Risk Factors for and Prediction of Hypoxemia during Tracheal Intubation of Critically Ill Adults

Andrew C. McKown¹, Jonathan D. Casey¹, Derek W. Russell², Aaron M. Joffe³, David R. Janz⁴, Todd W. Rice¹, and Matthew W. Semler¹; on behalf of the Pragmatic Critical Care Research Group

¹Division of Allergy, Pulmonary, and Critical Care Medicine, Vanderbilt University Medical Center, Nashville, Tennessee; ²Division of Pulmonary, Allergy, and Critical Care Medicine, University of Alabama at Birmingham, Birmingham, Alabama; ³Department of Anesthesiology and Pain Medicine, University of Washington, Harborview Medical Center, Seattle, Washington; and ⁴Section of Pulmonary/Critical Care and Allergy/Immunology, Louisiana State University School of Medicine, New Orleans, Louisiana

ORCID IDs: 0000-0002-7017-3615 (A.C.M.); 0000-0001-7188-4200 (J.D.C.); 0000-0002-2716-1344 (D.W.R.); 0000-0002-7136-5408 (T.W.R.); 0000-0002-7664-8263 (M.W.S.).

Abstract

Rationale: Hypoxemia is a common complication during tracheal intubation of critically ill adults and is a frequently used endpoint in airway management research. Identifying patients likely to experience low oxygen saturations during tracheal intubation may be useful for clinical practice and clinical trials.

Objectives: To identify risk factors for lower oxygen saturations and severe hypoxemia during tracheal intubation of critically ill adults and develop prediction models for lowest oxygen saturation and hypoxemia.

Methods: Using data on 433 intubations from two randomized trials, we developed linear and logistic regression models to identify preprocedural risk factors for lower arterial oxygen saturations and severe hypoxemia between induction and 2 minutes after intubation. Penalized regression was used to develop prediction models for lowest oxygen saturation after induction and severe hypoxemia. A simplified six-point score was derived to predict severe hypoxemia.

Results: Among the 433 intubations, 426 had complete data and were included in the model. The mean (standard deviation) lowest oxygen saturation was 88% (14%); median (interquartile range) was 93% (83–98%). Independent predictors of severe hypoxemia

included hypoxemic respiratory failure as the indication for intubation (odds ratio [OR], 2.70; 95% confidence interval [CI], 1.58–4.60), lower oxygen saturation at induction (OR, 0.92 per 1% increase; 95% CI, 0.89–0.96 per 1% increase), younger age (OR, 0.97 per 1-year increase; 95% CI, 0.95–0.99 per 1-year increase), higher body mass index (OR, 1.03 per 1 kg/m²; 95% CI, 1.00–1.06 per 1 kg/m²), race (OR, 4.58 for white vs. black; 95% CI, 1.97–10.67; OR, 4.47 for other vs. black; 95% CI, 1.19–16.84), and operator with fewer than 100 prior intubations (OR, 2.83; 95% CI, 1.37–5.85). A six-point score using the identified risk factors predicted severe hypoxemia with an area under the receiver operating curve of 0.714 (95% CI, 0.653 to 0.778).

Conclusions: Lowest oxygen saturation and severe hypoxemia during tracheal intubation in the intensive care unit can be accurately predicted using routinely available preprocedure clinical data, with saturation at induction and hypoxemic respiratory failure being the strongest predictors. A simple bedside score may identify patients at risk for hypoxemia during intubation to help target preventative interventions and facilitate enrichment in clinical trials.

Keywords: intubation; intratracheal; hypoxia; predictive value of tests; critical care

(Received in original form February 16, 2018; accepted in final form July 30, 2018)

Supported by the Vanderbilt Institute for Clinical and Translational Research grant UL1 TR000445 from National Center for Advancing Translational Sciences/National Institutes of Health (NIH); National Heart, Lung, and Blood Institute (NHLBI) K12 award K12HL133117 (M.W.S.); NIH/NHLBI award T32HL105346-07 (D.W.R.); and NIH/NHLBI award 2T32HL087738-12 (J.D.C.). The funding institutions had no role in conception, design, or conduct of the study; collection, management, analysis, interpretation, or presentation of the data; or preparation, review, or approval of the manuscript.

