Emulating a Novel Clinical Trial Using Existing Observational Data

Predicting Results of the PreVent Study

Andrew J. Admon[1,2], John P. Donnelly[2,3,4], Jonathan D. Casey[5], David R. Janz[6], Derek W. Russell[7], Aaron M. Joffe[8], Derek J. Vonderhaar[9,10], Kevin M. Dischert[9], Susan B. Stempek[11], James M. Dargin[11], Todd W. Rice[5], Theodore J. Iwashyna[1,2,4,12‡], and Matthew W. Semler[5]; on behalf of the Pragmatic Critical Care Research Group*

[1]Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, [2]Institute for Healthcare Policy and Innovation, [3]Department of Learning Health Sciences, and [12]Institute for Social Research, University of Michigan, Ann Arbor, Michigan; [4]Veterans Affairs Center for Clinical Management Research, Health Services Research and Development Center of Innovation, Ann Arbor, Michigan; [5]Division of Pulmonary, Allergy, and Critical Care Medicine, Vanderbilt University Medical Center, Nashville, Tennessee; [6]Section of Pulmonary/Critical Care & Allergy/Immunology, Louisiana State University School of Medicine, New Orleans, Louisiana; [7]Division of Pulmonary, Allergy, & Critical Care Medicine, University of Alabama at Birmingham, Birmingham, Alabama; [8]Department of Anesthesiology and Pain Medicine, University of Washington, Seattle, Washington; [9]Department of Pulmonary and Critical Care Medicine, Ochsner Health System New Orleans, New Orleans, Louisiana; [10]Department of Medicine, Section of Emergency Medicine, Louisiana State University School of Medicine–New Orleans, New Orleans, Louisiana; and [11]Department of Medicine, Division of Pulmonary and Critical Care Medicine, Lahey Hospital and Medical Center, Burlington, Massachusetts

ORCID ID: 0000-0002-7432-3764 (A.J.A.).

Abstract

Rationale: "Target trial emulation" has been proposed as an observational method to answer comparative effectiveness questions, but it has rarely been attempted concurrently with a randomized clinical trial (RCT).

Objectives: We tested the hypothesis that blinded analysts applying target trial emulation to existing observational data could predict the results of an RCT.

Methods: PreVent (Preventing Hypoxemia with Manual Ventilation during Endotracheal Intubation) was a multicenter RCT examining the effects of positive-pressure ventilation during tracheal intubation on oxygen saturation and severe hypoxemia. Analysts unaware of PreVent's results used patient-level data from three previous trials evaluating airway management interventions to emulate PreVent's eligibility criteria, randomization procedure, and statistical analysis. After PreVent's release, results of this blinded observational analysis were compared with those of the RCT. Difference-in-differences estimates for comparison of treatment effects between the observational analysis and the PreVent trial are reported on the absolute scale.

Results: Using observational data, we were able to emulate PreVent's randomization procedure to produce balanced groups for comparison. The lowest oxygen saturation during intubation was higher in the positive-pressure ventilation group than the no positive-pressure ventilation group in the observational analysis ($n = 360$; mean difference = 1.8%; 95% confidence interval [CI] = −1.0 to 4.6) and in the PreVent trial ($n = 401$; mean difference = 3.9%; 95% CI = 1.4 to 6.4), though the observational analysis could not exclude no difference. Difference-in-differences estimates comparing treatment effects showed reasonable agreement for lowest oxygen saturation between the observational analysis and the PreVent trial (mean difference = −2.1%; 95% CI = −5.9 to 1.7). Positive-pressure ventilation resulted in lower rates of severe hypoxemia in both the observational analysis (risk ratio = 0.60; 95% CI = 0.38 to 0.93) and in the PreVent trial (risk ratio = 0.48; 95% CI = 0.30 to 0.77). The absolute reduction in the incidence of severe hypoxemia with positive-pressure ventilation was similar in the observational analysis (9.4%) and the PreVent trial (12.0%), though the difference between these estimates had wide CIs (mean difference = 2.5%; 95% CI = −8.0 to 13.6%).

Conclusions: Applying target trial emulation methods to existing observational data for the evaluation of a novel intervention produced results similar to those of a randomized trial. These findings support the use of target trial emulation for comparative effectiveness research.

Keywords: clinical trials; intubation; epidemiology; causal inference; target trial emulation

(Received in original form March 18, 2019; accepted in final form April 29, 2019)

*A complete list of members of the Pragmatic Critical Care Research Group Investigators may be found in the online supplement.

‡T.J.I. is a Section Editor of *AnnalsATS*. His participation complies with American Thoracic Society requirements for recusal from review and decisions for authored works.

Supported by U.S. National Heart, Lung, and Blood Institute grants T32HL007749 (A.J.A.), K12HL138039 (J.P.D.), T32HL087738 (J.D.C.), and K23HL143053 (M.W.S.), and by U.S. Department of Veterans Affairs Health Services Research and Development grant 17-045 (T.J.I.).

Ann Am Thorac Soc Vol 16, No 8, pp 998–1007, Aug 2019

DOI: 10.1513/AnnalsATS.201903-241OC
Internet address: www.atsjournals.org

Randomized clinical trials (RCTs) represent the strongest evidence for comparing the effectiveness of two interventions because randomization of large populations ensures that other factors that influence the outcome are equally distributed between the treatment and control groups. However, RCTs are costly, slow, and often impractical to generate evidence for many important questions (1, 2). Observational studies may fill this gap, though confounding and bias are persistent risks (3–5).

Applying modern epidemiological methods to detailed data from completed clinical trials may reduce the likelihood of misleading results in observational comparative effectiveness studies (6, 7). "Target trial emulation" is a method in which investigators explicitly mimic an idealized clinical trial using observational techniques (8–10). Because data from prior clinical trials typically include detailed, accurately collected clinical information, they may allow investigators to better control for relevant confounders and emulate the design features and "randomization" of an idealized RCT (11, 12). These methods may be particularly helpful for conducting observational studies of therapies in critical care where clinician judgment is a strong determinant of treatment received (13). Few target trial emulations have employed detailed data from completed clinical trials, and even fewer have occurred concurrently with the clinical trial they seek to emulate, rendering direct comparison of findings between the approaches difficult.

In this study, we combined data from several completed clinical trials examining unrelated airway management interventions to estimate the effects of a novel intervention: positive-pressure ventilation during tracheal intubation of critically ill adults. Whether positive-pressure ventilation should be provided between induction and laryngoscopy during tracheal intubation has been debated for decades (14). Positive-pressure ventilation has been hypothesized to prevent hypoxemia (a contributor to cardiac arrest and death), but has also been hypothesized to increase the risk of oropharyngeal or gastric aspiration. Because an operator's decision regarding whether to administer positive-pressure ventilation may be based on perceived risk of periprocedural hypoxemia and aspiration, observational studies of bag mask ventilation are highly prone to confounding. To overcome this limitation, we applied causal methods and target trial emulation to predict the results of the PreVent (Preventing Hypoxemia with Manual Ventilation during Endotracheal Intubation) trial before its release (15). To ensure a fair evaluation, investigators performing the observational analysis (A.J.A., J.P.D., and T.J.I.) were kept unaware of the results of the PreVent trial, and registered both their analytical plan and the results of their observational analyses before the PreVent trial results were published (16). We hypothesized that the effect of positive-pressure ventilation on lowest oxygen saturation and severe hypoxemia would not differ between the observational study cohort and the randomized trial.

Methods

Target trial emulation is an observational research method that incorporates design features from idealized randomized trials (i.e., target trials) to improve the quality and interpretability of observational research (8–10, 12). Specifically, this approach requires careful design of inclusion and exclusion criteria similar to those that would have been used in the target trial and consideration of the mode of treatment allocation and the timing and intensity of the exposure. Investigators recreate these trial design features in observational data and use statistical methods that capitalize on natural variability in provider practice to emulate random treatment assignment.

On February 7, 2019, investigators from the PreVent trial provided the PreVent study protocol and deidentified patient-level data from three prior randomized trials of unrelated airway management interventions to a team of observational analysts (A.J.A., J.P.D., and T.J.I.) who did not know the results of the PreVent study (17). The analysts registered an analytical plan, statistical code, and results of their observational analysis (*see* the online supplement) with an honest broker on February 17, 2019, before the results of the PreVent trial were released (16). This approach allowed examination of whether causal methods applied to rich observational data yield similar results to a clinical trial, while minimizing the risk of bias that could be introduced by observational analysts knowing the results of the gold standard trial. An overview of the design decisions made in the observational analysts to emulate the randomized trial is shown in Table 1.

Description of the Target Trial

The observational analysis was intended to emulate PreVent, a multicenter, parallel-group, unblinded, pragmatic, randomized trial comparing positive-pressure ventilation with a bag-mask device to no positive-pressure ventilation between induction and laryngoscopy during tracheal intubation of critically ill adults. Briefly, PreVent

Author Contributions: A.J.A. had full access to all of the data in the study and takes responsibility for the integrity of the data and accuracy of the data analysis; study concept and design—A.J.A., J.P.D., J.D.C., T.J.I., and M.W.S.; acquisition, analysis, or interpretation of the data—A.J.A., J.P.D., J.D.C., D.R.J., D.W.R., A.M.J., D.J.V., K.M.D., S.B.S., J.M.D., T.W.R., T.J.I., and M.W.S.; drafting of the manuscript—A.J.A., J.P.D., J.D.C., T.J.I., and M.W.S.; critical revision of the manuscript for important intellectual content—A.J.A., J.P.D., J.D.C., D.R.J., D.W.R., A.M.J., D.J.V., K.M.D., S.B.S., J.M.D., T.W.R., T.J.I., and M.W.S.; statistical analysis—A.J.A. and J.P.D.; obtained funding—M.W.S.; administrative, technical, or material support—J.D.C., D.R.J., D.W.R., A.M.J., D.J.V., K.M.D., S.B.S., J.M.D., T.W.R., and M.W.S.; study supervision—M.W.S.

The content is the responsibility of the authors alone and does not necessarily reflect the views or policies of the Department of Veterans Affairs or the U.S. Government.

Correspondence and requests for reprints should be addressed to Andrew J. Admon, M.D., M.P.H., University of Michigan, 2800 Plymouth Road., Bldg. 16-169c, Ann Arbor, MI 48109. E-mail: ajadmon@umich.edu.

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org.

Table 1. Comparison of target trial and observational analysis

Characteristic	Target Trial (PreVent)	Observational Analysis
Eligibility criteria	Included Admitted in a unit participating in PreVent Intubation planned 18 years old or older Excluded Pregnant Incarcerated Need for tracheal intubation too emergent Clinician deemed positive-pressure ventilation to be required (hypoxemia, severe acidemia, respiratory arrest) or contraindicated (active emesis, hematemesis, hemoptysis)	Included Admitted in a unit participating in Check-UP, FELLOW, or PrePARE Intubation planned 18 years old or older Any operator Excluded Pregnant Incarcerated Need for tracheal intubation too emergent Clinician deemed positive-pressure ventilation to be required (hypoxemia, severe acidemia, respiratory arrest) Sp_{O_2} at baseline <90% (49 people)
Treatment Strategies	-Intervention: positive-pressure ventilation with a bag-mask device (modified RSI) -Control: No positive-pressure ventilation except after a failed attempt or for Sp_{O_2} <90% (classic RSI)	-Exposure: positive-pressure ventilation with a bag-mask device or noninvasive ventilator (modified RSI) -Control: No positive-pressure ventilation except after a failed attempt or for Sp_{O_2} <90% (classic RSI)
Assignment procedures	Randomization in a 1:1 ratio to positive-pressure ventilation or no positive-pressure ventilation using permuted blocks of two, four, and six stratified by study site.	1. Exposure known for 125 patients (PrePARE patients) 2. For others, exposure imputed based on the following rules: a. Patients intubated on the first attempt without receiving bag-mask or noninvasive ventilation were assigned control group b. Patients who received bag-mask or noninvasive ventilation and were intubated on the first attempt were assigned intervention group, unless they had a prolonged intubation (top decile of intubation length) c. Patients who received bag-mask or noninvasive ventilation and required more than one intubation attempt or had a prolonged intubation were assigned using a predictive model based on known recipients of exposure
Follow-up period for primary outcome	From induction until 2 min after tracheal intubation	From induction until 2 min after tracheal intubation
Causal contrasts of interest	Lowest oxygen saturation between induction and 2 min after tracheal intubation	Lowest oxygen saturation between Induction and 2 min after tracheal intubation
	Other outcomes: Severe hypoxemia (Sp_{O_2} < 80%) Lowest Sp_{O_2}, highest FI_{O_2}, and highest PEEP between 6 and 24 h Operator reported aspiration New opacity on chest radiography within 48 h	Other outcomes: Severe hypoxemia (Sp_{O_2} < 80%) Lowest Sp_{O_2}, highest FI_{O_2}, and highest PEEP between 6 and 24 h Operator-reported aspiration
Analysis plan	Primary analysis: Mann-Whitney *U* test	Primary analysis: Mann-Whitney *U* test on propensity score matched population

Definition of abbreviations: Check-UP = Checklists and Upright Positioning in endotracheal intubation of critically ill patients; FELLOW = Facilitating EndotracheaL Intubation by Laryngoscopy Technique and Apneic Oxygenation within the Intensive Care Unit; FI_{O_2} = fraction of inspired oxygen; PEEP = positve end-expiratory pressure; PrePARE = Preventing Cardiovascular Collapse with Administration of Fluid Resuscitation before Tracheal Intubation; PreVent = Preventing Hypoxemia with Manual Ventilation during Endotracheal Intubation; RSI = rapid sequence intubation; Sp_{O_2} = oxygen saturation as measured by pulse oximetry.

