ORIGINAL ARTICLE

Effect of Esophageal Pressure–guided Positive End-Expiratory Pressure on Survival from Acute Respiratory Distress Syndrome
A Risk-based and Mechanistic Reanalysis of the EPVent-2 Trial

Todd Sarge[1], Elias Baedorf-Kassis[2], Valerie Banner-Goodspeed[1], Victor Novack[3], Stephen H. Loring[1], Michelle N. Gong[4], Deborah Cook[5], Daniel Talmor[1], and Jeremy R. Beitler[6]; for the EPVent-2 Study Group

[1]Department of Anesthesia, Critical Care, and Pain Medicine and [2]Division of Pulmonary and Critical Care Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts; [3]Soroka Clinical Research Center, Soroka University Medical Center, Beer-Sheva, Israel; [4]Montefiore Medical Center, Albert Einstein College of Medicine, Bronx, New York; [5]St. Joseph's Hospital, McMaster University, Hamilton, Ontario, Canada; and [6]Center for Acute Respiratory Failure, Division of Pulmonary, Allergy, and Critical Care Medicine, NewYork-Presbyterian Hospital, College of Physicians and Surgeons, Columbia University, New York, New York

ORCID ID: 0000-0003-0797-2374 (J.R.B.).

Abstract

Rationale: In acute respiratory distress syndrome (ARDS), the effect of positive end-expiratory pressure (PEEP) may depend on the extent to which multiorgan dysfunction contributes to risk of death, and the precision with which PEEP is titrated to attenuate atelectrauma without exacerbating overdistension.

Objectives: To evaluate whether multiorgan dysfunction and lung mechanics modified treatment effect in the EPVent-2 (Esophageal Pressure-guided Ventilation 2) trial, a multicenter trial of esophageal pressure (Pes)-guided PEEP versus empirical high PEEP in moderate to severe ARDS.

Methods: This *post hoc* reanalysis of the EPVent-2 trial evaluated for heterogeneity of treatment effect on mortality by baseline multiorgan dysfunction, determined via Acute Physiology and Chronic Health Evaluation II (APACHE-II). It also evaluated whether PEEP titrated to end-expiratory transpulmonary pressure near 0 cm H_2O was associated with survival.

Measurements and Main Results: All 200 trial participants were included. Treatment effect on 60-day mortality differed by

multiorgan dysfunction severity ($P = 0.03$ for interaction). Pes-guided PEEP was associated with lower mortality among patients with APACHE-II less than the median value (hazard ratio, 0.43; 95% confidence interval, 0.20–0.92) and may have had the opposite effect in patients with higher APACHE-II (hazard ratio, 1.69; 95% confidence interval, 0.93–3.05). Independent of treatment group or multiorgan dysfunction severity, mortality was lowest when PEEP titration achieved end-expiratory transpulmonary pressure near 0 cm H_2O.

Conclusions: The effect on survival of Pes-guided PEEP, compared with empirical high PEEP, differed by multiorgan dysfunction severity. Independent of multiorgan dysfunction, PEEP titrated to end-expiratory transpulmonary pressure closer to 0 cm H_2O was associated with greater survival than more positive *or* negative values. These findings warrant prospective testing in a future trial.

Keywords: acute respiratory distress syndrome; ventilator-induced lung injury; mechanical ventilation; positive end-expiratory pressure; randomized controlled trial

At a Glance Commentary

Scientific Knowledge on the Subject: Previous multicenter randomized trials evaluating positive end-expiratory pressure (PEEP) strategy in acute respiratory distress syndrome (ARDS) have either failed to demonstrate a significant treatment effect on mortality or have demonstrated increased mortality with an experimental (less conventional) PEEP strategy. Measuring esophageal pressure (Pes), an estimate of pleural pressure, may permit more precise PEEP titration that minimizes both atelectrauma and overdistension. However, the EPVent-2 (Esophageal Pressure-guided Ventilation 2) trial, the only multicenter randomized trial of Pes-guided PEEP, did not demonstrate a significant effect on mortality compared with an empirical high-PEEP strategy.

What This Study Adds to the Field: This *post hoc* reanalysis of EPVent-2 found the effect of PEEP strategy on mortality depended on pre-intervention severity of multiorgan dysfunction. Pes-guided PEEP was associated with lower mortality among patients with less severe multiorgan dysfunction and may have had the opposite effect in patients with more severe multiorgan dysfunction. Independent of baseline illness severity, mortality was lowest when PEEP titration achieved an end-expiratory transpulmonary pressure (P_L = airway − pleural pressure) near 0 cm H_2O, consistent with mechanistic understanding of biophysical lung injury pathogenesis. Higher end-inspiratory P_L, a marker of tidal overdistension, was independently correlated with risk of circulatory shock. A future clinical trial, designed to account for baseline heterogeneity of multiorgan dysfunction, should evaluate PEEP titration to an end-expiratory P_L near 0 cm H_2O with inspiratory support individually tailored to minimize overdistension.

Titration of positive end-expiratory pressure (PEEP) in acute respiratory distress syndrome (ARDS) ideally should balance potential benefits of lung recruitment (preventing atelectrauma, attenuating heterogeneous regional strain, and facilitating gas exchange) against risks of overdistension (biophysical trauma and hemodynamic instability) (1). Esophageal manometry, when used to estimate pleural pressure, may facilitate more precise PEEP titration by enabling measures of global lung stress (transpulmonary pressure, P_L = airway − pleural pressure), identifying the propensity for tidal recruitment and derecruitment and distinguishing lung from chest wall mechanics (2–5). How best to incorporate these measurements to guide PEEP titration in ARDS is not well established.

The EPVent-2 (Esophageal Pressure-guided Ventilation 2) trial, the largest multicenter randomized trial to date utilizing esophageal manometry, compared esophageal pressure (Pes)-guided PEEP versus empirical high PEEP in patients with moderate to severe ARDS. In the Pes-guided PEEP group, PEEP was adjusted to achieve end-expiratory P_L between 0 and 6 cm H_2O according to a P_L–$F_{I_{O_2}}$ table (*see* the online supplement) (4, 6). In the comparator group, PEEP was adjusted according to an empirical high PEEP– $F_{I_{O_2}}$ table (online supplement), and Pes and P_L were measured (though not acted on) to permit mechanistic comparison between groups. The overall trial result found no significant difference in mortality or key morbidity endpoints (7).

This study reanalyzes EPVent-2 to evaluate the impact of *1*) multiorgan dysfunction and *2*) lung mechanics on treatment effect. First, we hypothesized that the effect of Pes-guided PEEP may depend on baseline attributable risk of death, reasoning that patients with severe multiorgan dysfunction and comorbidities at baseline may be less likely to benefit if their outcomes are determined largely by other factors (8–10). Second, we hypothesized that, independent of treatment group, PEEP titrated to end-expiratory P_L near 0 cm H_2O would be associated with greater survival, reasoning that this value balances the competing risks of atelectrauma (from negative P_L) and overdistension (from overly positive P_L) (7, 11, 12). This reanalysis was not prespecified in the trial's statistical analysis plan, and findings should be considered hypothesis-generating.

Some of the results of this study have been presented previously at Critical Care Canada Forum and to the Pleural Pressure Working Group of the European Society of Intensive Care Medicine.

Methods

Study Participants
All patients enrolled in EPVent-2 (NCT01681225) were included (7). Patients were eligible for the trial if aged 16 years or older and undergoing invasive mechanical ventilation for early moderate or severe ARDS (Pa_{O_2}:$F_{I_{O_2}} \le 200$ mm Hg).

EPVent-2 Trial Procedures
Trial participants were randomly assigned to one of two PEEP titration strategies (online supplement). In the Pes-guided PEEP group, Pes was measured at least once daily, and PEEP was adjusted to achieve end-expiratory P_L between 0 and 6 cm H_2O according to a P_L–$F_{I_{O_2}}$ table until gas exchange improved sufficiently to initiate PEEP and $F_{I_{O_2}}$ weaning (4, 6). End-inspiratory P_L was not to exceed 20 cm H_2O.