Author Contributions: Study concept and design: A.C.M., T.W.R., and M.W.S.; acquisition of data: J.D.C., D.W.R., A.M.J., D.R.J., and M.W.S.; analysis and interpretation of data: A.C.M., T.W.R., and M.W.S.; drafting of the manuscript: A.C.M. and M.W.S.; critical revision of the manuscript for important intellectual content: J.D.C., D.W.R., A.M.J., D.R.J., and T.W.R.; statistical analysis: A.C.M.; A.C.M. and M.W.S. had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Correspondence and requests for reprints should be addressed to Matthew W. Semler, M.D., M.Sc., 1161 21st Avenue South, C-1216 MCN, Nashville, TN 37232-2650. E-mail: matthew.w.semmler@vanderbilt.edu.

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

Ann Am Thorac Soc Vol 15, No 11, pp 1320–1327, Nov 2018

Copyright © 2018 by the American Thoracic Society

DOI: 10.1513/AnnalsATS.201802-118OC

Internet address: www.atsjournals.org

Tracheal intubation is a frequently performed and potentially life-saving procedure among critically ill patients (1, 2). Hypoxemia is the most common complication of tracheal intubation in the intensive care unit (ICU) (2–5) and is associated with cardiac arrest (2, 6, 7). Thus, avoidance of hypoxemia during tracheal intubation is a goal in clinical practice, and hypoxemia is a frequently used endpoint in airway management research (8–11).

Identifying patients likely to experience low oxygen saturation during tracheal intubation in the ICU may improve clinical practice by encouraging changes in modifiable risk factors. Most prior studies have identified risk factors for “difficult airways” (5) and demonstrated an association between the number of laryngoscopy attempts required for tracheal intubation and hypoxemia (4, 12). Only one prior study has directly evaluated risk factors for hypoxemia (13). In addition to helping clinicians identify patients at high risk for peri-intubation hypoxemia (for whom changes to the procedural approach may be warranted), an accurate predictive model of oxygen saturation and severe hypoxemia could be used in clinical research to enrich trial populations or to assess for heterogeneity of treatment effect by baseline risk of the outcome (14).

We hypothesized that the lowest arterial oxygen saturation and severe hypoxemia during tracheal intubation of critically ill adults could be predicted using routinely available preprocedure clinical data. Some of these results have been previously presented in abstract form (15).

Methods

Study Design

We performed a retrospective analysis of individual patient-level data from two previously published prospective randomized trials, which examined emergent airway management among critically ill adults (8–11). The first was a single-center trial of 150 tracheal intubations randomized to video laryngoscopy or direct laryngoscopy, factorialized with randomization to apneic oxygenation or no apneic oxygenation (8, 9). The second was a multicenter trial of 292 intubations randomized to ramped or sniffing position, coenrolled with randomization to use of a written,

preprocedural checklist or usual care (10, 11). These trials are summarized in Table E1 in the online supplement. None of the four interventions studied were found to have a significant impact on lowest arterial oxygen saturation. All trials were approved by the institutional review board at participating sites and registered at ClinicalTrials.gov.

Patient Population

Complete details of the inclusion and exclusion criteria for the original trials have been published previously (8–11). Briefly, patients 18 years or older undergoing tracheal intubation in a participating ICU were eligible. Patients were excluded if tracheal intubation was deemed by the operator to be too emergent for safe conduct of study procedures (e.g., cardiac arrest), the operator believed any of the treatment assignments was contraindicated, or the initial intubation attempts were not made by a pulmonary and critical care fellow or anesthesiology trainee. The current study included all subjects from the original trials except those missing pulse oximetry data. Five patients in the dataset met the initial trials’ exclusion criteria because, although the planned initial operator was an anesthesiology trainee, the actual initial operator was a certified registered nurse anesthetist. We included these intubations in our primary analysis and performed sensitivity analyses excluding these intubations.