(registered with www.clinicaltrials.gov [NCT03026322]) enrolled 401 patients between March 15, 2017 and May 6, 2018, with complete details of the methods and results reported February 18, 2019 (15).

Data Sources

Data for the observational analysis were obtained from three completed randomized trials evaluating airway management interventions among patients similar to PreVent's intended study population. The Check-UP (Checklists and Upright Positioning in endotracheal intubation of critically ill patients) study, was a randomized,

multicenter, pragmatic, two-by-two factorial trial comparing *1*) the ramped position with the sniffing position and *2*) the use of a written preintubation checklist during endotracheal intubation of critically ill adults (18). The FELLOW (Facilitating EndotracheaL Intubation by Laryngoscopy Technique and Apneic Oxygenation within the Intensive Care Unit) study was a randomized, open-label, parallel-group, pragmatic, two-by-two factorial trial comparing apneic oxygenation with usual care and direct laryngoscopy with video laryngoscopy among critically ill adults (19). The PrePARE (Preventing Cardiovascular Collapse with Administration of Fluid Resuscitation before Tracheal Intubation) trial compared the effects of a fluid bolus administered before induction versus no preinduction fluid bolus on the incidence of cardiovascular collapse during tracheal intubation (registered with www.clinicaltrials.gov [NCT03026777]). Although patients could be coenrolled in PrePARE and PreVent, the only data from the PrePARE trial used in the current analysis were those of patients at centers not actively enrolling in PreVent. Despite being comprised of data from randomized trials, the cohort derived from these data was referred to as "observational" because the specific intervention of interest—positive-pressure ventilation between induction and laryngoscopy—was allocated based on clinician choice rather than by randomization.

Eligibility Criteria

To derive the cohort for the observational analysis, we included adults (≥18 yr of age) enrolled in Check-UP (Checklists and Upright Positioning in endotracheal intubation of critically ill patients), FELLOW, or PrePARE (Table 1). We excluded patients known to have met exclusion criteria for PreVent (e.g., active emesis, brisk upper gastrointestinal bleeding, or unstable facial fractures). This led to the exclusion of 43 patients from PrePARE, the only trial in which these data were collected. We also excluded patients with an oxygen saturation at induction of less than 90% as a technique for discriminating positive-pressure ventilation used for rescue (which was allowed in both groups in PreVent after a failed attempt or for treatment of oxygen saturations <90%) from *preventive* positive-pressure ventilation (*see* Assignment Procedures).

Treatment Strategies

The PreVent trial compared positive-pressure ventilation with a bag-mask device between induction and laryngoscopy (sometimes referred to as modified rapid sequence intubation [RSI]) to no positive-pressure ventilation between induction and laryngoscopy except as treatment for hypoxemia or after a failed intubation attempted (conventional RSI). Although positive-pressure ventilation between induction and laryngoscopy may be delivered with either a bag-mask device or a noninvasive ventilator, to limit imprecision in the delivery of ventilation, all ventilation in the PreVent trial was required to be delivered using a bag-mask device. Because clinicians could deliver positive-pressure ventilation during intubation using either a bag-mask device or noninvasive ventilation in the observational cohort, we included both modalities of positive-pressure ventilation in the treatment group in the observational analysis.

Assignment Procedures

We sought to emulate the PreVent trial's randomization to positive-pressure ventilation at induction using data from our observational cohort. This was done in two steps: first by imputing treatment status at induction among patients for whom timing of treatment was uncertain, and second by using coarsened exact matching to identify comparable groups.

Our exposure of interest was positive-pressure ventilation beginning at induction, which was prospectively recorded in the PreVent and PrePARE trials, but the Check-UP and FELLOW trials captured only use of positive-pressure ventilation at *any time* from induction to intubation. This included patients who received positive-pressure ventilation at induction (the intervention studied in PreVent) and those who received rescue ventilation for treatment of hypoxemia or between successive intubation attempts (which was permitted in *either* arm in PreVent).

As a result, timing of positive-pressure ventilation at induction was known only for patients in the PrePARE trial, and in patients in Check-UP and FELLOW who *never* received positive-pressure ventilation between induction and intubation (and so, received conventional RSI). For patients in Check-UP and FELLOW who received positive-pressure ventilation at some point from induction to intubation, we followed a series of rules to determine whether it was started at induction (e.g., modified RSI, the intervention studied in the PreVent trial) or later in the intubation procedure for rescue. First, anyone with a baseline oxygen saturation of less than 90% was excluded from the observational analysis, as these patients would have qualified for rescue positive-pressure ventilation regardless of randomization arm in PreVent. Second, patients who received positive-pressure ventilation and were intubated on the first attempt were assumed to have received it at induction, unless they had a prolonged intubation (in the top 10% of procedure durations). Finally, in the remaining patients (those who received positive-pressure ventilation and either required more than one intubation attempt or had a prolonged intubation), we used single imputation with a model based on those for whom timing of positive-pressure ventilation was known, assigning patients with predicted probabilities of less than 0.5 to the no positive-pressure ventilation at induction (e.g., conventional RSI or control) group. In a sensitivity analysis, we used multiple imputation with 100 imputed datasets.

Next, to mimic the randomization performed in the PreVent trial, we generated propensity scores for receiving positive-pressure ventilation using the independent variables of study site, intervention group in the original completed trial, age, body mass index (BMI), race, sex, presence of sepsis, presence of chronic obstructive pulmonary disease exacerbation, indication for intubation (hypoxia or hypercarbia), highest fraction of inspired oxygen in the prior 6 hours, and baseline oxygen saturation (20). These variables were selected because they were thought to influence an operator's decision to apply positive-pressure ventilation at induction (21). Using these scores, we matched patients in a K:K ratio with coarsened exact matching and verified covariate balance across matched groups.

Follow-Up Period

In both the observational study and the PreVent trial, oxygen saturations were measured from induction through

2 minutes after tracheal intubation. Secondary and safety outcomes were similarly measured for up to 24 hours after intubation.

Causal Contrasts of Interest

The primary outcome in both the observational analysis and the PreVent trial was the lowest arterial oxygen saturation between induction and two minutes after tracheal intubation. This outcome was captured in the same manner in all three trials (Check-UP, FELLOW, and PrePARE) comprising our observational cohort. Our secondary outcome, also matching PreVent, was the proportion of patients with severe hypoxemia, defined as oxygen saturation less than 80%. We also examined each of the PreVent exploratory outcomes (oxygen saturation <70% and 90%, decrease in saturation from induction to nadir, ventilator-free days, intensive care unit–free days, and in-hospital death) and safety outcomes (operator-reported aspiration during intubation and cardiac arrest within 1 h of intubation).

Analysis Plan

To arrive at results that were directly comparable between the observational analysis and PreVent, we applied the statistical analysis plan developed for PreVent to analyses of all outcomes in the observational cohort. We first compared baseline patient characteristics between this observational study and PreVent. After ensuring covariate balance, we used the Mann-Whitney *U* test to compare distributions of nadir oxygen saturation between our treatment and control groups and a series of regression analyses to estimate mean differences or risk ratios (RRs) with 95% confidence intervals (CIs) for each outcome. To more closely emulate the analyses reported in PreVent, linear models used for effect estimation included only one independent variable (treatment assignment). We also estimated differences in effect estimates (difference-in-differences [DIDs]) for outcomes between the observational study and PreVent using bootstrapping with 10,000 replicates and bias-corrected 95% CIs. All tests were two-tailed. Analyses were conducted using Stata 14.2 (StataCorp LLC). This secondary analysis of deidentified data was determined nonhuman subject research by the Vanderbilt Institutional Review Board (no. 160158).

Results

Derivation of the Observational Cohort

The initial cohort for the observational analysis included 567 total patients, including 292 from Check-UP, 150 from FELLOW, and 125 from PrePARE (Figure 1). Among these patients, 348 received positive-pressure ventilation between induction and intubation. Before any adjustment, the median lowest oxygen saturation was 92% among patients who received positive-pressure ventilation compared with 95% among patients who did not receive positive-pressure ventilation (mean difference = −3.23; 95% CI = −5.59 to −0.88; $P = 0.02$).

From these 567 patients, 105 were excluded for baseline oxygen saturation of less than 90%, contraindications to enrollment in PreVent, or missing data. Of the remaining 462 patients, treatment status was known in 374 patients and assigned by imputation in 88 patients. The imputation model for positive-pressure ventilation at induction had adequate discrimination (c-statistic = 0.80) and fit (pseudo-$R^2 = 0.21$). Propensity scores using the updated treatment variable were created and 360 patients were matched (180 in both the treatment and control groups). These groups were well balanced with regard to median age, sex, median BMI, median APACHE (Acute Physiology and Chronic Health Evaluation) II scores, indications for intubation (hypoxemic respiratory failure, hypoxic respiratory failure, or altered mental status), and other variables. (Table 2 and Figure E1 in the online supplement).

Primary Outcome

For the 360 matched patients in the observational target trial emulation, the median lowest oxygen saturation was numerically higher in the positive-pressure ventilation group (94%; interquartile range [IQR] = 86–98%) compared with the no positive-pressure ventilation group (93%; IQR = 82–99%) (mean difference = 1.8 percentage points; IQR = −1.0 to 4.6), though this difference was neither clinically nor statistically significant ($P = 0.76$; Figure 2A).

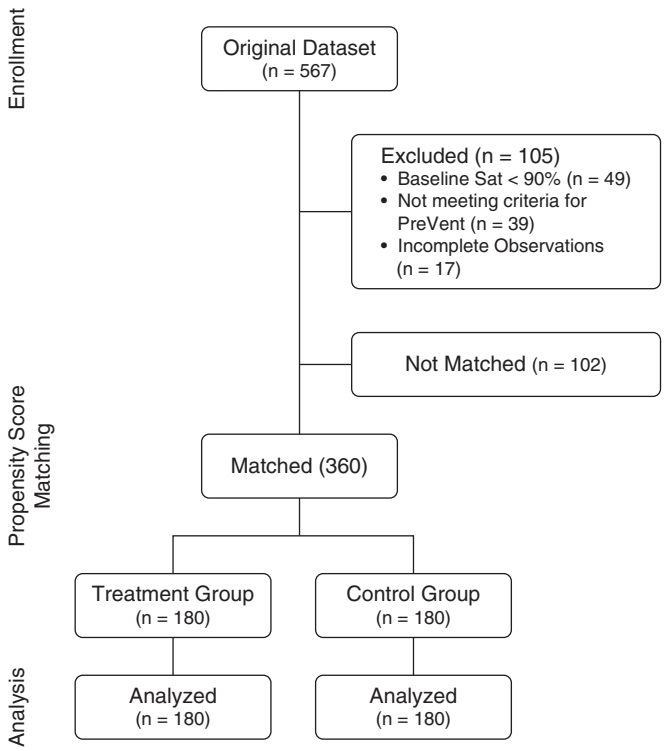


Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram. PreVent = Preventing Cardiovascular Collapse with Administration of Fluid Resuscitation before Tracheal Intubation; Sat = saturation.

