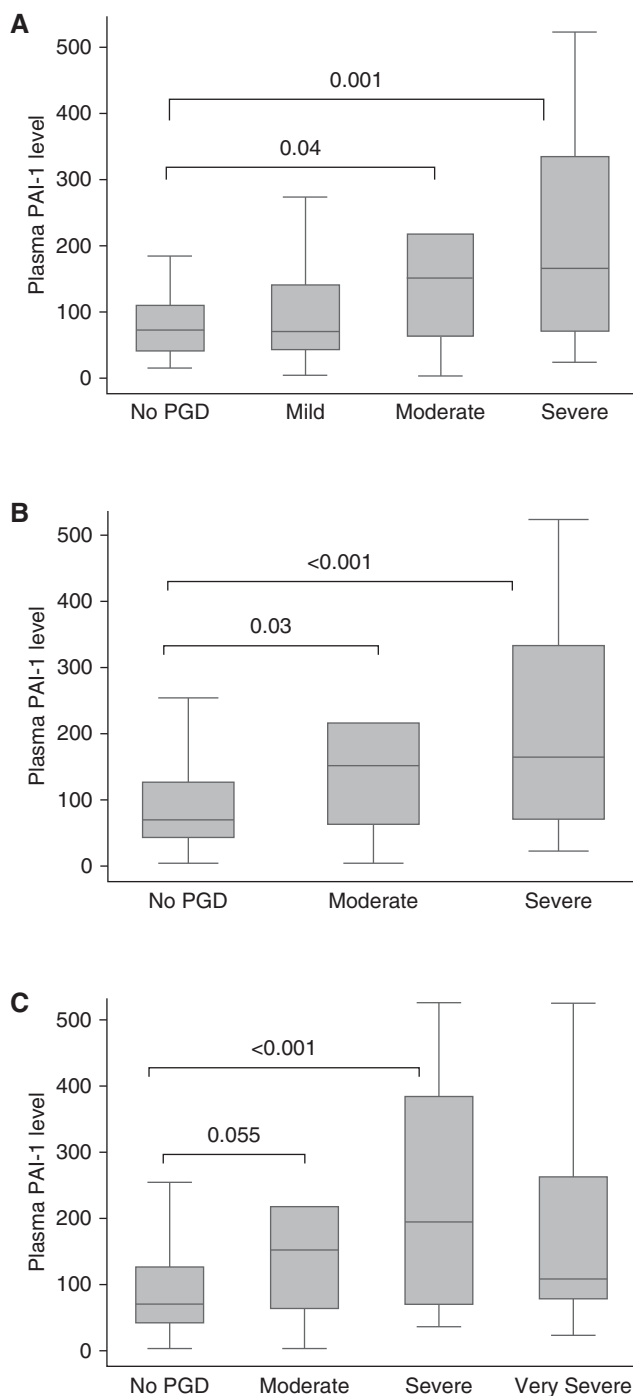
To evaluate the effects of ventilator status on definitions, all non-mechanically ventilated patients were recoded to have no PGD (grade 0), as has been previously suggested (18, 20). Noninvasive ventilation was only used in four patients during the first 3 days

**Table 2.** Construct Discriminant Validity across Grading Assumptions

	ISHLT		Revised		Berlin-Type	
	AUC	P Value	AUC	P Value	AUC	P Value
Timing						
0–72 h	0.685	—	0.679	—	0.708	—
48–72 h	0.755	0.002	0.744	0.008	0.749	0.103
Ventilator status						
Ventilation	0.755	—	0.744	—	0.749	—
No ventilation	0.747	0.770	0.743	0.941	0.751	0.727
Transplant type						
Single-lung Tx	0.841	—	0.814	—	0.811	—
Bilateral-lung Tx	0.717	0.118	0.713	0.261	0.720	0.319

*Definition of abbreviations:* AUC = area under the curve; ISHLT = International Society for Heart and Lung Transplantation; Tx = transplant.