In the empirical high-PEEP group, PEEP was set according to a PEEP–$F_{I_{O_2}}$ table, and plateau airway pressure was not to exceed 35 cm H_2O. Although not acted on in the empirical high-PEEP group, Pes and P_L were measured for the first 7 days to permit mechanistic comparison between groups.

In both groups, V_T of 6 (range, 4–8) ml/kg predicted body weight and ventilator weaning were determined by protocol. Protocol-directed ventilator titration continued through Day 28 or until successful extubation, discharge, or study withdrawal.

Multiorgan Dysfunction Assessment
The Acute Physiology and Chronic Health Evaluation II (APACHE-II) score was used to quantify baseline risk of death from multiorgan dysfunction and chronic morbidities (13). Predicted 60-day survival was calculated from APACHE-II via logistic regression. Analyses were repeated using a modified APACHE-II computed without the oxygenation subscore to reflect only extrapulmonary illness. Other baseline characteristics were compared according to APACHE-II, dichotomized by the median value (27.5) to facilitate data visualization.

Heterogeneity of Treatment Effect by Multiorgan Dysfunction
The primary analysis tested whether treatment effect on 60-day mortality differed

by baseline APACHE-II. Cox proportional hazards models were used entering APACHE-II (as a continuous variable), treatment group, and their interaction. Kaplan-Meier plots for mortality treatment effect by dichotomized APACHE-II were compared using the log-rank test.

Several sensitivity analyses were performed to confirm heterogeneity findings were robust to varying model specifications (online supplement), Logistic regression models for 28-day, 60-day, and 1-year mortality were built to confirm findings did not depend on regression technique. Models were rerun replacing APACHE-II with the Sequential Organ Failure Assessment (SOFA) score to determine if findings were robust to different measures of multiorgan dysfunction. In a *post hoc* analysis, to determine if shock severity alone might explain observed findings, additional Cox models tested for an interaction between treatment group and baseline SOFA-cardiovascular (SOFA-CV) subscore. SOFA-CV is an ordinal score ranging from 0, signifying no hypotension or vasopressor requirement, to 4, signifying shock requiring high-dose vasopressors (14). Given baseline differences in body mass index (BMI) by APACHE-II score, an additional Cox model was developed *post hoc* entering BMI, treatment group, and their interaction to test for heterogeneity of treatment effect according to BMI.

In evaluation of secondary outcomes, Poisson regression models evaluated for effect modification with ventilator-free days and shock-free days (online supplement), entering treatment group, APACHE-II, and their interaction as covariates. *Post hoc* sensitivity analyses of secondary outcomes are described in the online supplement. Models also were rerun replacing APACHE-II with SOFA-CV to determine whether shock severity alone might explain observed findings.

Mechanistic Analysis by Transpulmonary Pressure

In the primary mechanistic analysis, Cox models for 60-day mortality were developed entering the average of the absolute value of end-expiratory P_L from baseline through Day 3 as the predictor of interest. This time interval for averaging end-expiratory P_L was chosen *a priori* to minimize bias from extubation, death, or empirical lowering of PEEP during weaning. The absolute value of end-expiratory P_L (distance from 0 cm H_2O) was used to represent quantitatively the mechanistic hypothesis that maintaining

end-expiratory P_L near 0 cm H_2O might mitigate both atelectrauma (from negative P_L) and overdistension (from overly positive P_L), thereby improving outcomes.

Several sensitivity analyses were performed to challenge the hypothesis that end-expiratory P_L near 0 cm H_2O was associated with favorable clinical outcomes. Average end-expiratory P_L through Day 3 was reentered in models dichotomized as within protective range (end-expiratory P_L within ± 2 cm H_2O) or nonprotective range (end-expiratory P_L of < -2 cm H_2O or > 2 cm H_2O), a range selected *a priori* to approximate the change in cardiac oscillation pressure often seen with esophageal manometry. To evaluate whether a simpler linear relationship would sufficiently explain the data, additional Cox models evaluated average end-expiratory P_L as an untransformed variable without taking the absolute value. Models were developed without adjustment for covariates, and separately with APACHE-II and randomly assigned treatment group entered as covariates. An interaction between the absolute value of end-expiratory P_L and APACHE-II was tested in a Cox model entering each term and their multiplicative interaction to determine if effect of P_L depended on APACHE-II. Kaplan-Meier plots for the relationship between average P_L (< -2 cm H_2O, within ± 2 cm H_2O, or > 2 cm H_2O) and survival through Day 60 were compared using the log-rank test for equality across strata.

Given the concern for overdistension injury, additional models were developed entering end-inspiratory P_L (peak global lung stress) as a covariate of interest. An unadjusted Cox model for 60-day mortality was constructed to determine if end-inspiratory P_L correlated with mortality. Then, to determine if the effects of end-expiratory P_L distance from 0 cm H_2O were explained in part by overdistension, additional models entered both end-expiratory P_L and end-inspiratory P_L, adjusting for APACHE-II.

The possibility that ventilatory strategy might contribute to hemodynamic instability was explored via *generalized* linear mixed models with logit link function and random intercept to evaluate the association of P_L with occurrence of vasopressor-dependent shock over time (online supplement). The first set of models entered end-expiratory P_L, day, and their interaction as covariates. Considering tidal

overdistension a potential alternative explanatory factor, a second set of models instead entered end-inspiratory P_L, day, and their interaction as independent variables. Finally, a third set of models entered as independent variables end-expiratory P_L, end-inspiratory P_L, and the interaction of each of them with time. In all models, the comparison of interest was whether P_L influenced odds of shock on successive study days (*i.e.*, the interaction of P_L with time).

Common Statistical Procedures

Patient characteristics were compared using Fisher's exact test, Wilcoxon rank-sum test, or *t* test as appropriate. Final Cox models met the proportional hazards assumption, assessed via Schoenfeld residuals and evaluation for time dependence of model covariates (online supplement). For all analyses, $P < 0.05$ was considered statistically significant without adjustment for multiple comparisons. For models evaluating differential response to treatment, a statistically significant interaction term was considered evidence of heterogeneity of treatment effect. SAS 9.4 was used for all analyses.

Results

Patient Characteristics

In total, 200 patients were enrolled in EPVent-2 and included in this analysis. P_L data were missing in one patient, in whom the esophageal balloon catheter could not be inserted successfully; that patient was included in all analyses except those requiring P_L data.

Baseline characteristics by APACHE-II are presented in Tables 1 and E1. Patients with lower APACHE-II had higher BMI, higher pH, and decreased need for vasopressors at baseline. ARDS risk factors and respiratory mechanics did not differ significantly by APACHE-II. Treatment assignment was evenly distributed as expected from randomization.