Table 2. Patient characteristics at baseline—observational study and PreVent Trial

Patient Characteristic	Observational Study		PreVent	
	Positive-Pressure Ventilation (*n* = 180)	No Ventilation (*n* = 180)	Bag Mask Ventilation (*n* = 199)	No Ventilation (*n* = 202)
Age, median (IQR), yr	60 (49–68)	56 (48–68)	59 (45–67)	60 (48–68)
Male sex, *n* (%)	106 (58.9)	102 (56.7)	118 (59.3)	108 (53.5)
White race, *n* (%)	140 (77.8)	140 (77.8)	141 (71.9)	134 (66.3)
Body mass index, median (IQR), kg/m^2	28.1 (23.4–33.4)	27.4 (23.5–32.2)	27.1 (22.7–32.3)	27.6 (23.4–34.2)
APACHE II score, median (IQR)	20.5 (16–26)	21 (16–25)	22 (16–29)	22 (16–28)
Vasopressors, *n* (%)	27 (15.0)	42 (23.3)	35 (17.6)	40 (19.9)
Active medical conditions, *n* (%)				
Sepsis or septic shock	98 (54.4)	97 (53.9)	98 (49.3)	97 (48.0)
Gastrointestinal bleeding	21 (11.7)	33 (18.3)	28 (14.1)	16 (7.9)
Indications for intubation, *n* (%)				
Hypoxemic respiratory failure	103 (57.2)	101 (56.1)	117 (58.8)	116 (57.4)
Hypercarbic respiratory failure	23 (12.8)	25 (13.9)	39 (19.6)	55 (27.2)
Airway protection for decreased level of consciousness	60 (33.3)	60 (33.3)	74 (37.2)	76 (37.6)
BiPAP in prior 6 h, *n* (%)	71 (39.4)	37 (20.6)	44 (22.1)	57 (28.2)
Highest F_{IO_2} in prior 6 h, median (IQR)	0.4 (0.3–0.7)	0.4 (0.3–0.8)	0.4 (0.3–1.0)	0.5 (0.3–1.0)
Lowest oxygen saturation in prior 6 h, median (IQR), %	91 (88–94)	92 (89–95)	91 (87–95)	92 (88–95)

Definition of abbreviations: APACHE II = Acute Physiology and Chronic Health Evaluation II; BiPAP = bilevel positive airway pressure; F_{IO_2} = fraction of inspired oxygen; IQR = interquartile range; PreVent = Preventing Hypoxemia with Manual Ventilation during Endotracheal Intubation.

Secondary, Safety, and Exploratory Outcomes

Among the 360 matched patients in the observational cohort, fewer patients experienced the secondary outcome of severe hypoxemia in the positive-pressure ventilation group (13.9%) compared with the no positive-pressure ventilation group (23.3%) (RR = 0.6; 95% CI = 0.4 to 0.9; Table 3 and Figure 2B). The observational cohort did not have adequate statistical power to exclude clinically meaningful differences in the safety and exploratory outcomes. For example, rates of operator-reported aspiration during intubation were 1.7% in the positive-pressure ventilation group and 0.6% in the no positive-pressure ventilation group (RR = 3.0; 95% CI = 0.3 to 28.6) and rates of death before hospital discharge were 40% in both treatment and control groups (RR = 1.0; 95% CI = 0.7 to 1.3).

Results of the Observational Analysis Compared with the PreVent Trial

In comparing baseline characteristics of the positive-pressure ventilation group in the observational analysis (the group at highest risk for imbalances in disease severity) to the positive-pressure ventilation group in the PreVent trial, we found that median age (60 yr, IQR = 49–68 vs. 59, IQR = 45 to 67), median BMI (28.1, IQR = 32.4 to 33.5 vs. 27.1, IQR = 22.7 to 32.3), and median APACHE II score (20.5, IQR = 16 to 26 vs. 22.0, IQR = 16 to 29) were similar between the observational analysis and PreVent. There were differences in the rates of noninvasive ventilation in the prior 6 hours before intubation (39.4% vs. 22.1% in the observational analysis vs. PreVent, respectively), likely a result of our inclusion of noninvasive ventilation in the definition of positive-pressure ventilation during induction.

Similar to the findings of the observational analysis, the primary outcome of lowest oxygen saturation during intubation was higher for patients in the positive-pressure ventilation group of the PreVent trial (96%, IQR = 89 to 99%) compared with the no positive-pressure ventilation group of the PreVent trial (93%, IQR = 81 to 98%) (mean difference = 3.9 percentage points, IQR = 1.4 to 6.4). Unlike in the observational analysis, the PreVent trial's difference in the primary outcome between groups met statistical significance ($P = 0.01$). Similar to the findings of the observational analysis, rates of severe hypoxemia in the PreVent trial were lower in the positive-pressure ventilation group (10.9%) than in the no positive-pressure ventilation group (22.8%) (RR = 0.5; 95% CI = 0.3 to 0.8). The differences between positive-pressure ventilation and no positive-pressure ventilation in safety and exploratory outcomes in the observational analysis and PreVent trial are displayed in Table 3.

In our DIDs analyses directly comparing the differences between the intervention and control groups across both studies, we found that the effect of positive-pressure ventilation on the primary outcome of lowest oxygen saturation was similar between the observational analysis and the PreVent trial (DID = −2.1%; 95% CI = −5.9 to 1.7%; Table 4). We also found that the effect of positive-pressure ventilation on rates of severe hypoxemia was similar between the observational analysis and the PreVent trial (DID = 2.5%; 95% CI = −8.1% to 13.6%). There were no substantial differences for any of the other exploratory or safety outcomes between studies, though we could not exclude clinically important differences for several exploratory outcomes including median ventilator-free days (DID = −1.1; 95% CI = −4.5 to 2.2), median intensive care unit–free days (−0.7; 95% CI = −3.7 to 2.4), or death before hospital discharge (0.0%; 95% CI = −13.7 to 13.8).

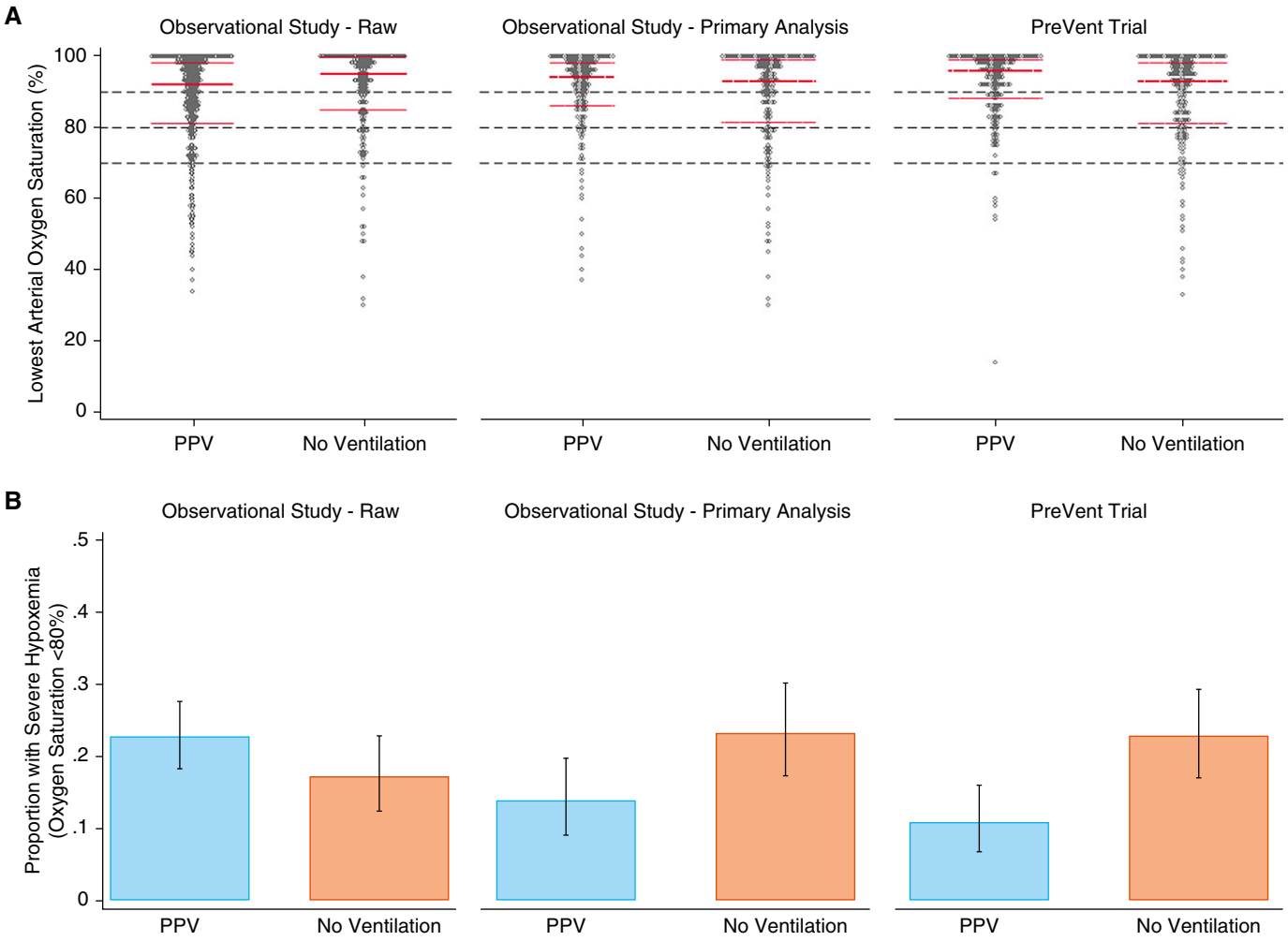


Figure 2. Lowest arterial oxygen saturation and severe hypoxemia in treatment versus control groups. (*A*) Dot plots showing the distributions of lowest oxygen saturation by group. Bold line represents the median; thinner lines represent the 25th and 75th percentile. (*B*) Bar charts showing the proportion with severe hypoxemia (lowest oxygen saturation <80%) by group. Error bars represent 95% confidence interval limits. PPV = positive pressure ventilation; PreVent = Preventing Hypoxemia with Manual Ventilation during Endotracheal Intubation.

Discussion

This study found that blinded application of target trial emulation analysis to observational data accurately predicted the results of a critical care clinical trial. The effects of positive-pressure ventilation on lowest oxygen saturation and severe hypoxemia during tracheal intubation of critically ill adults were similar in our analysis of observational data and in the subsequently published PreVent randomized trial. Our analysis of observational data, however, was unable to exclude the possibility of no difference between groups in the primary outcome of lowest oxygen saturation, and neither the observational cohort nor the clinical trial provided adequate power to assess rare safety or exploratory outcomes. To our knowledge, this is the first study to directly compare results of a blinded target trial emulation using observational data to the realized target trial. Collectively, the similarities between our results and those of the PreVent trial suggest that target trial emulation using existing observational data may offer a powerful tool for comparative effectiveness research using rich, clinical data.

Observational studies of critical care interventions face unique challenges due to confounding by indication, whereby providers allocate a treatment nonrandomly in a way that is dependent on other variables (e.g., clinical trajectory or severity of illness), which themselves influence the outcome of interest (22). In the absence of effective statistical adjustment (using propensity score matching, regression modeling, inverse probability of treatment weighting, or other techniques), confounding by indication can produce misleading results in observational comparative effectiveness studies (23). The influence of these biases was evident in our unadjusted analysis, where positive-pressure ventilation was associated with a significantly lower oxygen saturation than no positive-pressure ventilation—the opposite finding of our adjusted analysis and the PreVent trial. Despite this imbalance in the raw data, we were able to arrive at closely balanced intervention and control groups using