Outcomes

The primary endpoint for the current study was the lowest arterial oxygen saturation measured by continuous noninvasive pulse oximetry (SpO_2) from the time that induction medications were administered until 2 minutes after intubation. This outcome was the primary endpoint for the original trials and was routinely collected in real time by independent observers for all participants. The secondary endpoint for the current study was severe hypoxemia, defined as an SpO_2 less than 80% or a decrease in SpO_2 of greater than 10% for patients with an SpO_2 at induction less than 90% (7).

Data Collection

Other variables included in the current analysis were systematically collected during the original trials using a standardized data collection sheet and stored in a secure, online research electronic data capture

(REDCap) database (16). Methods for data collection are described in the original publications (8–11). In brief, independent observers not participating in the intubation procedure recorded data from the intubation, including the SpO_2 at induction, lowest SpO_2 , and duration of intubation. The operator self-reported his or her prior number of intubations at the time of each study intubation. Study personnel collected each patient’s baseline characteristics, comorbidities, severity of illness, ICU admission diagnoses, indication for tracheal intubation, and clinical outcomes from the electronic medical record.

Sample Size

To avoid overfitting, simulation studies suggest at least 10 to 15 outcome events per degree of freedom (17, 18). With a fixed sample size of 433 intubations, and the continuous primary outcome of lowest oxygen saturation, we limited the number of covariates, interactions, and nonlinear terms in our regression modeling to fewer than 29 degrees of freedom.

Development of the Regression Models for Risk Factor Identification

Potential predictors of lowest oxygen saturation were selected *a priori* for inclusion in multiple regression models on the basis of the authors’ clinical perceptions of likely risk factors for hypoxemia from demographic and physiologic variables available before intubation without univariate testing of relationships with the outcome (19). Among related variables, data reduction methods were used to avoid overfitting. Candidate variables for inclusion are listed in Table E2.

A linear regression model for the primary outcome of lowest SpO_2 between induction and 2 minutes after intubation was fit using the covariates of age, sex, race, body mass index (BMI), Acute Physiology and Chronic Health Evaluation (APACHE) II score, indication for intubation (hypoxemic respiratory failure, hypercarbic respiratory failure, or other), use of noninvasive ventilation in the preceding 6 hours, maximum fraction of inspired O_2 in preceding 6 hours, operator’s number of prior intubations, and presence of sepsis. Nonlinearity was allowed in continuous covariates using restricted cubic splines with three knots. Complete case analysis was performed. Significance for linearity and individual risk factors were assessed by *F* test

of the nested models. Accuracy of the model was assessed by coefficient of determination, corrected for optimism with 200-fold internal bootstrapping.

A logistic regression model was then constructed to identify risk factors for the secondary outcome of severe hypoxemia. To avoid overfitting, the logistic model used the same covariates as the linear regression but excluded sepsis, APACHE II, and recent fraction of inspired O₂ or noninvasive ventilation use. Continuous covariates were restricted to linearity, indication for intubation was collapsed to hypoxemic respiratory failure or other, and operator experience was dichotomized at 100 intubations on the basis of results of the primary analysis. The model fit was assessed

by area under the receiver operating curve (AUC).

After constructing the models, the cohort was described by number and proportion or median and interquartile range (IQR) for dichotomous and continuous variables. Chi-square and Wilcoxon rank-sum tests were applied for univariate comparison between intubations resulting in severe hypoxemia or not. Multiple sensitivity analyses were conducted with 1) inclusion of the original trial from which the data derived as an additive effect, 2) septic shock instead of sepsis, and 3) excluding intubations by certified registered nurse anesthetists from the data set. In addition, mixed effect models were constructed with random intercepts for

1) study sites, and 2) airway operators, to account for possible clustering effects.