Baseline Illness Severity and Risk of Death

APACHE II was normally distributed within each treatment group and evenly distributed between groups (Pes-guided vs. empirical high PEEP: 27.0 ± 7.7 vs. 27.7 ± 7.4; $P = 0.35$). The corresponding baseline risk of death by Day 60 was median 36.6% (interquartile range, 29.0–43.0%) in the Pes-guided PEEP group

Table 1. Characteristics of Trial Participants by APACHE-II

Variable	Low APACHE-II* (n = 99)	High APACHE-II* (n = 101)	Difference (95% CI)†
Age	57 (42 to 67)	58 (47 to 67)	−2 (−7 to 3)
Sex, F, n (%)	52 (52.5)	40 (39.6)	12.9 (−0.8 to 26.6)
Weight, kg	85.4 (72.0 to 110.3)	77.8 (69.0 to 96.5)	7.6 (1.3 to 14.1)
Body mass index, kg/m²	31.1 (26.4 to 39.7)	28.3 (24.8 to 33.4)	2.7 (0.7 to 4.9)
Time intubated prior to enrollment, h	21.5 (15 to 30)	21 (13 to 33)	0 (−4 to 4)
APACHE-II at enrollment	21 (4)	33 (4)	−12 (−14 to −11)
Acute physiology subscore	18 (4)	30 (4)	−12 (−13 to −11)
Chronic health subscore	3 (2)	3 (2)	−1 (−1 to 0)
SOFA at enrollment	8.5 (6 to 10)	13 (12 to 15)	−5 (−6 to −4)
ARDS risk factor, n (%)			
Sepsis	87 (87.9)	84 (83.2)	4.7 (−5.0 to 14.4)
Pneumonia	78 (78.8)	71 (70.3)	8.5 (−3.5 to 20.5)
Aspiration	17 (17.2)	25 (24.8)	−7.6 (−18.8 to 3.7)
Multiple transfusions	5 (5.1)	14 (13.9)	−8.8 (−16.8 to −0.8)
Acute pancreatitis	1 (1.0)	6 (5.9)	−4.9 (−9.9 to 0.1)
Trauma	3 (3.0)	3 (3.0)	−0.1 (−5.7 to 4.8)
Any pulmonary risk factor	87 (87.9)	83 (82.2)	5.7 (−4.2 to 15.6)
Respiratory characteristics at enrollment			
pH	7.36 (7.31 to 7.40)	7.29 (7.23 to 7.35)	0.06 (0.04 to 0.08)
Pa_{CO_2}, mm Hg	45 (37 to 52)	43 (37 to 50)	1 (−2 to 4)
Pa_{O_2}:F_{IO_2}	95 (75 to 132)	86 (69 to 119)	7 (−4 to 16)
V_T, ml/kg PBW	6.2 (5.9 to 6.8)	6.1 (5.9 to 7.0)	0.1 (−0.1 to 0.3)
Respiratory rate, breaths/min	25 (22 to 29)	26 (23 to 30)	−2 (−3 to 0)
V_E, L/min	9.4 (8.4 to 11.2)	10.2 (8.5 to 12.2)	−0.5 (−1.3 to 0.2)
Ventilatory ratio‡	1.9 (1.6 to 2.3)	2.0 (1.6 to 2.3)	0.0 (−0.2 to 0.1)
Set PEEP, cm H_2O	14 (10 to 16)	14 (10 to 16)	0 (−2 to 1)
Plateau airway pressure, cm H_2O	28 (24 to 31)	27 (25 to 31)	0 (−1 to 2)
Driving airway pressure, cm H_2O	13 (10 to 15)	13 (10 to 14)	0 (−1 to 1)
Pes, cm H_2O			
At end-expiration	16 (12 to 19)	16 (13 to 18)	−1 (−2 to 1)
At end-inspiration	18 (15 to 21)	19 (16 to 21)	−1 (−2 to 0)
Transpulmonary pressure, cm H_2O			
At end-expiration	−1 (−4 to 2)	−1 (−3 to 1)	0 (−1 to 1)
At end-inspiration	8 (5 to 12)	8 (5 to 11)	0 (−1 to 2)
Cointerventions at enrollment, n (%)			
Vasopressors	45 (45.5)	69 (68.3)	−22.9 (−36.2 to −9.5)
Neuromuscular blockade	28 (28.3)	38 (37.6)	−9.3 (−22.3 to 3.6)
Systemic corticosteroids	34 (34.3)	35 (34.7)	−0.3 (−13.5 to 12.9)
Treatment assignment, n (%)			
Pes-guided PEEP	51 (51.5)	51 (50.5)	1.0 (−12.8 to 14.9)
Empirical PEEP	48 (48.5)	50 (49.5)	−1.0 (−14.9 to 12.8)
Outcomes			
28-d mortality, n (%)	21 (21.2)	42 (41.6)	−20.4 (−32.9 to −7.8)
60-d mortality, n (%)	29/98 (29.6)	46/101 (45.5)	−15.9 (−29.2 to −2.7)
Ventilator-free days through Day 28	20 (0 to 24)	9 (0 to 22)	2 (0 to 6)
Shock-free days through Day 28	18 (7 to 22)	9 (0 to 20)	4 (1 to 7)

Definition of abbreviations: APACHE-II = Acute Physiology and Chronic Health Evaluation II; ARDS = acute respiratory distress syndrome; CI = confidence interval; PBW = predicted body weight; PEEP = positive end-expiratory pressure; Pes = esophageal pressure; SOFA = sequential organ failure assessment.
Data are shown as median (interquartile range) unless otherwise indicated.
*Low and high APACHE-II refer to patients with values less or greater than the median APACHE-II score, which was 27.5.
†Hodges-Lehmann estimation was used to determine unbiased median differences and 95% CIs for variables reported as median (interquartile range). The result represents the median difference among all possible pairs formed by one patient from each group and thus does not necessarily equal the crude difference between group medians.
‡Ventilatory ratio is a surrogate for dead-space fraction calculated as [$\dot{V}_E$ (ml/min) × Pa_{CO_2} (mm Hg)]/(predicted body weight × 100 × 37.5).

and 37.6% (31.9–44.3%) in the empirical high-PEEP group (*P* = 0.34).

A modified APACHE-II computed without the oxygenation subscore did not appreciably change the distribution of overall illness severity or baseline risk of death (Figure E1). Modified APACHE-II and associated mortality risk remained evenly distributed between treatment groups.

Heterogeneity of Treatment Effect on Mortality
The effect of Pes-guided PEEP on 60-day mortality differed by baseline overall illness severity (*P* = 0.03 for interaction of treatment group with APACHE-II as a continuous variable). Pes-guided PEEP was associated with a 57% decrease in the hazard for death among patients with a low APACHE-II score (less than median value) and appeared to have the opposite
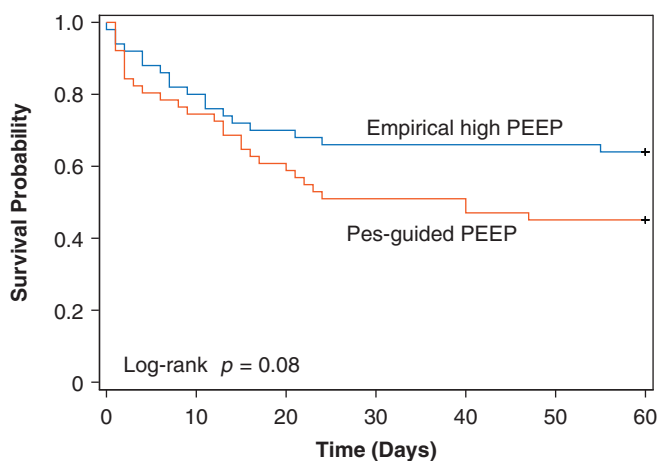
A Low APACHE-II
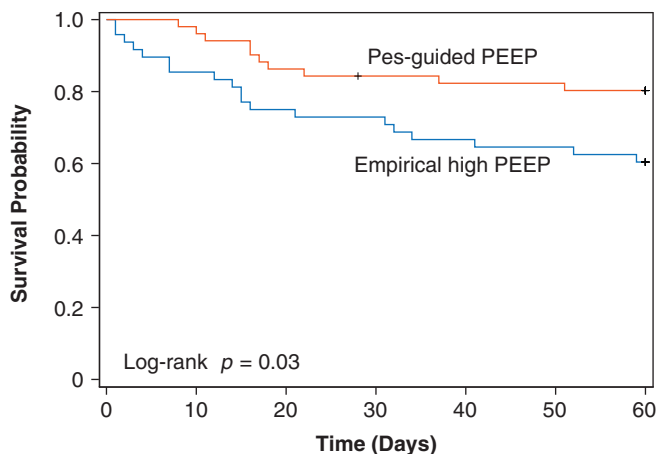
No. at Risk

Pes-guided PEEP	51	50	44	42	41	41	40
Empirical high PEEP	48	41	36	35	32	31	29

B High APACHE-II

No. at Risk

Pes-guided PEEP	51	38	31	26	26	23	23
Empirical high PEEP	50	40	35	33	33	33	32

Figure 1. Kaplan-Meier survival analysis by randomly assigned treatment group, stratified by APACHE-II score. Low and high APACHE-II scores refer to patients with values less or greater than the median APACHE-II score, respectively. Median APACHE-II score was 27.5. APACHE-II = Acute Physiology and Chronic Health Evaluation II; PEEP = positive end-expiratory pressure; Pes = esophageal pressure.

effect in patients with a high APACHE-II score (hazard ratio [HR] for a low APACHE-II score, 0.43, 95% confidence interval [CI], 0.20–0.92; HR for high APACHE-II score, 1.69; 95% CI 0.93–3.05; $P < 0.01$ for interaction with dichotomized APACHE-II) (Figure 1).