Table 3. Target trial outcomes—observational cohort and PreVent Trial

Outcomes	Observational Study			PreVent Trial		
	Positive-Pressure Ventilation (*n* = 180)	No Ventilation (*n* = 180)	RR or Mean Difference (95% CI)	Bag Mask Ventilation (*n* = 199)	No Ventilation (*n* = 202)	RR or Mean Difference (95% CI)
Primary outcome						
Lowest oxygen saturation, median (IQR), %	94 (86 to 98)	93 (81.5 to 99)	1.8 (−1.0 to 4.6)	96 (88 to 99)	93 (81 to 98)	3.9 (1.4 to 6.4)
Secondary outcome						
Lowest oxygen saturation <80%, *n* (%)	25/180 (13.9)	42/180 (23.3)	0.60 (0.4 to 0.9)	21/193 (10.9)	45/197 (22.8)	0.48 (0.3 to 0.8)
Exploratory oxygen saturation outcomes						
Lowest oxygen saturation <90%, *n* (%)	59/180 (32.8)	66/180 (36.7)	0.89 (0.7 to 1.2)	57/193 (29.5)	79/197 (40.1)	0.74 (0.6 to 1.0)
Lowest oxygen saturation <70%, *n* (%)	12/180 (6.7)	19/180 (10.6)	0.63 (0.3 to 1.3)	8/193 (4.2)	20/197 (10.2)	0.41 (0.2 to 0.9)
Decrease in oxygen saturation, median (IQR), %	3 (0 to 11)	3 (0 to 13)	−1.9 (−4.5 to 0.8)	1 (0 to 7)	5 (0 to 14)	−4.5 (−6.8 to −2.2)
Exploratory safety outcomes						
Operator-reported aspiration, *n* (%)	3/180 (1.7)	1/180 (0.6)	3.00 (0.3 to 28.6)	5/198 (2.5)	8/202 (4.0)	0.64 (0.2 to 1.9)
Cardiac arrest 1 h after intubation, *n* (%)	4/180 (2.2)	2/180 (1.1)	2.00 (0.4 to 10.8)	3/199 (1.5)	4/202 (2.0)	0.76 (0.2 to 3.4)
Exploratory clinical outcomes						
Ventilator-free days, median (IQR)	15 (0 to 25)	15 (0 to 25)	−0.5 (−3.0 to 1.9)	19 (0 to 25)	17.5 (0 to 25)	0.6 (−1.7 to 2.9)
ICU-free days, median (IQR)	10 (0 to 23)	10 (0 to 22)	0.1 (−2.2 to 2.4)	16 (0 to 22)	13.5 (0 to 22)	0.8 (−1.3 to 2.9)
Died before hospital discharge, *n* (%)	72/180 (40.0)	72/180 (40.0)	1.00 (0.8 to 1.3)	71/199 (35.7)	72/202 (35.6)	1.00 (0.8 to 1.3)

Definition of abbreviations: CI = confidence interval; ICU = intensive care unit; IQR = interquartile range; PreVent = Preventing Hypoxemia with Manual Ventilation during Endotracheal Intubation; RR = risk ratio.

propensity score and coarsened exact matching. This was dependent on the availability of several important confounders (e.g., APACHE II score, reason for intubation, oxygen saturation at induction, and BMI) in our dataset (11, 20). In general, clinical data are better suited to target trial analysis than administrative claims or other datasets with less complete capture of relevant confounders (24).

Other groups have used observational data to explicitly emulate randomized experiments. Hernán and colleagues (8) analyzed data from the Nurses' Health Study to emulate an intention-to-treat analysis of an idealized randomized trial of postmenopausal hormone therapy on coronary heart disease. By emulating design features of an idealized RCT, the authors show that differences in results between observational and randomized studies of hormone replacement therapy may be due to discrepancies in the included patient population and other design factors, and not to unmeasured confounding. A criticism of prior target trial emulations, however, is that they are conducted with knowledge of the resultant randomized experiment. By emulating PreVent *before* its results were known to the analysts, we eliminated the risk that prior knowledge of study results would influence our analytical approach even inadvertently.

Target trial emulation also draws attention to the timing of eligibility and treatment assignment, two standard design features of randomized experiments that can lead to bias when mismatched in observational analyses (12). García-Albéniz and colleagues (9) applied target trial emulation to evaluate the effects of emulated "randomization" to screening colonoscopy on subsequent colorectal cancer risk. By determining both eligibility and assigning treatment simultaneously at the start of a study, and by analyzing patients who "break protocol" with their initially assigned group, the authors show that they are able to arrive at more credible effect estimates. We similarly assigned treatment at the "start" of induction, using imputation to ascertain intervention status when this was not known, before subsequently emulating the randomization process using coarsened exact matching.

Our study does have several limitations. A major limitation is that criteria for

Table 4. Difference-in-differences estimates comparing the observational study and PreVent

Outcome	Treatment Effect		Difference-in-Differences (95% CI)
	Observational Study	PreVent Trial	
Primary outcome			
Lowest oxygen saturation, %	1.8	3.9	−2.1 (−5.9 to 1.7)
Secondary outcome			
Lowest oxygen saturation <80%	−9.4	−12.0	2.5 (−8.1 to 13.6)
Exploratory oxygen saturation outcomes			
Lowest oxygen saturation <90%	−3.9	−10.6	6.7 (−7.3 to 20.4)
Lowest oxygen saturation <70%	−3.9	−6.0	2.1 (−5.4 to 10.1)
Decrease in oxygen saturation, %	−1.9	−4.5	2.6 (−0.8 to 6.2)
Exploratory safety outcomes			
Operator-reported aspiration	1.1	−1.4	2.6 (−1.5 to 6.8)
Cardiac arrest 1 h after intubation	1.1	−0.5	1.6 (−2.0 to 5.3)
Exploratory clinical outcomes			
Ventilator-free days, median	−0.5	0.6	−1.1 (−4.5 to 2.2)
ICU-free days, median	0.1	0.8	−0.7 (−3.7 to 2.4)
Died before hospital discharge	0.0	0.0	0.0 (−13.7 to 13.8)

Definition of abbreviations: CI = confidence interval; ICU = intensive care unit; PreVent = Preventing Hypoxemia with Manual Ventilation during Endotracheal Intubation.

formally assessing the agreement between the observational analysis and the PreVent trial were not prespecified as part of our analytical plan. Authoritative guidance for comparing target trial emulations with realized target trials does not yet exist; we suggest using a prespecified minimal clinically important difference in outcomes and testing agreement with prespecified statistical tests, such as a DIDs estimator (25). In light of this limitation, however, the findings of this analysis should be considered exploratory. Second, we included positive-pressure ventilation delivered via noninvasive ventilation in our treatment group, though the PreVent trial required that positive pressure be delivered via bag-mask device for all patients randomized to positive-pressure ventilation. Although this decision was made with the intention of approximating the causal question in PreVent without excluding the large group of patients receiving noninvasive ventilation in the prior trials we analyzed, there could be differences in positive-pressure ventilation delivered by noninvasive ventilation versus bag-mask device that render this a sufficiently different intervention. Third, although we used both content knowledge and published evidence to generate a causal model and controlled for all relevant confounders in that model, this approach does not eliminate the possibility that unmeasured confounding may have influenced our findings. Finally, our observational analysis could not exclude there being no effect of positive-pressure ventilation on the primary outcome, median lowest oxygen saturation, although the direction and magnitude of our result was similar to findings from PreVent. Potential reasons for this may be our more restrictive inclusion criteria (including only patients with oxygen saturations of ⩾90%), effects of the other interventions (e.g., apneic oxygenation) tested in the trials that comprised our observational cohort, or our smaller effective sample size after matching. Nevertheless, our findings yielded a similar interpretation, that positive-pressure ventilation applied starting at induction reduced rates of severe hypoxemia.

Despite these limitations, our findings suggest several potential applications for target trial emulations. When RCTs are not available or are impractical, rigorously conducted target trial emulation using rich clinical data may yield stronger evidence to support clinical decisions than other types of observational research (9, 12). In the case where an RCT is planned, target trial emulation may inform study design, including sample size calculations, outcome selection, and selection of inclusion/exclusion criteria. Specifically, when selecting outcomes for clinical trials, investigators often attempt to balance a number of factors, including relevance to patients, providers, and other stakeholders, measurement characteristics, and statistical efficiency (26). This latter consideration typically favors the selection of continuous outcomes.

With regard to PreVent, target trial emulation suggested that the intervention may have had a greater effect on the outcome of severe hypoxemia than on the outcome of lowest oxygen saturation. This finding might have led investigators to consider severe hypoxemia as a dichotomous primary outcome. Based on the observational analysis, the relatively high baseline of severe hypoxemia in the control group coupled with the large difference in outcome rates between the intervention and control groups would have yielded an estimated sample size of 359 patients per group to achieve 90% power for detecting a similar difference at the α level of 0.05 (27). Though this is greater than the sample size ultimately used for the continuous outcome, and would have led to a lengthier trial, it may have prompted investigators to favor severe hypoxemia given its clinical relevance. Finally, clinical trialists interested in evaluating several closely related interventions may consider building in prespecified observational studies into prospective randomized trials. These methods might generate stronger evidence than would be possible with other types of observational data, with a significant gain

in efficiency over the cost or time involved in conducting separate clinical trials. For example, our observational analysis was designed and conducted in 11 days, whereas the PreVent trial itself required more than 2 years.

In summary, we used preexisting clinical data to evaluate a novel intervention and arrived at results comparable to those of a contemporaneous randomized experiment. This suggests that target trial emulation using modern causal methods and rich clinical trial data can provide informative results for comparative effectiveness research. These results also support the routine use of positive-pressure ventilation during tracheal intubation of critically ill adults to prevent severe hypoxemia. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

References

1 Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med* 2000;342:1887–1892.
2 Vincent J-L. We should abandon randomized controlled trials in the intensive care unit. *Crit Care Med* 2010;38(10 suppl):S534–S538.
3 Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *N Engl J Med* 2000;342:1878–1886.
4 Shikata S, Nakayama T, Noguchi Y, Taji Y, Yamagishi H. Comparison of effects in randomized controlled trials with observational studies in digestive surgery. *Ann Surg* 2006;244:668–676.
5 Hemkens LG, Contopoulos-Ioannidis DG, Ioannidis JPA. Agreement of treatment effects for mortality from routinely collected data and subsequent randomized trials: meta-epidemiological survey. *BMJ* 2016;352:i493.
6 Moodie EEM, Stephens DA. Using directed acyclic graphs to detect limitations of traditional regression in longitudinal studies. *Int J Public Health* 2010;55:701–703.
7 Pearl J. An introduction to causal inference. *Int J Biostat* 2010;6:7.
8 Hernán MA, Alonso A, Logan R, Grodstein F, Michels KB, Willett WC, *et al*. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiology* 2008;19:766–779.
9 García-Albéniz X, Hsu J, Hernán MA. The value of explicitly emulating a target trial when using real world evidence: an application to colorectal cancer screening. *Eur J Epidemiol* 2017;32:495–500.
10 Labrecque JA, Swanson SA. Target trial emulation: teaching epidemiology and beyond. *Eur J Epidemiol* 2017;32:473–475.
11 Dahabreh IJ, Sheldrick RC, Paulus JK, Chung M, Varvarigou V, Jafri H, *et al*. Do observational studies using propensity score methods agree with randomized trials? A systematic comparison of studies on acute coronary syndromes. *Eur Heart J* 2012;33:1893–1901.
12 Hernán MA, Sauer BC, Hernández-Díaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol* 2016;79:70–75.
13 Walkey AJ. Expanding the dimensions of effectiveness research in sepsis. *Am J Respir Crit Care Med* 2014;190:970–971.
14 El-Orbany M, Connolly LA. Rapid sequence induction and intubation: current controversy. *Anesth Analg* 2010;110:1318–1325.
15 Casey JD, Janz DR, Russell DW, Vonderhaar DJ, Joffe AM, Dischert KM, *et al*. Bag-mask ventilation during tracheal intubation of critically ill adults. *N Engl J Med* 2019;380:811–821.
16 Admon AJ, Donnely JP, Iwashyna, TJ. Emulating a novel target trial using previously collected clinical trial data: an application to positive pressure ventilation during tracheal intubation. Charlottesville, VA: The Center for Open Science; 18 February 2019 [accessed 2019 Feb 18]. Available from: http://osf.io/p8cz3.
17 Casey JD, Janz DR, Russell DW, Vonderhaar DJ, Joffe AM, Dischert KM, *et al*.; PreVent Investigators and the Pragmatic Critical Care Research Group. Manual ventilation to prevent hypoxaemia during endotracheal intubation of critically ill adults: protocol and statistical analysis plan for a multicentre randomised trial. *BMJ Open* 2018;8:e022139.
18 Semler MW, Janz DR, Russell DW, Casey JD, Lentz RJ, Zouk AN, *et al*.; Check-UP Investigators(*); Pragmatic Critical Care Research Group. A multicenter, randomized trial of ramped position vs sniffing position during endotracheal intubation of critically ill adults. *Chest* 2017;152:712–722.
19 Semler MW, Janz DR, Lentz RJ, Matthews DT, Norman BC, Assad TR, *et al*.; FELLOW Investigators; Pragmatic Critical Care Research Group. Randomized trial of apneic oxygenation during endotracheal intubation of the critically ill. *Am J Respir Crit Care Med* 2016;193:273–280.
20 Garrido MM, Kelley AS, Paris J, Roza K, Meier DE, Morrison RS, *et al*. Methods for constructing and assessing propensity scores. *Health Serv Res* 2014;49:1701–1720.
21 McKown AC, Casey JD, Russell DW, Joffe AM, Janz DR, Rice TW, *et al*. Risk factors for and prediction of hypoxemia during tracheal intubation of critically ill adults. *Ann Am Thorac Soc* 2018;15:1320–1327.
22 Sjoding MW, Luo K, Miller MA, Iwashyna TJ. When do confounding by indication and inadequate risk adjustment bias critical care studies? A simulation study. *Crit Care* 2015;19:195.
23 Wunsch H, Linde-Zwirble WT, Angus DC. Methods to adjust for bias and confounding in critical care health services research involving observational data. *J Crit Care* 2006;21:1–7.
24 Poses RM, Smith WR, McClish DK, Anthony M. Controlling for confounding by indication for treatment. Are administrative data equivalent to clinical data? *Med Care* 1995;33(4 suppl):AS36–AS46.
25 Chan LS. Minimal clinically important difference (MCID)—adding meaning to statistical inference. *Am J Public Health* 2013;103:e24–e25.
26 Iwashyna TJ, McPeake J. Choosing outcomes for clinical trials: a pragmatic perspective. *Curr Opin Crit Care* 2018;24:428–433.
27 Schulz KF, Grimes DA. Sample size calculations in randomised trials: mandatory and mystical. *Lancet* 2005;365:1348–1353.