Comparisons of significance within definitions are listed directly to the right of the AUC. There were no significant differences in discriminant ability between definitions for each assumption. Note: single and bilateral cohorts are abstracted from the 48- to 72-hour interval definitions, and the significance is determined using a chi-square test of the receiver operating characteristic curves comparing single to bilateral.



**Figure 3.** Plasminogen activator inhibitor (PAI)-1 biomarker association with primary graft dysfunction (PGD) definitions. Displayed are comparisons of PGD grade assessed between 48 and 72 hours for (A) International Society for Heart and Lung Transplantation, (B) revised, and (C) Berlin-type definitions. In this box-and-whisker plot, the box defines the upper/lower quartiles, the line depicts the median, and the whiskers the maximum and minimum values. Significance was determined using ANOVA.

in this cohort, and was thus treated similarly to mechanical forms of ventilation. The resulting model showed no significant loss of discrimination (Table 2). Among all definitions with varying P/F

thresholds, results were similar regardless of presence or absence of mechanical ventilation. In addition, we were unable to demonstrate any interaction of ventilator status and 30-day mortality. Further

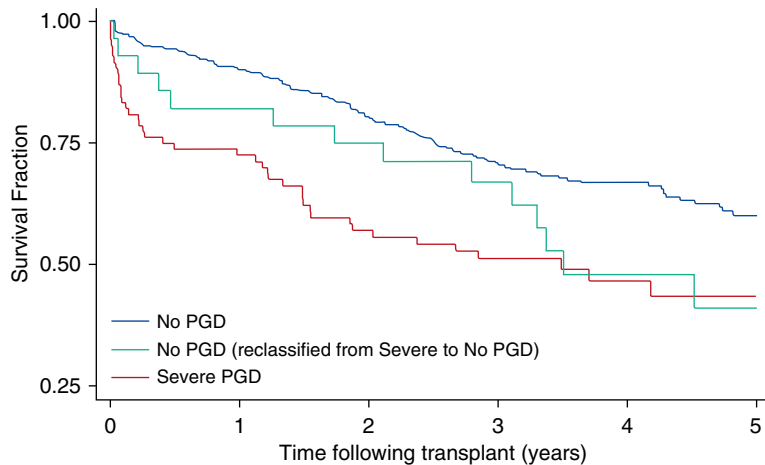
sensitivity analysis using less extreme variations of definition determined by infiltrates on chest radiograph (no infiltrates, no PGD: grade 0; infiltrates, mild PGD: grade 1) similarly had little effects on mortality discrimination.

Because the decision to continue or discontinue mechanical ventilation can potentially affect PGD grading, particularly when used at 72 hours as an endpoint for unblinded clinical trials, further analysis of the severe PGD groups was conducted. Among those in the 72-hour severe PGD (grade 3) group who were reclassified as no PGD (grade 0) due to not receiving mechanical ventilation, the 30-day mortality risk appeared to be intermediate when compared with those receiving mechanical ventilation: 7% for nonmechanically ventilated reclassified versus 14% for mechanically ventilated severe PGD ( $P=0.32$ ) and versus 1% for no PGD ( $P=0.03$ ), as displayed in Figure 4.

With respect to transplant procedure type, incremental risk per grade was consistent across all definitions within both single and bilateral lung transplants. We were unable to demonstrate an interaction between transplant type with PGD grade for any definition ( $P=0.09$ – $0.88$ ). Likewise, discrimination was not significantly different between single or bilateral transplants, despite a nominally higher AUC for mortality discrimination within single-lung transplants (Table 2).