In sensitivity analyses, separate Cox proportional hazards models

demonstrated a significant interaction between treatment group and modified APACHE-II with the pulmonary subscore removed ($P = 0.04$), and when APACHE-II was replaced by SOFA in the model ($P = 0.04$) (Table E2). Logistic regression confirmed the significant treatment–APACHE-II interaction was independent of modeling technique

(Table E2) and consistent across 28-day, 60-day, and 1-year mortality (Figure 2). Across analyses, the interaction consistently indicated Pes-guided PEEP was associated with more favorable outcomes among patients with less multiorgan dysfunction, and less favorable outcomes among patients with more multiorgan dysfunction.

Although BMI was higher in patients with a low APACHE-II score (Table 1), there was no evidence for heterogeneity of treatment effect on 60-day mortality by BMI ($P = 0.74$ for interaction of BMI with treatment group). The interaction between treatment group and APACHE-II score remained significant in multivariable models adjusting for BMI (Table E2).

Heterogeneity of Treatment Effect on Secondary Endpoints

The effect of Pes-guided PEEP on both ventilator-free days and shock-free days was dependent on baseline overall illness severity ($P < 0.01$ for interaction of treatment group with APACHE-II score in Poisson regression models).

The directionality of treatment effect on both ventilator-free days and shock-free days mirrored that for mortality (Figure 2). Among patients with lower APACHE-II scores (less than median value), assignment to Pes-guided PEEP was associated with more ventilator-free days and shock-free days, whereas the opposite effect was seen among patients with higher APACHE-II scores (Figure 3).

However, APACHE-II-dependent treatment effects on ventilator-free days and shock-free days were not robust to *post hoc* sensitivity analyses. Neither the zero-inflated negative binomial model nor the stratified rank-sum van Elteren test provided evidence to suggest the effect of treatment group on either ventilator-free days or shock-free days was dependent on APACHE-II score.

Effects of Baseline Shock Severity on Mortality and Secondary Endpoints

The effect of Pes-guided PEEP on 60-day mortality did not depend on baseline hemodynamic instability, assessed via SOFA-CV score ($P = 0.12$ for interaction of SOFA-CV with treatment group in Cox model) (Table E3). However, baseline SOFA-CV score did modify the effect of Pes-guided PEEP on both shock-free days and ventilator-free days ($P < 0.01$ for
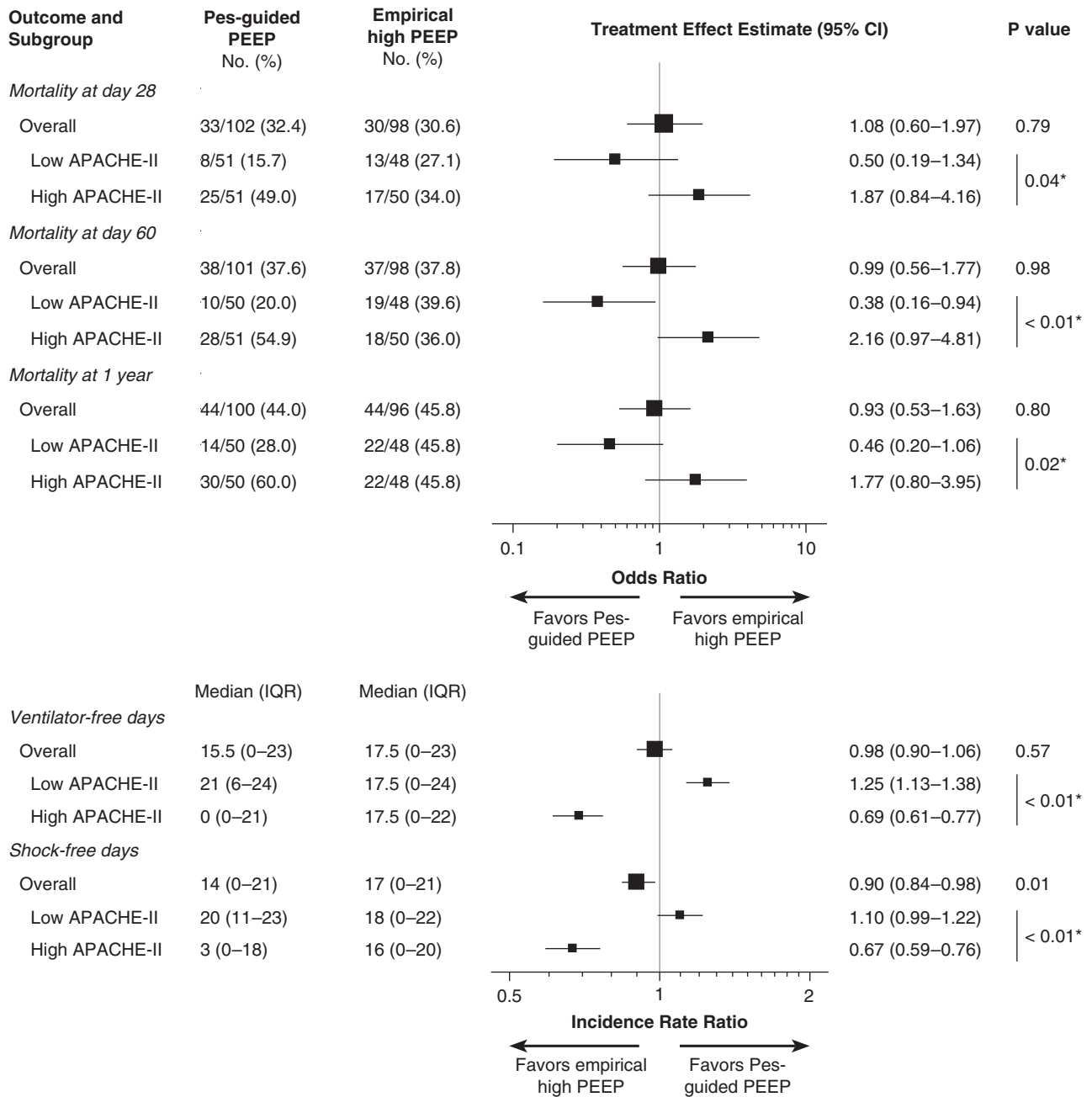
Outcome and Subgroup	Pes-guided PEEP No. (%)	Empirical high PEEP No. (%)	Treatment Effect Estimate (95% CI)	P value
Mortality at day 28				
Overall	33/102 (32.4)	30/98 (30.6)	1.08 (0.60–1.97)	0.79
Low APACHE-II	8/51 (15.7)	13/48 (27.1)	0.50 (0.19–1.34)	0.04*
High APACHE-II	25/51 (49.0)	17/50 (34.0)	1.87 (0.84–4.16)	
Mortality at day 60				
Overall	38/101 (37.6)	37/98 (37.8)	0.99 (0.56–1.77)	0.98
Low APACHE-II	10/50 (20.0)	19/48 (39.6)	0.38 (0.16–0.94)	< 0.01*
High APACHE-II	28/51 (54.9)	18/50 (36.0)	2.16 (0.97–4.81)	
Mortality at 1 year				
Overall	44/100 (44.0)	44/96 (45.8)	0.93 (0.53–1.63)	0.80
Low APACHE-II	14/50 (28.0)	22/48 (45.8)	0.46 (0.20–1.06)	0.02*
High APACHE-II	30/50 (60.0)	22/48 (45.8)	1.77 (0.80–3.95)	