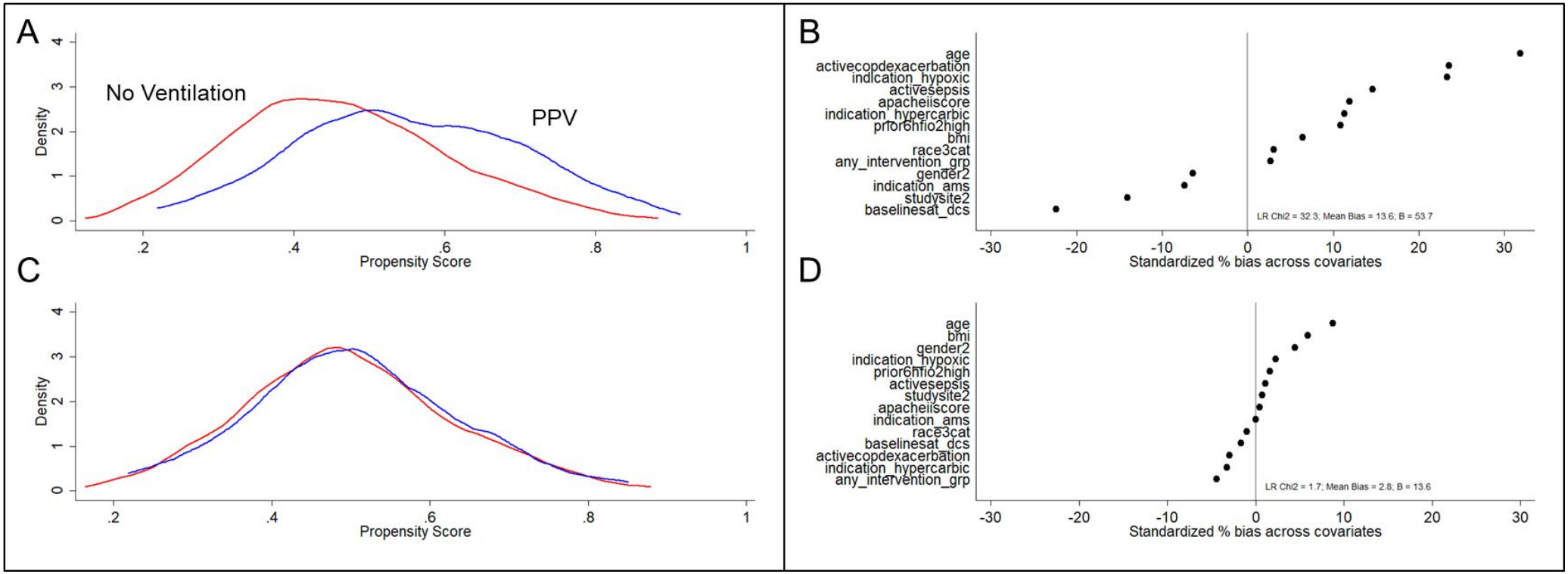
Online Supplement

Emulating a Novel Clinical Trial using Existing Observational Data: Predicting Results of the PreVent Study

Andrew J. Admon, MD, MPH, John P. Donnelly, PhD, Jonathan D. Casey, MD, David. R. Janz, MD, MSc, Derek W. Russell, MD, Aaron M. Joffe, DO, Derek J. Vonderhaar, MD, Kevin M. Dischert, MD, Susan B. Stempek, MMSc, PA, James M. Dargin, MD, Todd W. Rice, MD, MSc, Theodore J. Iwashyna, MD, PhD, and Matthew W. Semler, MD, MSc for the Pragmatic Critical Care Research Group.

Table of Contents

Supplementary Figure 1 - Propensity score distributions and balance of covariates between matched groups. (A) Densities of propensity score by treatment and control group before matching. (B) Standardized percent bias before matching. (C) Densities of propensity score by treatment and control group after matching. (D) Standardized percent bias after matching.



List of Investigators for the Pragmatic Critical Care Research Group

Vanderbilt University Medical Center, Nashville, TN – Ryan M. Brown, Jonathan D. Casey*, Todd W. Rice*, Wesley H. Self, Matthew W. Semler*; Lahey Hospital & Medical Center, Burlington, MA – James Dargin*, Susan Stempek*, Joanne Wozniak; Louisiana State University School of Medicine, New Orleans, LA – Bennett P. deBoisblanc, David R. Janz*, Yasin A. Khan; Ochsner Health System, New Orleans, LA – Kevin M. Dischert*, Derek J. Vonderhaar*; University of Alabama, Birmingham, AL – Swati Gulati, Derek W. Russell*, William S. Stigler, Aline N. Zouk; University of Washington, Seattle, WA – Itay Bentov, Aaron M. Joffe*

*Denotes members of the Writing Committee.

Correspondences with Dr. David Lederer (Honest Broker)

Initial submission from analysts to honest broker:

From: Admon, Andrew
Date and Time: February 17, 2019 11:58 PM, ET
Subject: pre-registering a prospective Target Trial

Hi Dr. Lederer,

Thank you again for your willingness to serve as an 'honest broker'. In the attached document is a short introduction, our complete analysis plan, and our results with a brief interpretation. We've also included all of our statistical code in the appendix. We're looking forward to seeing the results of the actual target trial (PreVent) tomorrow!

Regards,
Andy Admon

—
Andrew J. Admon, MD, MPH
Fellow, Pulmonary and Critical Care Medicine
Department of Internal Medicine
University of Michigan

Response from honest broker:

From: Lederer, David
Date and Time: February 18, 2019 6:20 AM, ET
Subject: pre-registering a prospective Target Trial

Hi Andy,

I am confirming that I received your report including the complete analysis plan and results. Received February 18, 2019. 6:20am Eastern Time.

Best
Dave

Pre-Registered Description of Analytical Plan and Results (Registered 2/18/2019)

Introduction

Well-conducted randomized, controlled trials (RCTs) represent the strongest level of evidentiary support for or against medical interventions. Unfortunately, RCTs are costly, lengthy, and, for some interventions, impractical, leaving us limited in our ability to apply this gold standard to a number of important clinical questions. Target trial emulation using observational data may fill this important gap, though confounding by indication, whereby patients selectively receive the intervention due to factors also associated with the study's outcome, may lead to misleading results.(1–4)

Modern causal methods applied to rich, clinical trial data may reduce this important source of confounding. Specifically, clinical trial data often include detailed, accurately collected information obtained with the intent of informing trial interpretation. In the context of critical care trials, this often includes several patient-level (i.e., detailed physiologic data), provider-level (i.e., experience and specialty), and unit-level (i.e., medical versus surgical) variables not often available in other datasets. By capitalizing on these data, investigators may extend the impact of a clinical trial by accurately emulating related target trials to answer additional clinical questions.

In this study, we use existing clinical trial data to estimate the effects of positive pressure ventilation (PPV) added to conventional rapid sequence intubation on procedural hypoxia. To do so, we analyze data collected from three prior randomized trials to emulate the recently completed Preventing Hypoxemia with Manual Ventilation during Endotracheal Intubation (PreVent) trial. The PreVent study protocol was provided to the analysts (AJA and JPD) by trial investigators (JDC and MWS) for the purpose of target trial emulation before the results of the trial were published. The analysts subsequently registered an analytical plan, statistical code,

and results without knowledge of the target trial's findings in order to evaluate whether causal methods could be applied to the observational data to 1) emulate the randomization procedure of the target trial and 2) arrive at similar conclusions.

Methods

Description of the Target Trial

The target trial is the PreVent trial, a multi-center, parallel-group, unblinded, pragmatic randomized trial comparing bag-mask ventilation to no ventilation between induction and laryngoscopy during endotracheal intubation among critically ill adults (Citation Pending). Complete details about the target trial are reported elsewhere.

Data Source

Data were obtained from three randomized trials evaluating interventions related to endotracheal intubation in similar patient populations and study settings to the PreVent trial. The Checklists and Upright Positioning in endotracheal intubation of critically ill patients (Check-UP) study was a randomized, multicenter, pragmatic two-by-two factorial trial comparing 1) the ramped position with the sniffing position and 2) the use of an a written pre-intubation checklist during endotracheal intubation of critically ill adults.(5) The Facilitating Endotracheal intubation by Laryngoscopy technique and apneic Oxygenation Within the intensive care unit (FELLOW) study was a randomized, open-label, parallel group, pragmatic two-by-two factorial trial comparing apneic oxygenation with usual care and direct laryngoscopy with video laryngoscopy among critically ill adults.(6) The PrePARE trial compared administration of a fluid bolus prior to induction to no administration of a fluid bolus with regard to the incidence of cardiovascular collapse during tracheal intubation of critically ill adults. Although patients could be co-enrolled in the PrePARE and PreVent trials, the only data from the PrePARE trial used in the current analysis was data from centers not actively enrolling in the PreVent trial. The cohort derived

from this dataset is referred to as the observational cohort because it was observational with respect to the primary intervention, which was not allocated by randomization.

Eligibility Criteria

We included adult (aged >= 18) patients enrolled in Check-UP, FELLOW, or PrePARE. (**Table 1**). Patients enrolled in PreVent were not included in this analysis. We excluded pregnant or incarcerated patients, patients for whom the need for tracheal intubation was deemed too urgent for randomization, or cases where bag-valve mask ventilation was deemed necessary or contraindicated (known and true for 43 patients). For purposes of the observational study, we also excluded patients with a baseline oxygen saturation below 90% by pulse oximetry. This was done in order to more accurately separate use of PPV for rescue (done in those with oxygen saturations below 90% in PreVent) from those treated with *preventive* PPV, at the risk of biasing our observational results towards the null (see: *Assignment Procedures,* below).

Treatment Strategies

The PreVent trial compared modified RSI (defined as RSI with positive pressure ventilation via bag-valve mask between induction and laryngoscopy) to standard RSI (defined as RSI without positive pressure between induction and laryngoscopy). To limit imprecision in the exposure of interest (positive pressure ventilation) stemming from differences in pressure settings, all PPV in the PreVent trial was delivered using non-invasive ventilation. Because the decision to deliver PPV during intubation in the observational cohort could have led to either bag valve mask or non-invasive ventilation, we included both of these modalities in our treated group. Patients receiving neither modality of PPV were included in the untreated, or control group.

Assignment Procedures

We sought to emulate the PreVent trial's randomization to PPV at induction using data from our observational cohort.

First, although the exposure of interest was PPV beginning at induction, Check-UP and FELLOW captured only use of positive pressure ventilation at any time from induction to intubation. This included patients who received PPV at induction (e.g., the intervention arm in PreVent) and those who received PPV for rescue during prolonged intubations, in between successive intubation attempts, or when oxygen saturation dropped below 90% (either the intervention or control arms in PreVent). As a result, although we knew intubation status for subjects who never received PPV between induction and intubation, we followed a series of rules to recreate treatment status for patients who received PPV at an uncertain point during intubation. First, anyone with a baseline oxygen saturation of <90% was excluded, as these patients would have qualified for PPV regardless of randomization arm in PreVent. Second, subjects who never received PPV between induction and intubation were assigned to the control group. Third, those subjects without a prolonged intubation (defined as either > 1 attempt or in the top 10% of procedure durations) who received PPV were assumed to have received this at or close to the time of induction, and so were assigned to the treatment arm. Finally, intubation status among the 371 patients for whom exposure status was known more confidently was used to predict exposure status at intubation among the remaining 91 patients using a logistic model. This model had good discrimination (c-statistic: 0.80) and fit (pseudo R^2: 0.21) and resulted in reclassifying 27 people from the treatment to control groups.