## Discussion

The ISHLT consensus definition of 2005 has been instrumental in enhancing the transplant community's ability to study and understand the incidence, related outcomes, risk factors, biomarkers, and genomic factors influencing PGD. However, clinical practice has changed, and clinical trials are now being conducted with increasing frequency using PGD as the outcome for approval by governing bodies (22, 24). Thus, a valid PGD grading method is essential in objectively assessing the risks and benefits of a therapy. In this study, we have shown that definitions combining categories with P/F greater than 300 improve convergent and discriminant validity in the context of associated biomarkers and mortality compared with the existing definition. Furthermore, the addition of more severe categories did not enhance the definition, except at the earliest



**Figure 4.** Kaplan-Meier survival of International Society for Heart and Lung Transplantation (ISHLT) definition after adjustment for mechanical ventilation status. Demonstrated are comparisons of the ISHLT definition for the 72-hour discrete time point. Significance was determined by the log-rank test. Curves are discriminated by color: blue, no primary graft dysfunction (PGD); green, severe PGD reclassified as no PGD owing to absence of mechanical ventilation status; and red, severe PGD.

grading times. We have also demonstrated that later time points for assessment of PGD grade (48–72 h) have better discrimination for mortality, and all grading constructs based on P/F ratios perform well, regardless of transplant type or mechanical ventilation status.

With respect to 30-day, 90-day and 1-year mortality, there were no significant differences in risk between no (grade 0) and mild (grade 1) PGD at any time point or interval. Differences between these categories are determined by findings on chest radiography; thus, for grades with P/F ratios over 300 mm Hg, there was no difference in mortality, regardless of chest radiograph results. Biomarker data likewise demonstrated no significant differences in plasma levels of PAI-1 between no PGD and mild PGD. Furthermore, addition of a very severe PGD category did not reduce discriminant ability of 30-day mortality, did not enhance biomarker discriminant validity, and demonstrated similar overall mortality risk as severe PGD.

There was improved mortality and biomarker discrimination when grading was employed at later time points. It is reasonable to recommend that PGD prevention and risk studies focus on 48- to 72-hour grading time points, given the demonstrated superior discriminant validity. However, in trials attempting early identification of high-risk individuals for interventions (e.g., for enrollment based on PGD grade within 24 h), it is reasonable to use the first discrete blood gas on arrival to the intensive care unit (T0) severe PGD

grade to identify high-risk recipients for early postoperative interventions. This is justifiable because all definitions were capable of defining high-risk candidates for early intervention using the T0 grade, albeit with some sacrifice in discrimination.

The effect of mechanical ventilation on validity of PGD definition was minimal. In addition, there was no interaction between ventilator status and PGD grade on mortality discrimination. However, subjects with severe PGD based on P/F and chest radiograph at 72 hours who were reclassified as a lower PGD grade due to lack of mechanical ventilation use demonstrated mortality that was significantly higher than “traditional” no PGD recipients. Therefore, presence of mechanical ventilation should be explicitly reported, and should not result in an empirical change in PGD grading. Furthermore, it would seem prudent that regulatory agencies recommend reporting short-term mortality along with PGD grade.

We also investigated the effects of transplant type on definition performance, given the concern for potential native lung contribution to P/F ratios in PGD grading. Transplant procedure type had no interaction with PGD grade on mortality discrimination. Furthermore, adequate discriminant validity was demonstrated for PGD grading within transplant type. There is evidence, therefore, to support reporting PGD grades in aggregate based on transplant type; however, adjuncts reporting within individual transplant type should be considered, or care should be taken to

ensure balanced distribution of transplant type between comparator groups.

There are several potential limitations of this study. There is the potential for misclassification of PGD due to missing or incorrect data. Because data were collected prospectively, audited regularly, and are missing in less than 0.5% of the variables of interest, misclassification is less likely. The use of extracorporeal membrane oxygenation, inhaled vasodilator medicines, and noninvasive ventilation in this cohort was limited to very small numbers of patients (<20), preventing any substantive evaluation of effects on grading and mortality. Given that practice patterns have changed during the period of our study, future studies of the effects of extracorporeal membrane oxygenation, pulmonary vasodilators, and noninvasive ventilation strategies are warranted, particularly as these strategies have become more frequently used in recent years. The LTOG dataset did not include endpoints assessing chronic lung allograft dysfunction, which has been shown to be associated with PGD in prior studies (38–40). Given the relationship of lower PGD grades with chronic lung allograft dysfunction in these prior studies, future investigation is warranted. Though definition performance with a lung injury biomarker previously shown to be associated with binary PGD definitions was excellent, other biomarkers may not also demonstrate a dose–response effect based on degree of lung injury. Additional work will be needed to evaluate other identified biomarkers in separate populations. The generalizability of these results to non-LTOG sites has not been proven; however, this cohort is specifically designed to study PGD, and includes transplant recipients from 10 U.S. centers of varying size and location; therefore, it provides the largest sample size for study of PGD currently available.