Odds Ratio (scale 0.1 – 1 – 10)

← Favors Pes-guided PEEP   Favors empirical high PEEP →

	Median (IQR)	Median (IQR)		
Ventilator-free days				
Overall	15.5 (0–23)	17.5 (0–23)	0.98 (0.90–1.06)	0.57
Low APACHE-II	21 (6–24)	17.5 (0–24)	1.25 (1.13–1.38)	< 0.01*
High APACHE-II	0 (0–21)	17.5 (0–22)	0.69 (0.61–0.77)	
Shock-free days				
Overall	14 (0–21)	17 (0–21)	0.90 (0.84–0.98)	0.01
Low APACHE-II	20 (11–23)	18 (0–22)	1.10 (0.99–1.22)	< 0.01*
High APACHE-II	3 (0–18)	16 (0–20)	0.67 (0.59–0.76)	

Incidence Rate Ratio (scale 0.5 – 1 – 2)

← Favors empirical high PEEP   Favors Pes-guided PEEP →

Figure 2. Treatment effects by APACHE-II score. Odds ratio for mortality <1 indicates treatment effect favors Pes-guided PEEP. Incidence rate ratio >1 indicates treatment effect favors Pes-guided PEEP. Effect estimates are reported from the overall study population or unadjusted subgroup logistic or Poisson models as indicated. *P value for interaction term. APACHE-II = Acute Physiology and Chronic Health Evaluation II; IQR = interquartile range; PEEP = positive end-expiratory pressure; Pes = esophageal pressure.

interaction in Poisson models). In patients without shock or hypotension at baseline (SOFA-CV score of 0), Pes-guided PEEP was associated with more shock-free days and ventilator-free days. By contrast, in patients with vasopressor-dependent shock at baseline (SOFA-CV score of 3 or 4), Pes-guided PEEP was associated with fewer shock-free days and ventilator-free days (Figure E2).

Distribution of End-Expiratory P$_L$
End-expiratory P$_L$ values were reported in 191 of 200 patients at baseline, and 199 of 200 patients on protocol. At baseline, 57.1% of patients had a negative P$_L$ value, and fewer than half of patients (45.5%) had end-expiratory P$_L$ within the hypothesized protective range of ±2 cm H$_2$O (Figure 4). End-expiratory P$_L$ did not differ significantly between treatment groups at baseline.

Upon initiating study-directed PEEP, end-expiratory P$_L$ was maintained nearer to 0 cm H$_2$O through the majority of
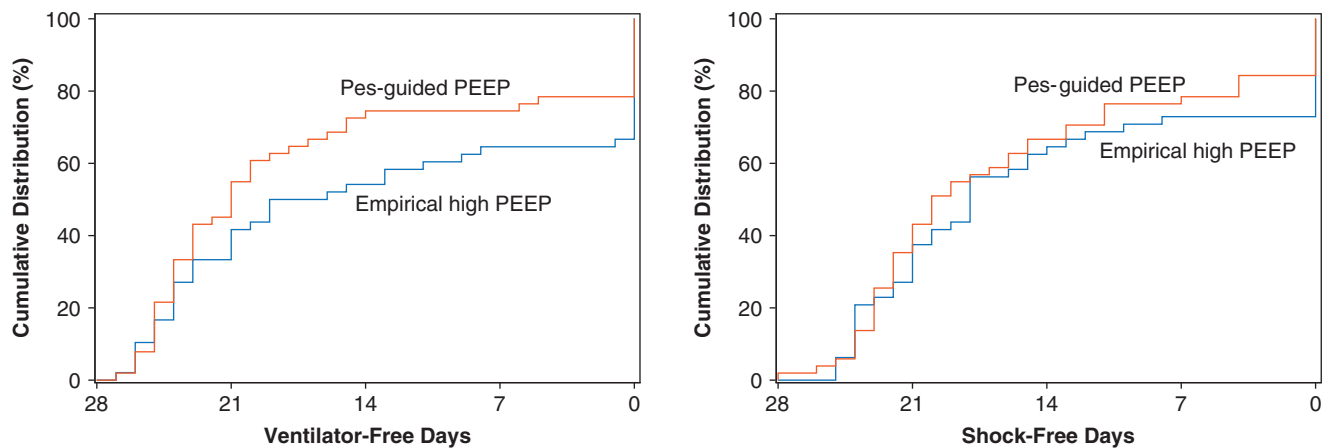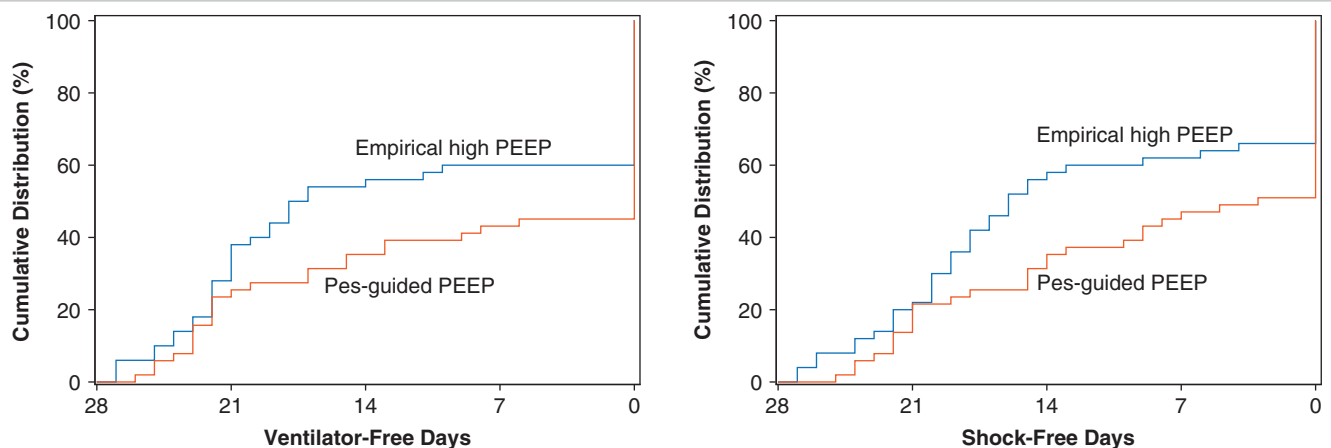
Figure 3. Cumulative distribution of ventilator-free days and shock-free days by treatment group, stratified by APACHE-II. APACHE-II = Acute Physiology and Chronic Health Evaluation II; PEEP = positive end-expiratory pressure; Pes = esophageal pressure.

study days over the first week in patients assigned to Pes-guided PEEP irrespective of APACHE-II score (Figures 4 and E3). Patients assigned to Pes-guided PEEP also were substantially more likely to have end-expiratory P_L within the protective range of ± 2 cm H_2O through Study Day 7 (Figure 4 and Table E4).