Next, to mimic the randomization performed in the PreVent trial, we generated propensity scores using study site, intervention group, age, BMI, race, gender, presence of sepsis, presence of chronic obstructive pulmonary disease exacerbation, indication for intubation (hypoxia or hypercarbia), highest fraction of inspired oxygen in the prior six hours, and baseline

oxygen saturation. These variables were selected because they were thought to influence both an operator's decision to apply positive pressure ventilation and a subject's lowest oxygen saturation (i.e., they are possible confounders).(7) We matched patients in a 1:1 ratio using coarsened exact matching and verified that covariates were balanced across matched groups. Of 462 patients in the analytical sample, 360 patients were matched (180 in both the treatment and control groups).

Follow-up Period

Follow-up periods were similar between this observational study and the PreVent trial. Specifically, oxygen saturations were measured from induction through two minutes after tracheal intubation. Secondary and safety outcomes were measured for up to 24 hours after intubation.

Causal Contrasts of Interest

The primary outcome in both this observational study and in PreVent was nadir oxygen saturation. This was the primary outcome in all three trials (Check-UP, FELLOW, and PrePARE) comprising our analytical cohort. Secondary outcomes included proportion of patients with severe hypoxemia (below 80%),

Analysis Plan

We first evaluated our propensity matching procedure by evaluating balance of baseline patient characteristics ("Table 1" variables from PreVent) across our treatment and control groups using two sample t-tests. After assessing balance of these covariates across groups, we next sought to apply the statistical analysis plan used in the PreVent trial in analyzing primary and all secondary/safety outcomes. As a result, we used the Mann-Whitney U test to compare distributions of nadir oxygen saturation between our treatment and control groups and a series

of regression analyses to estimate mean differences or risk ratios for each outcome. Because propensity score matching was employed to adjust for confounding, linear models used for effect estimation included only one independent variable (treatment assignment). All tests were two-tailed. Analyses were conducted using Stata 14.2 (StataCorp LLC, College Station, Texas).

Brief Interpretation of Results

Using propensity score matching, we were able to arrive at two balanced groups for comparison that differed in their use of PPV at induction (**Figure 1, Table 2**). The median oxygen saturation, our primary outcome, was slightly higher in the intervention (PPV) group, though we could not exclude there being no difference. (**Figure 2**). There were fewer cases of severe hypoxia, defined as an arterial oxygen saturation of <80% by pulse oximeter, in the intervention group with no substantial differences in either of the safety outcomes (**Table 3**).

References

1. Kyriacou DN, Lewis RJ. Confounding by Indication in Clinical Research. *JAMA* 2016;316:1818.
2. Hernán MA, Sauer BC, Hernández-Díaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol* 2016;doi:10.1016/j.jclinepi.2016.04.014.
3. Cain LE, Saag MS, Petersen M, May MT, Ingle SM, Logan R, Robins JM, Abgrall S, Shepherd BE, Deeks SG, Gill MJ, Touloumi G, Vourli G, Dabis F, Vandenhende MA, Reiss P, van Sighem A, Samji H, Hogg RS, Rybniker J, Sabin CA, Jose S, del Amo J, Moreno S, Rodríguez B, Cozzi-Lepri A, Boswell SL, Stephan C, Pérez-Hoyos S, *et al.* Using observational data to emulate a randomized trial of dynamic treatmentswitching strategies: An application to antiretroviral therapy. *Int J Epidemiol* 2016;doi:10.1093/ije/dyv295.
4. Hernán MA, Alonso A, Logan R, Grodstein F, Michels KB, Willett WC, Manson JE, Robins JM. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiology* 2008;19:766–79.
5. Janz DR, Semler MW, Joffe AM, Casey JD, Lentz RJ, deBoisblanc BP, Khan YA, Santanilla JI, Bentov I, Rice TW, Check-UP Investigators* GP, Pragmatic Critical Care Research Group JL, Vonderhaar DJ, Lapinel NC, Samant SD, Paccione R, Dischert K, Majid-Moosa A, Crespo J, Fashho MB, Matthews DT, Berg JZ, Assad TR, McKown AC, Huerta LE, Kocurek EG, Halliday SJ, Kerchberger VE, Merrick CM, *et al.* A Multicenter Randomized Trial of a Checklist for Endotracheal Intubation of Critically Ill Adults. *Chest* 2018;153:816–824.
6. Semler MW, Janz DR, Lentz RJ, Matthews DT, Norman BC, Assad TR, Keriwala RD, Ferrell BA, Noto MJ, McKown AC, Kocurek EG, Warren MA, Huerta LE, Rice TW,

FELLOW Investigators, Pragmatic Critical Care Research Group. Randomized Trial of Apneic Oxygenation during Endotracheal Intubation of the Critically Ill. *Am J Respir Crit Care Med* 2016;193:273–80.

7. McKown AC, Casey JD, Russell DW, Joffe AM, Janz DR, Rice TW, Semler MW. Risk factors for and prediction of hypoxemia during tracheal intubation of Critically Ill Adults. *Ann Am Thorac Soc* 2018;15:1320–1327.

Pre-Registered Comparison of Target Trial and Observational Study

Characteristic	Target Trial	Observational Study

Eligibility Criteria	• Included
	  ○ Admitted in a unit participating in PreVent
	  ○ Intubation planned
	  ○ 18 years old or older
	  ○ Qualified operator
	• Excluded
	  ○ Pregnant
	  ○ Incarcerated
	  ○ Need for tracheal intubation too emergent
	  ○ Clinician deemed patient to need BVM (hypoxemia, severe acidemia, respiratory arrest)
	  ○ BVM Contraindicated or intubation too urgent
		• Included
		  ○ Admitted in a unit participating in Check-UP, FELLOW, or PreVent
		  ○ Intubation planned
		  ○ 18 years old or older
		  ○ Any operator
		• Excluded
		  ○ Pregnant
		  ○ Incarcerated
		  ○ Need for tracheal intubation too emergent
		  ○ Clinician deemed patient to need BVM (hypoxemia, severe acidemia, respiratory arrest)
		  ○ BVM contraindicated or intubation too urgent, when known
		  ○ SaO2 at baseline < 90% [49 people]

Treatment Strategies	• Modified RSI [PPV via BVM at induction] • Classic RSI [control] with BVM Rescue (if SaO2 < 90%) and in between attempts	• Modified RSI [PPV via BVM or NIV at induction] • Classic RSI [control] with BVM Rescue (if SaO2 < 90%) and in between attempts

Assignment Procedures	Randomization 1:1 ratio to BVM using permuted blocks of two, four, and six stratified by study site.	1. Exposure known for 125 people
		2. For others, exposure imputed based on the following rules:
		  a. BVM or BiPAP during first attempt without prolonged intubation -> assigned intervention group
		  b. No BVM or BiPAP during first attempt without prolonged intubation -> assigned control group
		  c. For those with prolonged intubation (Top decile of intubation length or >1 attempt) assigned using a predictive model based on known recipients of exposure

Follow-up Period	During RSI to two minutes after tube placement	During RSI to two minutes after tube placement
Causal Contrasts of Interest	Lowest oxygen saturation between induction and two minutes after tracheal intubation Other outcomes: • Severe hypoxemia (SaO2 < 80%)	Lowest oxygen saturation between induction and two minutes after tracheal intubation Other outcomes: • Severe hypoxemia (SaO2 < 80%)

	• Lowest SaO2, highest FiO2, and highest PEEP between 6-24 hours • Operator reported aspiration • New opacity within 48 hours	• Lowest SaO2, highest FiO2, and highest PEEP between 6-24 hours • Operator reported aspiration
Analysis Plan	Primary Analysis: Mann-Whitney U Test	Primary Analysis: Mann-Whitney U Test on propensity score matched population

Pre-Registered Results

Table 2: Patient Characteristics at Baseline – Observational Cohort

Patient Characteristic	**Positive Pressure Ventilation (n=180)**	**No Ventilation (n=180)**	**p-value**
Age, median [IQR], years	60 (49-68)	56 (48-68)	0.41
Male sex, No. (%)	106 (58.9)	102 (56.7)	0.67
White race, No. (%)	140 (77.8)	140 (77.8)	0.97
Body mass index, median [IQR], kg/m^2	28.1 (23.4-33.4)	27.4 (23.5-32.2)	0.57
APACHE II score, median [IQR]	20.5 (16-26)	21 (16-25)	0.96
Vasopressors, No. (%)	27 (15.0)	42 (23.3)	0.05
Active medical conditions, No. (%)			
Sepsis or septic shock	98 (54.4)	97 (53.9)	0.92
Gastrointestinal bleeding	21 (11.7)	33 (18.3)	0.08
Indications for intubation, No. (%)			
Hypoxemic respiratory failure	103 (57.2)	101 (56.1)	0.83
Hypercarbic respiratory failure	23 (12.8)	25 (13.9)	0.76
Airway protection for decreased level of consciousness	60 (33.3)	60 (33.3)	1.0
BiPAP in prior 6 hours, No. (%)	71 (39.4)	37 (20.6)	<0.001
Highest FiO_2 in prior 6 hours, median [IQR]	0.4 (0.3-0.7)	0.4 (0.3-0.8)	0.88
Lowest oxygen saturation in prior 6 hours, median [IQR], %	91 (88-94)	92 (89-95)	0.27

Table 3: Outcomes after Tracheal Intubation – Observational Cohort

Outcomes	Positive Pressure Ventilation (n=180)	No Ventilation (n=180)	Relative Risk or Mean Difference (95% Confidence Intervals)
Primary Outcome			
Lowest oxygen saturation, median [IQR], %	94 (86-98)	93 (81.5-99)	1.81 (-0.97, 4.59)
Secondary Outcome			
Lowest oxygen saturation < 80%, No. (%)	25 (13.9)	42 (23.3)	0.60 (0.38-0.93)
Exploratory Oxygen Saturation Outcomes			
Lowest oxygen saturation < 90%, No. (%)	59 (32.8)	66 (36.7)	0.89 (0.67-1.19)
Lowest oxygen saturation < 70%, No. (%)	12 (6.7)	19 (10.6)	0.63 (0.32-1.26)
Decrease in oxygen saturation, median [IQR], %	3 (0-11)	3 (0-13)	-1.86 (-4.50, 0.78)
Exploratory Safety Outcomes			
Operator-reported aspiration, No. (%)	3 (1.7)	1 (0.6)	3.00 (0.32-28.57)
Cardiac arrest within 1 hour after intubation, No. (%)	4 (2.2)	2 (1.1)	2.00 (0.37-10.78)
Exploratory Clinical Outcomes			
Ventilator-free days, median [IQR]	15 (0-25)	15 (0-25)	-0.53 (-3.01, 1.95)
ICU-free days, median [IQR]	10 (0-23)	10 (0-22)	0.12 (-2.17, 2.40)
Died before hospital discharge, No. (%)	72 (40.0)	72 (40.0)	1.00 (0.78-1.29)