In summary, we have demonstrated construct validity of the ISHLT definition, regardless of presence of mechanical ventilation and procedure type. Combining grading categories with P/F greater than 300 is justified in most circumstances, and grading at later time points provides the most impactful clinical definition for PGD risk or prevention studies. Further addition of PGD severity categories below P/F ratios of 200 had minimal to no effects on mortality or biomarker discrimination, except when graded at the earliest time points. ■

**Author disclosures** are available with the text of this article at [www.atsjournals.org](http://www.atsjournals.org).

## References

- Christie JD, Kotloff RM, Ahya VN, Tino G, Pochettino A, Gaughan C, et al. The effect of primary graft dysfunction on survival after lung transplantation. *Am J Respir Crit Care Med* 2005;171:1312–1316.
- Christie JD, Kotloff RM, Pochettino A, Arcasoy SM, Rosengard BR, Landis JR, et al. Clinical risk factors for primary graft failure following lung transplantation. *Chest* 2003;124:1232–1241.
- Christie JD, Sager JS, Kimmel SE, Ahya VN, Gaughan C, Blumenthal NP, et al. Impact of primary graft failure on outcomes following lung transplantation. *Chest* 2005;127:161–165.
- King RC, Binns OA, Rodriguez F, Kanithanon RC, Daniel TM, Spotnitz WD, et al. Reperfusion injury significantly impacts clinical outcome after pulmonary transplantation. *Ann Thorac Surg* 2000;69:1681–1685.
- Yusen RD, Edwards LB, Kucheryavaya AY, Benden C, Dipchand AI, Dobbels F, et al. International Society for Heart and Lung Transplantation. The registry of the International Society for Heart and Lung Transplantation: thirty-first adult lung and heart–lung transplant report—2014; focus theme: retransplantation. *J Heart Lung Transplant* 2014;33:1009–1024.
- Lee JC, Christie JD, Keshavjee S. Primary graft dysfunction: definition, risk factors, short- and long-term outcomes. *Semin Respir Crit Care Med* 2010;31:161–171.
- Whitson BA, Prekker ME, Herrington CS, Whelan TP, Radosevich DM, Hertz MI, et al. Primary graft dysfunction and long-term pulmonary function after lung transplantation. *J Heart Lung Transplant* 2007;26:1004–1011.
- Christie JD, Carby M, Bag R, Corris P, Hertz M, Weill D; ISHLT Working Group on Primary Lung Graft Dysfunction. Report of the ISHLT Working Group on Primary Lung Graft Dysfunction part II: definition: a consensus statement of the International Society for Heart and Lung Transplantation. *J Heart Lung Transplant* 2005;24:1454–1459.
- Arcasoy SM, Fisher A, Hachem RR, Scavuzzo M, Ware LB; ISHLT Working Group on Primary Lung Graft Dysfunction. Report of the ISHLT Working Group on Primary Lung Graft Dysfunction part V: predictors and outcomes. *J Heart Lung Transplant* 2005;24:1483–1488.
- Barr ML, Kawut SM, Whelan TP, Girgis R, Böttcher H, Sonett J, et al.; ISHLT Working Group on Primary Lung Graft Dysfunction. Report of the ISHLT Working Group on Primary Lung Graft Dysfunction part IV: recipient-related risk factors and markers. *J Heart Lung Transplant* 2005;24:1468–1482.
- de Perrot M, Bonser RS, Dark J, Kelly RF, McGiffin D, Menza R, et al.; ISHLT Working Group on Primary Lung Graft Dysfunction. Report of the ISHLT Working Group on Primary Lung Graft Dysfunction part III: donor-related risk factors and markers. *J Heart Lung Transplant* 2005;24:1460–1467.
- Christie JD, Van Raemdonck D, de Perrot M, Barr M, Keshavjee S, Arcasoy S, et al.; ISHLT Working Group on Primary Lung Graft Dysfunction. Report of the ISHLT Working Group on Primary Lung Graft Dysfunction part I: introduction and methods. *J Heart Lung Transplant* 2005;24:1451–1453.
- Christie JD, Bellamy S, Ware LB, Lederer D, Hadjilias D, Lee J, et al. Construct validity of the definition of primary graft dysfunction after lung transplantation. *J Heart Lung Transplant* 2010;29:1231–1239.
- Bernard GR, Artigas A, Brigham KL, Carlet J, Falke K, Hudson L, et al. The American–European Consensus Conference on ARDS: definitions, mechanisms, relevant outcomes, and clinical trial coordination. *Am J Respir Crit Care Med* 1994;149:818–824.
- Ranieri VM, Rubenfeld GD, Thompson BT, Ferguson ND, Caldwell E, Fan E, et al.; ARDS Definition Task Force. Acute respiratory distress syndrome: the Berlin definition. *JAMA* 2012;307:2526–2533.