Association of End-Expiratory P_L with Mortality

Greater distance of end-expiratory P_L from 0 cm H_2O through Day 3 was associated with increased 60-day mortality (HR 1.10, 95% CI 1.01–1.21 per 1-cm H_2O increase in absolute value of P_L; $P = 0.04$). End-expiratory P_L distance from 0 cm H_2O remained significantly associated with mortality in separate models adjusting for APACHE-II score and treatment group (Table E5). The

association between P_L distance from 0 cm H_2O (absolute value of P_L) and 60-day mortality was not dependent on APACHE-II ($P = 0.91$ for interaction).

Additional models were developed recoding end-expiratory P_L according to whether values were within the hypothesized protective range of ± 2 cm H_2O. Patients within the protective range on average through Study Day 3 had significantly lower mortality (HR 0.48, 95% CI 0.29–0.79; $P < 0.01$). In exploratory pairwise comparisons, patients with end-expiratory P_L within the protective range had lower mortality compared with either those with P_L less than -2 cm H_2O (log-rank $P = 0.04$) or separately with P_L greater than 2 cm H_2O (log-rank $P = 0.02$) (Figure 5). The association between protective-range end-expiratory P_L and lower mortality remained

statistically significant in additional models adjusting for APACHE-II score and treatment group (Figure E4 and Table E5) and was not dependent on APACHE-II score ($P = 0.68$ for interaction with APACHE-II score) (Figure E5).

By contrast, average end-expiratory P_L through Day 3 as a continuous nontransformed variable (not transformed as absolute value) was not associated with 60-day mortality (HR 1.02, 95% CI 0.94–1.10; $P = 0.70$). There was no evidence to suggest nontransformed end-expiratory P_L had a differential effect on mortality dependent on APACHE-II ($P = 0.82$ for interaction).

The relationship between end-expiratory P_L distance from 0 cm H_2O and both ventilator-free and shock-free days is presented in Figure 6.
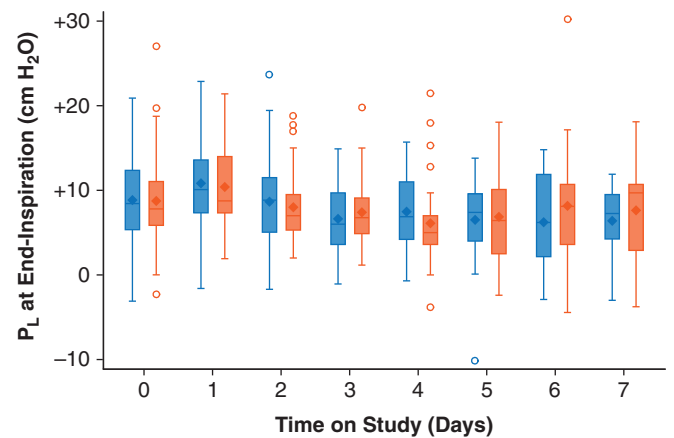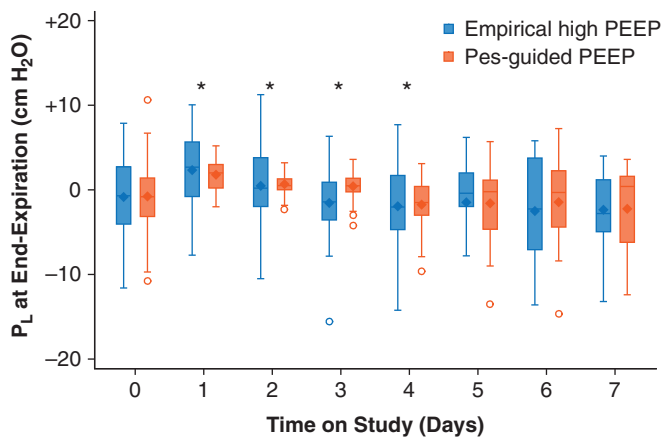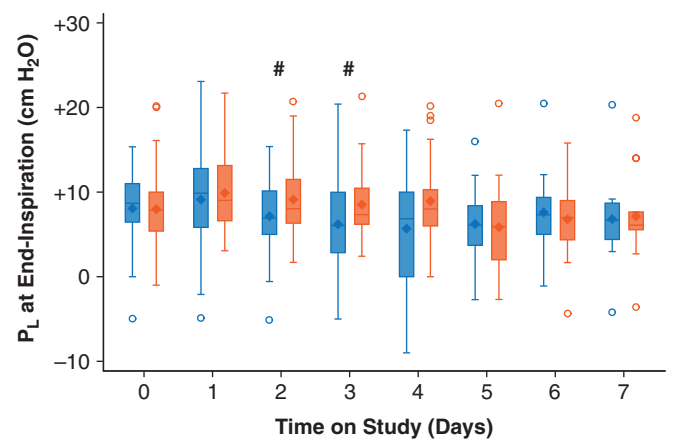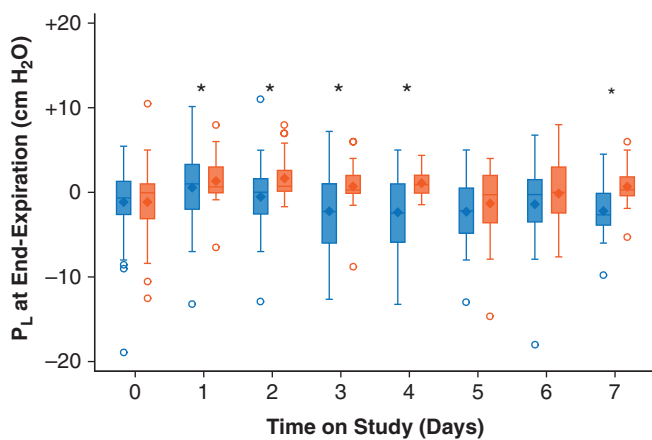
Figure 4. End-expiratory and end-inspiratory P_L over time by treatment assignment. (*A*) Patients with an APACHE-II score less than median value. (*B*) Patients with an APACHE-II score greater than median value. Note the much narrower range of end-expiratory P_L values on each of the first 4 days on protocol with the Pes-guided PEEP strategy, indicative of more precise PEEP titration to P_L. Day 0 denotes baseline preintervention values. Once oxygenation improved and remained stable for 24 hours on minimum ventilator requirements, PEEP was weaned empirically without regard for P_L in either arm, partially contributing to the lower values observed on later study days. *Statistically significant difference between treatment groups in the absolute value of end-expiratory P_L (distance from 0 cm H_2O) on that study day. #Statistically significant difference between treatment groups in P_L at end-inspiration on that study day. APACHE-II = Acute Physiology and Chronic Health Evaluation II; PEEP = positive end-expiratory pressure; Pes = esophageal pressure; P_L = transpulmonary pressure.

Distribution of End-Inspiratory P_L

Baseline end-inspiratory P_L values were reported in 190 of 200 patients and did not differ substantially between treatment groups or by APACHE-II score at baseline (Tables 1 and E1). In both treatment groups, end-inspiratory P_L increased by 2 cm H_2O on average once protocol ventilator settings were introduced. Overall, there was no significant difference in end-inspiratory P_L through Day 7. However, among the subgroup of patients with high APACHE-II scores, end-inspiratory P_L was significantly higher in patients assigned to Pes-guided PEEP during 2 of the first 3 days on study protocol (Figure 4). No significant difference

in end-inspiratory P_L was observed among the subgroup with low APACHE-II scores on any study day.

Association of End-Inspiratory P_L, End-Expiratory P_L, and Outcomes

End-inspiratory P_L averaged through Day 3 was not significantly associated with 60-day mortality in unadjusted analysis (HR 1.03, 95% CI 0.98–1.09; $P = 0.28$).

In the multivariable model including end-inspiratory P_L and APACHE-II score, the association between greater distance of end-expiratory P_L and higher mortality remained statistically significant (HR 1.13, 95% CI 1.02–1.24 per 1-cm H_2O increase in

absolute value of end-expiratory P_L; $P = 0.02$). In this model, higher end-inspiratory P_L also appeared associated with higher mortality, although the relationship did not quite achieve statistical significance (HR 1.05, 95% CI 1.00–1.11 per 1-cm H_2O increase in end-inspiratory P_L; $P = 0.07$).