Development of the Regression Models and Simplified Score for Risk Prediction

To facilitate prediction of lowest oxygen saturation or hypoxemia in external patient populations, we used the same variables as above to develop risk prediction models for lowest SpO₂ and severe hypoxemia, using penalized maximum likelihood estimation to improve external generalizability. Finally, we incorporated all significant variables from the logistic model for severe hypoxemia into a simple point score for bedside risk stratification for severe hypoxemia. Breakpoints for continuous

Table 1. Baseline characteristics

	Complete Cases (N = 433)	No Severe Hypoxemia (n = 340)	Severe Hypoxemia (n = 93)	P Value
Age, yr	58 (47; 67)	59 (48; 68)	53 (45; 62)	0.002
Body mass index, kg/m ² (n = 427)	27.8 (23.7; 32.8)	27.4 (23.4; 32.2)	28.9 (24.8; 34.2)	0.08
Male	241 (56)	192 (56)	49 (53)	0.52
Race (n = 431)				0.03
White	321 (75)	243 (72)	78 (84)	
Black	87 (20)	77 (23)	10 (11)	
Other	23 (5)	18 (5)	5 (5)	
APACHE II score	22 (17; 26)	22 (17; 26)	21 (17; 25)	0.29
Sepsis	212 (49)	156 (46)	56 (60)	0.02
Septic shock	111 (26)	79 (23)	32 (34)	0.03
Indication for intubation				<0.001
Hypercarbic respiratory failure	34 (8)	34 (10)	0 (0)	
Hypoxemic respiratory failure	154 (36)	105 (31)	49 (53)	
Other	245 (57)	201 (59)	44 (47)	
Preoxygenation				
100% Non-rebreather mask	198 (46)	152 (45)	46 (49)	0.42
Noninvasive ventilation	119 (27)	95 (28)	24 (26)	0.70
Bag-valve mask	112 (26)	88 (26)	24 (26)	1.00
Nasal cannula	23 (5)	19 (5)	5 (5)	0.94
Other	11 (2)	4 (1)	7 (8)	0.003
None recorded	21 (5)	18 (5)	3 (3)	0.44
Laryngoscopy type				0.64
Direct	281 (65)	223 (66)	58 (62)	
Video	151 (35)	116 (34)	35 (38)	
Flexible fiber optic	1 (<1)	1 (<1)	0 (0)	
Neuromuscular blockade				0.92
Succinylcholine	190 (44)	149 (44)	41 (44)	
Rocuronium	230 (53)	181 (53)	49 (53)	
Other	2 (0.5)	2 (0.6)	0 (0)	
None	11 (3)	8 (2)	3 (3)	
Operator experience, No. of intubations	61 (40; 85)	67 (40; 90)	56 (29; 75)	0.02
Operator with <100 prior intubations	341 (79)	259 (76)	82 (88)	0.01
Operator training				0.13
Pulmonary and critical care fellow	401 (93)	311 (91)	90 (97)	
Critical care fellow	4 (1)	3 (1)	1 (1)	
Anesthesia resident	23 (5)	22 (6)	1 (1)	
CRNA	5 (1)	4 (1)	1 (1)	

Definition of abbreviations: APACHE II = Acute Physiology and Chronic Health Evaluation II; CRNA = certified registered nurse anesthetist.

Data are presented as median (25th percentile; 75th percentile) or n (%). P value for significance obtained by chi-square or Fisher exact test for categorical data and Wilcoxon rank-sum test for continuous data. Percentages may not sum to 100 because of rounding. For preoxygenation, patients may have received more than one category.