- Christie J, Keshavjee S, Orens J, Arcasoy S, DePerrot M, Barr M, et al.; ISHLT Working Group on PGD. Potential refinements of the International Society for Heart and Lung Transplantation primary graft dysfunction grading system. *J Heart Lung Transplant* 2008;27:138.
- Oto T, Griffiths AP, Levvey BJ, Pilcher DV, Williams TJ, Snell GI. Definitions of primary graft dysfunction after lung transplantation: differences between bilateral and single lung transplantation. *J Thorac Cardiovasc Surg* 2006;132:140–147.
- Oto T, Levvey BJ, Snell GI. Potential refinements of the International Society for Heart and Lung Transplantation primary graft dysfunction grading system. *J Heart Lung Transplant* 2007;26:431–436.
- Prekker ME, Herrington CS, Hertz MI, Radosevich DM, Dahlberg PS. Early trends in PaO<sub>2</sub>/fraction of inspired oxygen ratio predict outcome in lung transplant recipients with severe primary graft dysfunction. *Chest* 2007;132:991–997.
- Prekker ME, Nath DS, Walker AR, Johnson AC, Hertz MI, Herrington CS, et al. Validation of the proposed International Society for Heart and Lung Transplantation grading system for primary graft dysfunction after lung transplantation. *J Heart Lung Transplant* 2006;25:371–378.
- Suzuki Y, Cantu E, Christie JD. Primary graft dysfunction. *Semin Respir Crit Care Med* 2013;34:305–319.
- Chertoff J. High-flow oxygen, PEEP, and the Berlin definition of ARDS: are they mutually exclusive? *Am J Respir Crit Care Med* 2017;196:396–397.
- Cypel M, Yeung JC, Liu M, Anraku M, Chen F, Karolak W, et al. Normothermic ex vivo lung perfusion in clinical lung transplantation. *N Engl J Med* 2011;364:1431–1440.
- Diamond JM, Lee JC, Kawut SM, Shah RJ, Localio AR, Bellamy SL, et al.; Lung Transplant Outcomes Group. Clinical risk factors for primary graft dysfunction after lung transplantation. *Am J Respir Crit Care Med* 2013;187:527–534.
- Ardehali A, Warnecke G, Van Raemdonck D, Loo G, Kukreja J, Smith M, et al. Impact of OCS lung portable EVLP on pulmonary function and bronchiolitis obliterans in standard lung transplant recipients - prospective evidence from the OCS lung INSPIRE Trial patients [abstract]. *J Heart Lung Transplant* 2016;35:S69.
- Cantu E, Diamond J, Nellen J, Beduhn B, Suzuki Y, Borders C, et al. Redefining primary graft dysfunction after lung transplantation [abstract]. *J Heart Lung Transplant* 2016;35:S89.
- Shah RJ, Bellamy SL, Localio AR, Wickersham N, Diamond JM, Weinacker A, et al. A panel of lung injury biomarkers enhances the definition of primary graft dysfunction (PGD) after lung transplantation. *J Heart Lung Transplant* 2012;31:942–949.
- Christie JD, Robinson N, Ware LB, Plotnick M, De Andrade J, Lama V, et al. Association of protein C and type 1 plasminogen activator inhibitor with primary graft dysfunction. *Am J Respir Crit Care Med* 2007;175:69–74.
- Christie JD, Shah CV, Kawut SM, Mangalmurti N, Lederer DJ, Sonett JR, et al.; Lung Transplant Outcomes Group. Plasma levels of receptor for advanced glycation end products, blood transfusion, and risk of primary graft dysfunction. *Am J Respir Crit Care Med* 2009;180:1010–1015.
- Covarrubias M, Ware LB, Kawut SM, De Andrade J, Milstone A, Weinacker A, et al.; Lung Transplant Outcomes Group. Plasma intercellular adhesion molecule-1 and von Willebrand factor in primary graft dysfunction after lung transplantation. *Am J Transplant* 2007;7:2573–2578.
- Pelaez A, Force SD, Gal AA, Neujahr DC, Ramirez AM, Naik PM, et al. Receptor for advanced glycation end products in donor lungs is associated with primary graft dysfunction after lung transplantation. *Am J Transplant* 2010;10:900–907.
- Sims MW, Beers MF, Ahya VN, Kawut SM, Sims KD, Lederer DJ, et al.; Lung Transplant Outcomes Group. Effect of single vs bilateral lung transplantation on plasma surfactant protein D levels in idiopathic pulmonary fibrosis. *Chest* 2011;140:489–496.
- Diamond JM, Akimova T, Kazi A, Shah RJ, Cantu E, Feng R, et al.; Lung Transplant Outcomes Group. Genetic variation in the prostaglandin E2 pathway is associated with primary graft dysfunction. *Am J Respir Crit Care Med* 2014;189:567–575.
- DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837–845.