Figure 1 – CONSORT Diagram

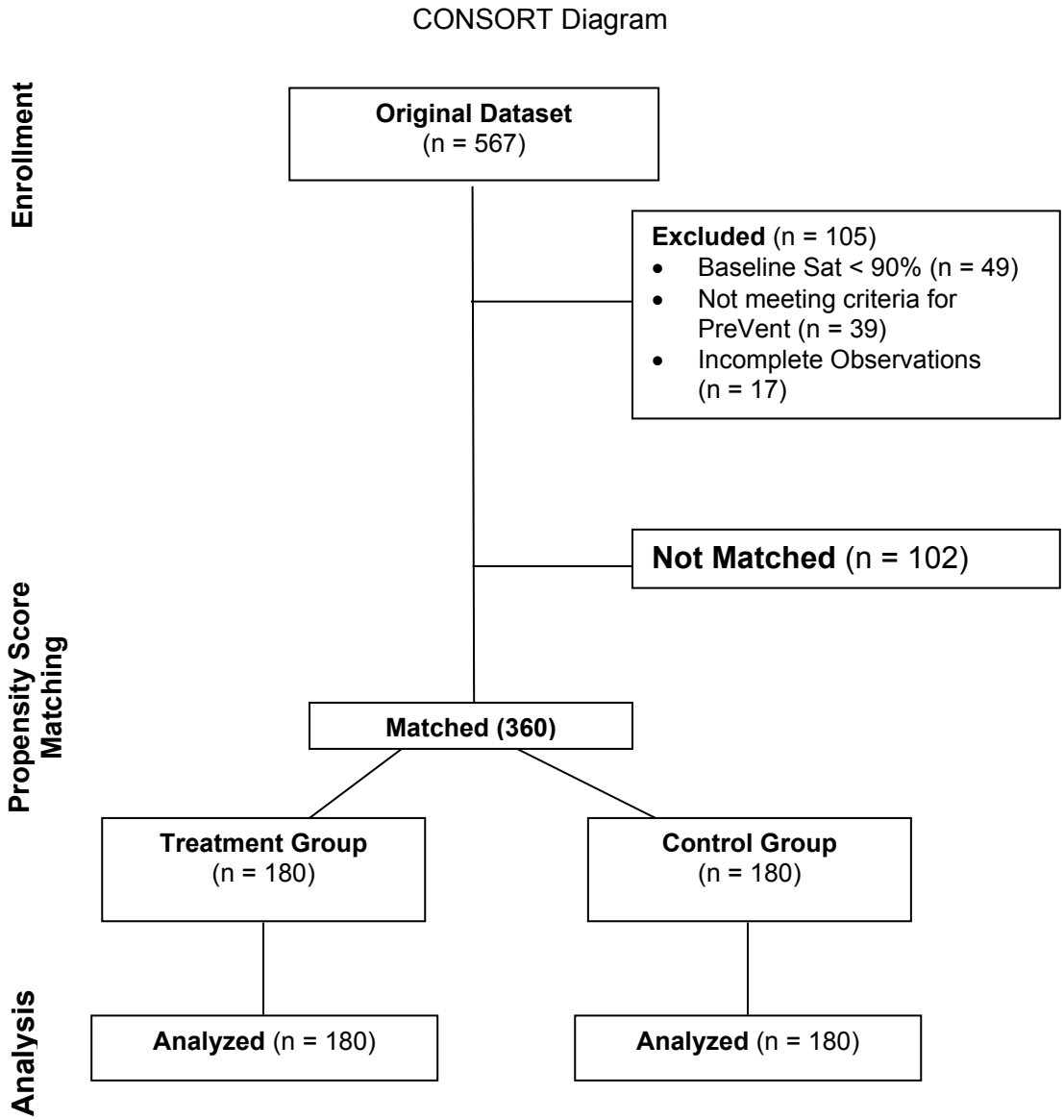
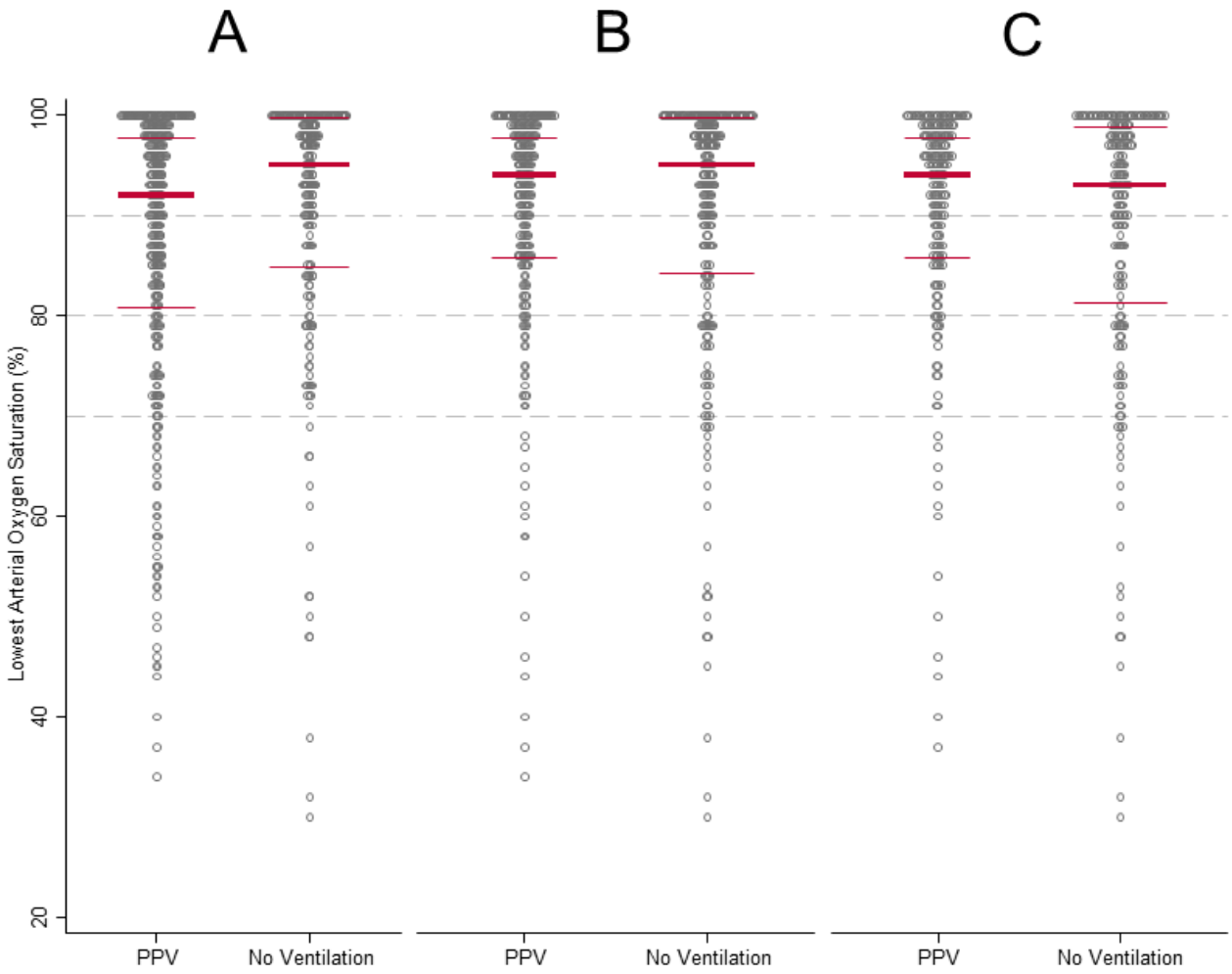


Figure 2 – Lowest arterial oxygen saturation in treatment versus control groups.



A reflects unadjusted results. B reflects results after treatment assignment procedure. C reflects results after propensity score matching.