In generalized linear mixed models, the risk over time of vasopressor-dependent shock was increased with either higher end-expiratory P_L or higher end-inspiratory P_L when entered separately into models (Table E6). When both end-expiratory and end-inspiratory P_L were entered together into the model, higher end-inspiratory P_L remained significantly associated with
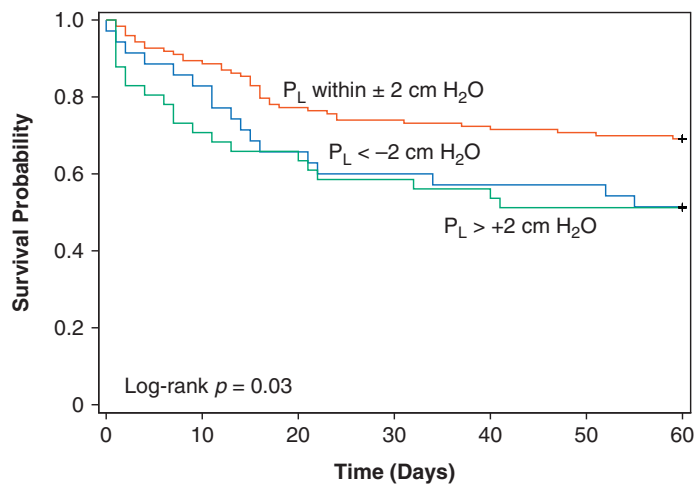
Figure 5. Kaplan-Meier survival analysis by end-expiratory P_L. The average daily value of P_L from baseline through Day 3 was used for analyses. P_L data were missing in one patient, assigned to Pes-guided PEEP, in whom the esophageal balloon catheter could not be inserted successfully. PEEP = positive end-expiratory pressure; Pes = esophageal pressure; P_L = transpulmonary pressure.

greater risk of shock over time, whereas end-expiratory P_L was no longer associated with daily risk of shock (Table E6).

Discussion

In this *post hoc* reanalysis of the EPVent-2 trial, the effect on survival of Pes-guided PEEP compared with empirical high PEEP differed by baseline severity of multiorgan dysfunction. Pes-guided PEEP was associated with greater survival in patients with less multiorgan dysfunction and may have had the opposite effect in patients with severe baseline multiorgan dysfunction. Differences in end-expiratory P_L and hemodynamic instability influenced by tidal overdistension may explain these findings. Independent of treatment group and multiorgan dysfunction, PEEP titration that achieved end-expiratory P_L close to 0 cm H_2O was associated with greater survival than more positive *or* negative end-expiratory P_L values.

The probability that a given intervention may improve chances of survival in the patient with ARDS depends on several factors: *1)* the patient's overall risk of death; *2)* the extent to which risk of death is driven by ARDS versus other factors (disease-attributable risk); *3)* the extent to which risk of death from ARDS is driven by the mechanism targeted by the intervention (mechanism-attributable risk); *4)* efficacy of the intervention to modify the mechanism-attributable risk; and *5)* risk of off-target effects from the intervention.

In patients with severe ARDS without extensive extrapulmonary multiorgan dysfunction, risk of death is attributable primarily to lung injury, and precise PEEP titration might be likelier to improve outcomes. In patients with severe multiorgan failure, heightened risk of adverse hemodynamic effects with higher PEEP and/or resulting tidal overdistension may exceed any lung-protective benefit. Differences in disease-attributable risk and susceptibility to off-target effects thus might explain the heterogeneity of treatment effect by APACHE-II score in EPVent-2.

The mechanistic analysis by P_L suggests another potential explanation. Higher end-inspiratory P_L, indicative of tidal hyperinflation, was associated with increased risk of shock. Among patients with high APACHE-II scores, end-inspiratory P_L was significantly higher in the Pes-guided PEEP

group on 2 of the first 3 days on protocol, suggesting tidal hyperinflation could have contributed to poorer outcomes in the subgroup of patients with high APACHE-II scores. No meaningful difference in end-inspiratory P_L was noted on protocol among patients with low APACHE-II scores.

Independent of APACHE-II score or treatment group, risk of death was less the closer end-expiratory P_L was to 0 cm H_2O, consistent with preclinical data and pathophysiological understanding of the competing goals to mitigate both atelectrauma and hyperinflation with PEEP titration (12, 15–18). Although not strictly targeted in the Pes-guided strategy, end-expiratory P_L was nearer to 0 cm H_2O in the Pes-guided PEEP group, suggesting a treatment-driven protective effect.

Whether concomitantly reducing V_T when increasing PEEP might mitigate risk and maximize patient benefit warrants consideration in future studies. Until then, high PEEP should be used with caution in patients with severe shock, particularly if the chosen PEEP generates end-expiratory P_L far exceeding 0 cm H_2O.

Limitations
This analysis was not prespecified in the trial protocol. *Post hoc* subgroup analyses are often viewed with skepticism (19). Still, analyzing trial results for risk-based treatment effect heterogeneity is essential to interpreting findings (8, 20) and weighing whether they apply to a given patient in clinical practice (21).

The trial's modest sample size of 200 participants may underpower to detect other clinically important differences and lessen precision of effect estimates. Statistical analyses did not correct for multiple testing, risking type I error inflation. All statistical procedures conducted after the primary analyses were sensitivity analyses to evaluate robustness of findings and guide interpretation of main results.

Findings of APACHE-II–dependent heterogeneity of treatment effect on secondary outcomes ventilator-free and shock-free days should be interpreted cautiously. Poisson regression was chosen *a priori* for modeling ventilator-free and shock-free days to preserve interpretability and facilitate tests of interaction. However, there is no consensus on how best to model "failure-free days" outcomes (22), and treatment effects on ventilator-free and shock-free days were not robust to sensitivity
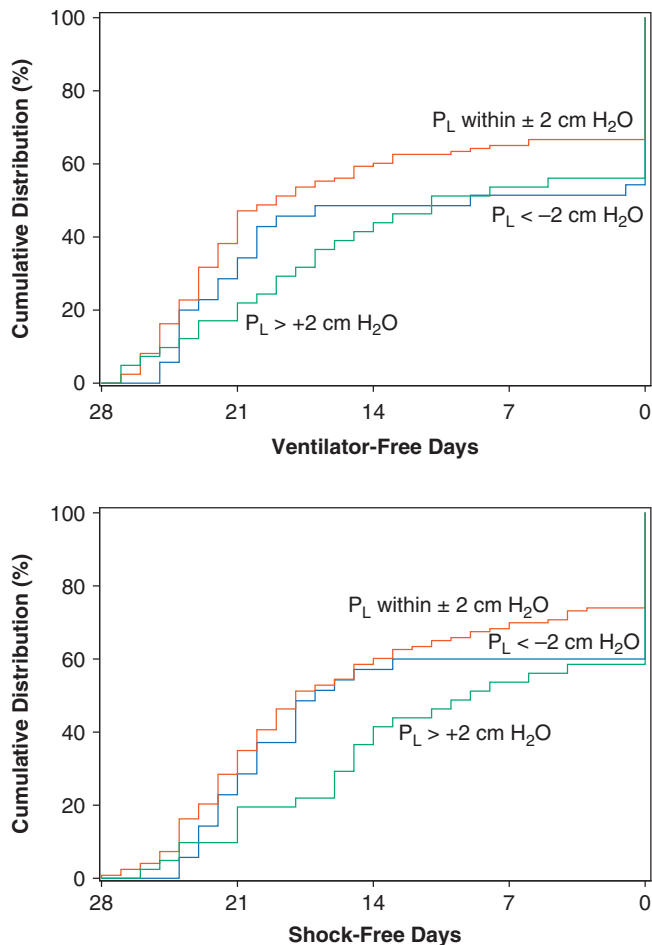
Figure 6. Cumulative distribution of ventilator-free days and shock-free days by end-expiratory P_L. P_L data were missing in one patient, assigned to Pes-guided PEEP, in whom the esophageal balloon catheter could not be inserted successfully. PEEP = positive end-expiratory pressure; Pes = esophageal pressure; P_L = transpulmonary pressure.