35. Shah RJ, Diamond JM, Cantu E, Flesch J, Lee JC, Lederer DJ, *et al.* Objective estimates improve risk stratification for primary graft dysfunction after lung transplantation. *Am J Transplant* 2015;15: 2188–2196.
36. Andersen KH, Schultz HH, Nyholm B, Iversen MP, Gustafsson F, Carlsen J. Pulmonary hypertension as a risk factor of mortality after lung transplantation. *Clin Transplant* 2016;30: 357–364.
37. Cleves MA. Comparative assessment of three common algorithms for estimating the variance of the area under the nonparametric receiver operating characteristic curve. *Stata J* 2002;2:280–289.
38. Bharat A, Kuo E, Steward N, Aloush A, Hachem R, Trulock EP, *et al.* Immunological link between primary graft dysfunction and chronic lung allograft rejection. *Ann Thorac Surg* 2008;86:189–195. (Discussion, pp. 196–197.).
39. Daud SA, Yusef RD, Meyers BF, Chakinala MM, Walter MJ, Aloush AA, *et al.* Impact of immediate primary lung allograft dysfunction on bronchiolitis obliterans syndrome. *Am J Respir Crit Care Med* 2007; 175:507–513.
40. Huang HJ, Yusef RD, Meyers BF, Walter MJ, Mohanakumar T, Patterson GA, *et al.* Late primary graft dysfunction after lung transplantation and bronchiolitis obliterans syndrome. *Am J Transplant* 2008;8:2454–2462.