Pre-Registered Statistical Code

```
/*************************************************************/
/*************************************************************/
/*************************************************************/
/*TARGET TRIAL EMULATION*/
/*2/17/19*/
/*VERSION 2.0 (Pre-Trial Release)*/
/*ADMON AND DONNELLY*/
/*************************************************************/
/*************************************************************/
/*************************************************************/

cd "C:\Users\joond\Desktop\UMich\Target Trial Emulation\"
set more off

version 14.2

set seed 43452

/*IMPORT*/
import delimited using Database_without_PreVent_Lincoln_ED_for_UM.csv

/*MAKE EXCLUSION BASED ON SAT AND PREVENT STATUS*/
gen baselinesat90 = .
replace baselinesat90=0 if baselinesat_dcs<90
replace baselinesat90=1 if baselinesat_dcs>=90 & baselinesat_dcs<.

gen exclude = 0
replace exclude=1 if baselinesat90==0 | prevent_nonenroll!="Center not participating in trial"

keep if exclude==0

/*GENERATE STARTING TREATMENT VARIABLE*/
/*OVERRIDE WITH BMV_INDUCTION FOR PATIENTS WHO WE KNOW DID NOT RECEIVE TREATMENT STARTING AT
INDUCTION*/

gen bvmbipap_yesno =.
replace bvmbipap_yesno=0 if ventilation_induct_laryng_none==1
replace bvmbipap_yesno=1 if ventilation_induct_laryng_none==0
replace bvmbipap_yesno=0 if bmv_induction==0

/*TAG POPULATION TO PREDICT EXPOSURE*/
/*DERIVATION POPULATION: ONE ATTEMPT WITH SHORT DURATION OR KNOWN TREATMENT OR KNOWN NO
TREATMENT*/
gen sample_tag=0
replace sample_tag=1 if (attempts_dcs==1 & duration_total_seconds<=215) | bmv_induction!=. |
bvmbipap_yesno==0

/*GENERATE OUTCOME VARIABLES*/
gen hypox80 = .
replace hypox80=0 if nadirsat_dcs>=80 & nadirsat_dcs<=.
replace hypox80=1 if nadirsat_dcs<80

gen hypox90 = .
replace hypox90=0 if nadirsat_dcs>=90 & nadirsat_dcs<=.
replace hypox90=1 if nadirsat_dcs<90

gen hypox70 = .
replace hypox70=0 if nadirsat_dcs>=70 & nadirsat_dcs<=.
replace hypox70=1 if nadirsat_dcs<70

gen sat_decrease = baselinesat_dcs - nadirsat_dcs

*aspiration_new_dcs

gen carrest1hrintub = 0
replace carrest1hrintub = 1 if cardiacarrest1hr==1 | dcscomplicationscardiacarrest==1

*vfds
*icufds
```

```
*died_in_hospital

/*CHECK VARIABLE AVAILABILITY ACROSS TRIALS*/
set more off
codebook indication_hypoxic indication_hypercarbic indication_hypercarb_hypox indication_ams
indication_airwaycompromise indication_resparrest activesepsis activesepticshock
activecardiogenicshock activedistributiveshock activeneurogenicshock activehemorrhagicshock
activehypovolemicshock activegibleeding uppergibleed lowergibleed activehepaticencephalopathy
activedelirium activecopdexacerbation activeasthmaexacerbation activecfexacerbation
activeildexacerbation activestemi activenstemi activeunstableangi activedrugoverdose
activepancreatitis campriorintubaiton lowmapsixhoursprior pressorssixhoursprior pressors_6hrs_epi
pressors_6hrs_norepi pressors_6hrs_phenyl pressors_6hrs_dopa pressors_6hrs_dobuta
pressors_6hrs_milrinone prior6hbipap prior6hfio2high prior6hspo2low prior6habg reintubation_24h
reintubation_72h preoxnone preoxnrb preoxbipap preoxbmv preoxnc preoxhfnc nmbrocuronium
nmbvecuronium nmbsuccinylcholine nmbcisatracurium anynmb induction_ativan induction_dilaudid
induction_etomidate induction_fentanyl induction_ketamine induction_lidocaine induction_propofol
induction_versed tubesize laryngoscopetype laryngoscopesize laryngoscopeblade tubetapelevel
confirmauscultation confirmetco2 confirmbronch grade_view_emr difficultystarpanel bougie_emr
rescuedevice_emr nd_proceduralist_emr complications_none_emr fail_to_intubate_emr
complications_aspiration_emr complications_equip_fail_emr complications_bleeding_emr
complications_laryngospasm_emr da_emesis_emr da_aspiration_emr da_ugib_emr da_epistaxis_emr
da_airway_mass_emr da_hn_radiation_emr da_limitedneckmobility_emr da_limitedmouthopen_emr
cardiacarrest10min enroll_prepare prepare_nonenroll prevent_nonenroll aoassigned
deviceassigneddcs baselinesat_dcs duration_total_seconds nadirsat_dcs experience_dcs
fellowtrainmonths deviceexperience_dcs bmv_ever_dcs ventilation_induct_laryng_none ao_none_dcs
ao_nc bipap_induc_laryng_dcs device_dcs grade_view_dcs attempts_dcs bougie_dcs
addition_equip_vl_dcs addition_equip_dl_dcs additional_equp_lma_dcs additional_equip_bronch
secondproceduralist_dcs aspiration_new_dcs complications_sbp80_dcs dcscomplicationscardiacarrest
dcscomplicationsesophagealintuba dcscomplicationsairwaytrauma operatorme specialty
intubationmonth monthofstudy nightintubation diedever died_1hr died_icu died_in_hospital vfds
iculos icufds

/*RECODE VARIABLES TO NUMERIC*/
foreach var of varlist trialgroup studysite gender race aoassigned position_randomized {
encode `var', gen(`var'2)
}

/*GENERATE PREDICTION AND ADJUSTMENT VARIABLES*/

gen race3cat = 3
replace race3cat = 1 if race2==4
replace race3cat = 2 if race2==2

gen difficultystarpanel2 = .
replace difficultystarpanel2 = 1 if difficultystarpanel == "Difficult"
replace difficultystarpanel2 = 2 if difficultystarpanel == "Easy"
replace difficultystarpanel2 = 3 if difficultystarpanel == "Moderate"

gen intervention_group = 5
replace intervention_group=1 if aoassigned2==1
replace intervention_group=2 if aoassigned2==2
replace intervention_group=3 if position_randomized2==1
replace intervention_group=4 if position_randomized2==2

egen studysite_ivgroup = group(studysite intervention_group)
tab studysite_ivgroup bvmbipap_yesno if sample_tag==1
*COMBINE ICUS AT SINGLE SITE DUE TO SMALL NUMBERS
replace studysite2 = 5 if studysite2==6

gen any_intervention_grp = 0
replace any_intervention_grp=1 if position_randomized2==1 | aoassigned2==1 |
fluidbolus_preintub_dcs==1

gen preoxpp_group =0
replace preoxpp_group = 1 if preoxbipap==1 | preoxbmv==1
replace preoxpp_group = . if preoxbipap==. | preoxbmv==.

/*SAVE DATA SET WITHOUT NEW TREATMENT VARIABLE*/
save emulation_cohort_premodel.dta, replace

/*USE MODEL IN PATIENTS WITH KNOWN TREATMENT STATUS TO PREDICT PROBABILITIES OUT OF SAMPLE*/
```

```
logistic bvmbipap_yesno i.studysite2 any_intervention_grp ///
age bmi apacheiiscore i.race3cat i.gender2 ///
indication_hypoxic indication_hypercarbic indication_ams ///
activesepsis activegibleeding activecopdexacerbation ///
prior6hbipap prior6hfio2high pressorssixhoursprior ///
baselinesat_dcs preoxpp_group if sample_tag==1
lroc
predict p
keep if p!=.

keep studyid p nadirsat_dcs bvmbipap_yesno sample_tag studysite2 age bmi apacheiiscore race3cat
gender2 indication_hypoxic any_intervention_grp ///
indication_hypercarbic indication_ams activesepsis activegibleeding activecopdexacerbation ///
prior6hbipap prior6hfio2high prior6hspo2low baselinesat_dcs intervention_group preoxbmv
preoxbipap bmv_induction exclude hypox80 bmv_ever_dcs preoxpp_group pressorssixhoursprior ///
hypox90 hypox70 sat_decrease aspiration_new_dcs carrest1hrintub vfds icufds died_in_hospital

/*ALTERNATIVE STRATEGY BASED ON MULTIPLE IMPUTATION*/
/*NEED VALUES FOR THIS ANALYSIS SO NEEDED TO IDENTIFY MOST COMMON VALUE ACROSS IMPUTATIONS*/
/*RESULTED IN EXACT SAME TREATMENT ASSIGNMENT*/
/*FOR SUBSEQUENT ANALYSES WILL CONSIDER POOLING*/
/*
gen bvmbipap_yesno_im = bvmbipap_yesno
replace bvmbipap_yesno_im = . if sample_tag==0
mi set flong
mi register imputed bvmbipap_yesno_im
mi register regular studysite2 any_intervention_grp /*i.intervention_group*/ ///
age bmi apacheiiscore race3cat gender2 ///
indication_hypoxic indication_hypercarbic indication_ams ///
activesepsis activegibleeding activecopdexacerbation ///
prior6hbipap prior6hfio2high /*prior6hspo2low*/ pressorssixhoursprior baselinesat_dcs
preoxpp_group
mi impute chained (logit) bvmbipap_yesno_im = i.studysite2 any_intervention_grp
/*i.intervention_group*/ ///
age bmi apacheiiscore i.race3cat i.gender2 ///
indication_hypoxic indication_hypercarbic indication_ams ///
activesepsis activegibleeding activecopdexacerbation ///
prior6hbipap prior6hfio2high /*prior6hspo2low*/ pressorssixhoursprior baselinesat_dcs
preoxpp_group, add(1000) rseed(6068)

gsort - sample_tag + studyid + _mi_m
order p, last

drop if _mi_m==0

egen total_pos = sum(bvmbipap_yesno_im), by(_mi_id)

gen bvmbipap_yesno_im2=.
replace bvmbipap_yesno_im2 = 0 if total_pos<500
replace bvmbipap_yesno_im2 = 1 if total_pos>=500

mi extract 1, clear
*/

/*GENERATE NEW TREATMENT VARIABLE BASED ON PREDICTED PROBABILITIES*/
gen bvmbipap_yesno2 = bvmbipap_yesno
replace bvmbipap_yesno2 =0 if sample_tag==0 & p<0.5

/*BASE MODEL*/
/*BASED ON PRIOR PAPER IN SAME POPULATION LOOKING AT RISK FACTORS*/
/*SOME BALANCE PROBLEMS*/
/*
logistic bvmbipap_yesno2 i.studysite2 any_intervention_grp ///
age bmi i.race3cat ///
indication_hypoxic indication_hypercarbic baselinesat_dcs
lroc
*/

/*FINAL MODEL*/
logistic bvmbipap_yesno2 i.studysite2 any_intervention_grp ///
age bmi i.race3cat i.gender2 ///
```

```
indication_hypoxic indication_hypercarbic activesepsis activecopdexacerbation prior6hfio2high ///
baselinesat_dcs
lroc

predict pa

cem pa, tr(bvmbipap_yesno2) k2k

tabstat pa if cem_matched==1, by(bvmbipap_yesno2) stat(mean sd p50 p25 p75)
pstest baselinesat_dcs studysite2 any_intervention_grp age bmi apacheiiscore race3cat gender2
indication_hypoxic indication_hypercarbic indication_ams activesepsis activecopdexacerbation
prior6hfio2high, t(bvmbipap_yesno2) raw graph
pstest baselinesat_dcs studysite2 any_intervention_grp age bmi apacheiiscore race3cat gender2
indication_hypoxic indication_hypercarbic indication_ams activesepsis activecopdexacerbation
prior6hfio2high if cem_matched==1, t(bvmbipap_yesno2) raw graph

/*ALTERNATIVE MATCHING STRATEGY BASED ON CALIPER WIDTH*/
/*SEEMS TO BE MORE UNBALANCED*/
/*
sum pa, det
gen logitpa = logit(pa)
egen sdlogitpa = sd(logitpa)
gen calip = 0.25*sdlogitpa
gen calip_alt = sqrt(sdlogitpa)/4
di calip calip_alt
psmatch2 bvmbipap_yesno2, caliper() pscore(pa) noreplacement

tabstat nadirsat_dcs pa if _weight==1, by(bvmbipap_yesno2) stat(mean sd p50 p25 p75)
ranksum nadirsat_dcs if _weight==1, by(bvmbipap_yesno2)
tab hypox80 bvmbipap_yesno2 if _weight==1, col chi2 exact

pstest baselinesat_dcs studysite2 any_intervention_grp age bmi apacheiiscore race3cat gender2
indication_hypoxic indication_hypercarbic indication_ams activesepsis activecopdexacerbation
prior6hfio2high, t(bvmbipap_yesno2) raw graph
pstest baselinesat_dcs studysite2 any_intervention_grp age bmi apacheiiscore race3cat gender2
indication_hypoxic indication_hypercarbic indication_ams activesepsis activecopdexacerbation
prior6hfio2high if _weight==1, t(bvmbipap_yesno2) raw graph
*/

/*SAVE FINAL ANALYTIC DATASET*/
save emulation_cohort_analytic.dta, replace
clear

/*ANALYSES*/
set more off
import delimited using Database_without_PreVent_Lincoln_ED_for_UM.csv, clear

gen bvmbipap_yesno =.
replace bvmbipap_yesno=0 if ventilation_induct_laryng_none==1
replace bvmbipap_yesno=1 if ventilation_induct_laryng_none==0

label define vent 0 "No Ventilation" 1 "PPV"
label values bvmbipap_yesno vent
revrs bvmbipap_yesno, replace

dotplot nadirsat_dcs, over(bvmbipap_yesno) mcolor(gs7) center median bar nogroup msize(small)
mlwidth(vvthin) mfcolor(none) name(g1, replace) ///
ytitle("Lowest Arterial Oxygen Saturation (%)", size(small)) xtitle("") ylabel(, labsize(small))
xlabel(, labsize(small)) graphregion(color(white)) bgcolor(white)  yline(90 80 70, lw(vvthin)
lcolor(gs12) lpattern(dash))
gr_edit plotregion1.plot4.style.editstyle marker(fillcolor(cranberry)) editcopy
gr_edit plotregion1.plot4.style.editstyle marker(linestyle(color(cranberry))) editcopy
gr_edit plotregion1.plot4.style.editstyle marker(size(vsmall)) editcopy
gr_edit plotregion1.plot4.style.editstyle marker(symbol(smsquare)) editcopy
gr_edit plotregion1.plot3.style.editstyle label(textstyle(color(cranberry))) editcopy
gr_edit plotregion1.plot2.style.editstyle label(textstyle(color(cranberry))) editcopy
clear

use emulation_cohort_analytic.dta

/*TABLE 1*/
```

```
tabstat age bmi apacheiiscore prior6hfio2high prior6hspo2low if cem_matched==1,
by(bvmbipap_yesno2) stat(n p50 p25 p75)
foreach var of varlist age bmi apacheiiscore prior6hfio2high prior6hspo2low {
reg `var' bvmbipap_yesno2 if cem_matched==1
}
foreach var of varlist gender race3cat activesepsis activegibleeding pressorssixhoursprior
indication_hypoxic indication_hypercarbic indication_ams prior6hbipap {
tab `var' bvmbipap_yesno2 if cem_matched==1, col chi2
}

/*TABLE 2*/
tabstat nadirsat_dcs sat_decrease vfds icufds if cem_matched==1, by(bvmbipap_yesno2) stat(n p50
p25 p75)
foreach var of varlist nadirsat_dcs sat_decrease vfds icufds {
reg `var' bvmbipap_yesno2 if cem_matched==1
}

foreach var of varlist nadirsat_dcs sat_decrease vfds icufds {
ranksum `var' if cem_matched==1, by(bvmbipap_yesno2)
}

foreach var of varlist hypox80 hypox90 hypox70 aspiration_new_dcs carrest1hrintub
died_in_hospital {
tab `var' bvmbipap_yesno2 if cem_matched==1, col chi2 exact
glm `var' bvmbipap_yesno2 if cem_matched==1, link(log) fam(bin) eform
}

/*DOTPLOT FIGURE*/
label define vent 0 "No Ventilation" 1 "PPV"
label values bvmbipap_yesno2 vent
revrs bvmbipap_yesno2, replace

dotplot nadirsat_dcs, over(bvmbipap_yesno2) mcolor(gs7) center median bar nogroup msize(small)
mlwidth(vvthin) mfcolor(none) name(g2, replace) ///
yscale(off) xtitle("") ylabel(, labsize(small)) xlabel(, labsize(small))
graphregion(color(white)) bgcolor(white) yline(90 80 70, lw(vvthin) lcolor(gs12) lpattern(dash))
gr_edit plotregion1.plot4.style.editstyle marker(fillcolor(cranberry)) editcopy
gr_edit plotregion1.plot4.style.editstyle marker(linestyle(color(cranberry))) editcopy
gr_edit plotregion1.plot4.style.editstyle marker(size(vsmall)) editcopy
gr_edit plotregion1.plot4.style.editstyle marker(symbol(smsquare)) editcopy
gr_edit plotregion1.plot3.style.editstyle label(textstyle(color(cranberry))) editcopy
gr_edit plotregion1.plot2.style.editstyle label(textstyle(color(cranberry))) editcopy
dotplot nadirsat_dcs if cem_matched==1, over(bvmbipap_yesno2) mcolor(gs7) center median bar
nogroup msize(small) mlwidth(vvthin) mfcolor(none) name(g3, replace) ///
yscale(off) xtitle("") ylabel(, labsize(small)) xlabel(, labsize(small))
graphregion(color(white)) bgcolor(white) yline(90 80 70, lw(vvthin) lcolor(gs12) lpattern(dash))
gr_edit plotregion1.plot4.style.editstyle marker(fillcolor(cranberry)) editcopy
gr_edit plotregion1.plot4.style.editstyle marker(linestyle(color(cranberry))) editcopy
gr_edit plotregion1.plot4.style.editstyle marker(size(vsmall)) editcopy
gr_edit plotregion1.plot4.style.editstyle marker(symbol(smsquare)) editcopy
gr_edit plotregion1.plot3.style.editstyle label(textstyle(color(cranberry))) editcopy
gr_edit plotregion1.plot2.style.editstyle label(textstyle(color(cranberry))) editcopy
graph combine g1 g2 g3, ycommon row(1) name(combined, replace) graphregion(color(white))
imargin(small)
graph export Trial_Emulation_Dotplot_2172019.png, replace
graph export Trial_Emulation_Dotplot_2172019.pdf, replace
```