Conclusions

Treatment effects in the EPVent-2 trial depended on baseline severity of multiorgan dysfunction. Pes-guided PEEP was associated with greater survival among patients with less severe baseline extrapulmonary organ dysfunction. Pes-guided PEEP also may have exacerbated hemodynamic instability in those with baseline severe shock via tidal hyperinflation. Independent of treatment assignment and multiorgan dysfunction, lower risk of death was observed among patients for whom PEEP achieved an end-expiratory P_L near 0 cm H_2O. This *post hoc* analysis was conceived after trial completion, and results should be interpreted as hypothesis-generating. These findings warrant further examination prospectively in a clinical trial designed to account for baseline heterogeneity of multiorgan dysfunction, in which the intervention sets PEEP to achieve end-expiratory P_L near 0 cm H_2O while titrating inspiratory support to attenuate overdistension. ∎

analyses using alternative analytic approaches. As with other analyses, findings should be considered hypothesis-generating and warrant prospective testing in a future trial.

The trial's Pes-guided PEEP strategy allowed end-expiratory P_L up to 6 cm H_2O with submaximal F_{IO_2}, but the present findings suggest end-expiratory P_L nearer 0 cm H_2O might be more protective (1). The Pes-guided PEEP strategy also allowed end-inspiratory P_L as high as 20 cm H_2O, a value comparable to that observed at total lung capacity in healthy individuals and which might exceed the safe limit in patients at risk of biophysical lung injury (3). Adopting a lower limit before reducing V_T to attenuate overdistension might enhance lung protection with Pes-guided PEEP.

Positive end-expiratory P_L could be a marker of lung injury severity without causal relation to outcome, as higher PEEP and end-expiratory P_L were targeted per protocol when higher F_{IO_2} was needed to maintain oxygenation. However, end-expiratory P_L was not linearly associated with outcome. Rather, mortality increased the further end-expiratory P_L deviated from 0 cm H_2O in the positive *or* negative direction.

APACHE-II includes factors that are not direct markers of multiorgan dysfunction (*e.g.*, age, temperature), although similar heterogeneity of treatment effect was observed by baseline SOFA score, another measure of multiorgan dysfunction. Still, differences in illness severity scores could underlie other occult subpopulation differences (*e.g.*, ARDS precipitant, immune response to lung injury, or regional strain distribution) that might be the main drivers of differential treatment effects.

References

1. Madahar P, Talmor D, Beitler JR. Transpulmonary pressure–guided ventilation to attenuate atelectrauma and hyperinflation in acute lung injury. *Am J Respir Crit Care Med* 2021;203:934–937.
2. Chiumello D, Carlesso E, Cadringher P, Caironi P, Valenza F, Polli F, *et al.* Lung stress and strain during mechanical ventilation for acute respiratory distress syndrome. *Am J Respir Crit Care Med* 2008;178: 346–355.
3. Beitler JR, Majumdar R, Hubmayr RD, Malhotra A, Thompson BT, Owens RL, *et al.* Volume delivered during recruitment maneuver predicts lung stress in acute respiratory distress syndrome. *Crit Care Med* 2016;44:91–99.
4. Loring SH, O'Donnell CR, Behazin N, Malhotra A, Sarge T, Ritz R, *et al.* Esophageal pressures in acute lung injury: do they represent artifact or useful information about transpulmonary pressure, chest wall mechanics, and lung stress? *J Appl Physiol (1985)* 2010;108:515–522.
5. Talmor D, Sarge T, O'Donnell CR, Ritz R, Malhotra A, Lisbon A, *et al.* Esophageal and transpulmonary pressures in acute respiratory failure. *Crit Care Med* 2006;34:1389–1394.
6. Loring SH, Topulos GP, Hubmayr RD. Transpulmonary pressure: the importance of precise definitions and limiting assumptions. *Am J Respir Crit Care Med* 2016;194:1452–1457.
7. Beitler JR, Sarge T, Banner-Goodspeed VM, Gong MN, Cook D, Novack V, *et al.*; EPVent-2 Study Group. Effect of titrating positive end-expiratory pressure (PEEP) with an esophageal pressure-guided strategy vs an empirical high PEEP-FiO2 strategy on death and days free from mechanical ventilation among patients with acute respiratory distress syndrome: a randomized clinical trial. *JAMA* 2019;321:846–857.
8. Kent DM, Hayward RA. Limitations of applying summary results of clinical trials to individual patients: the need for risk stratification. *JAMA* 2007; 298:1209–1212.
9. Ioannidis JPA, Lau J. The impact of high-risk patients on the results of clinical trials. *J Clin Epidemiol* 1997;50:1089–1098.
10. Ioannidis JPA, Lau J. Heterogeneity of the baseline risk within patient populations of clinical trials: a proposed evaluation algorithm. *Am J Epidemiol* 1998;148:1117–1126.
11. Fish E, Novack V, Banner-Goodspeed VM, Sarge T, Loring S, Talmor D. The Esophageal Pressure-Guided Ventilation 2 (EPVent2) trial protocol: a multicentre, randomised clinical trial of mechanical ventilation guided by transpulmonary pressure. *BMJ Open* 2014;4:e006356.
12. Beitler JR, Talmor D. Strategies to adjust positive end-expiratory pressure in patients with ARDS-reply. *JAMA* 2019;322:580–582.
13. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Crit Care Med* 1985;13: 818–829.
14. Ferreira FL, Bota DP, Bross A, Mélot C, Vincent JL. Serial evaluation of the SOFA score to predict outcome in critically ill patients. *JAMA* 2001; 286:1754–1758.
15. Loring SH, Pecchiari M, Della Valle P, Monaco A, Gentile G, D'Angelo E. Maintaining end-expiratory transpulmonary pressure prevents worsening of ventilator-induced lung injury caused by chest wall constriction in surfactant-depleted rats. *Crit Care Med* 2010;38: 2358–2364.
16. Muscedere JG, Mullen JB, Gan K, Slutsky AS. Tidal ventilation at low airway pressures can augment lung injury. *Am J Respir Crit Care Med* 1994;149:1327–1334.
17. Behazin N, Jones SB, Cohen RI, Loring SH. Respiratory restriction and elevated pleural and esophageal pressures in morbid obesity. *J Appl Physiol (1985)* 2010;108:212–218.
18. Fumagalli J, Berra L, Zhang C, Pirrone M, Santiago RRS, Gomes S, *et al.* Transpulmonary pressure describes lung morphology during decremental positive end-expiratory pressure trials in obesity. *Crit Care Med* 2017;45:1374–1381.
19. Assmann SF, Pocock SJ, Enos LE, Kasten LE. Subgroup analysis and other (mis)uses of baseline data in clinical trials. *Lancet* 2000;355: 1064–1069.
20. Iwashyna TJ, Burke JF, Sussman JB, Prescott HC, Hayward RA, Angus DC. Implications of heterogeneity of treatment effect for reporting and analysis of randomized trials in critical care. *Am J Respir Crit Care Med* 2015;192:1045–1051.
21. Rothwell PM. Treating individuals 2. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. *Lancet* 2005;365:176–186.
22. Yehya N, Harhay MO, Curley MAQ, Schoenfeld DA, Reeder RW. Reappraisal of ventilator-free days in critical care research. *Am J Respir Crit Care Med* 2019;200:828–836.