Table 2. Multivariable risk factor model for lowest oxygen saturation

Variable	Mean Difference in Predicted Lowest SpO ₂ (95% CI)	Adjusted P Value
Age, 67 yr vs. 47 yr	3.0 (1.2 to 4.8)	0.006
Sex, female vs. male	−1.9 (−4.2 to 0.5)	0.12
Body mass index, 32.8 vs. 23.7 kg/m ²	−2.1 (−3.9 to −0.3)	0.04
Sepsis, present vs. absent	−1.0 (−3.5 to 1.4)	0.42
Operator experience, 85 vs. 40 intubations	2.1 (0.3 to 3.8)	0.04
APACHE II score, 26 vs. 17	1.0 (−0.6 to 2.7)	0.41
Race		0.02
Black vs. white	4.1 (1.1 to 7.1)	
Other vs. white	−1.8 (−7.0 to 3.4)	
Indication for procedure		<0.001
Hypercapnia vs. neither	7.6 (2.9 to 12.2)	
Hypoxemia vs. neither	−3.1 (−5.8 to −0.4)	
Noninvasive ventilation in preceding 6 h vs. none	−0.9 (−3.7 to 1.9)	0.54
Highest fraction inspired O ₂ in preceding 6 h, 80% vs. 30%	−1.7 (−4.6 to 1.2)	0.28
Arterial oxygen saturation at induction, 100% vs. 95%	7.4 (4.4 to 10.4)	<0.001

Definition of abbreviations: APACHE II = Acute Physiology and Chronic Health Evaluation II; CI = confidence interval; SpO₂ = arterial oxygen saturation as measured by continuous pulse oximetry. The table displays the mean difference in predicted lowest oxygen saturation for each covariate for specified contrast. For continuous variables, the difference in lowest oxygen saturation is compared between the 75th and 25th percentiles. For example, the lowest oxygen saturation is predicted to be 3% higher for a 67-year-old patient compared with a 47-year-old patient, after adjusting for prespecified confounders. The *P* values are for the nested model comparison of the overall model, with removal of the covariate term from the model.

variables were selected by visual inspection of the partial effects plots. We assigned one point for each risk factor to balance the relative weighting of the risk factors with the goal of a reduced-complexity score. We evaluated the test characteristics of the score at various cutoff values.

Statistical analysis was performed with R Version 3.3.2 (R Foundation for Statistical Computing).

Results

Prevalence of Severe Hypoxemia

A total of 442 intubations were included in the original trials, of which 9 were excluded from the current study for incomplete pulse oximetry data—in each case because of the inability to pick up the pulse waveform required to obtain an SpO₂ reading. Of the remaining 433 intubations, the mean (standard deviation) lowest SpO₂ was 88% (14%); the median (IQR) was 93% (83–98%). Ninety-three patients (21.5%) experienced severe hypoxemia.

Operator and Patient Characteristics

Operator and patient characteristics are summarized in Table 1. Thirty-six percent

of patients were intubated for hypoxemic respiratory failure. The median (IQR) APACHE II score was 22 (17–26), 49% of patients had sepsis, and 26% had septic shock. A total of 58 unique operators contributed intubations from six ICUs in four academic medical centers.

Multivariable Modeling for Risk Factor Identification

The linear multiple regression model for the primary outcome of lowest SpO₂ between induction and 2 minutes after intubation included 426 intubations and excluded 7 intubations for missing data on BMI (*n* = 5), race (*n* = 1), or both (*n* = 1). Table 2 summarizes the independent effects of each risk factor, holding other variables to their median. Because of nonlinearity, continuous variable effects are summarized by comparing 75th to 25th percentiles, and the partial effects of independent predictors are displayed in their entirety in Figure 1. Lower SpO₂ during intubation was independently associated with lower SpO₂ at induction (mean difference in predicted lowest SpO₂ comparing an induction SpO₂ of 100 compared with 95%: 7.4%; 95% confidence interval [CI], 4.4% to 10.4%), hypoxemic respiratory failure as the

indication for intubation (hypoxemia compared with neither hypoxemia nor hypercapnia as indication: −3.1%; 95% CI, −5.8 to −0.4), younger patient age (67 yr compared with 47 yr: 3.0%; 95% CI, 1.2–4.8%), higher patient BMI (32.8 kg/m² compared with 23.7 kg/m²: −2.1%; 95% CI, −3.9 to −0.3), race (black compared with white: 4.1%; 95% CI, 1.1% to 7.1%), and more limited prior operator intubating experience (85 intubations compared with 40: 2.1%; 95% CI, 0.3–3.8%). Increasing operator prior intubating experience was significantly associated with a higher SpO₂ during intubation until an apparent plateauing of the dose–response effect around 100 intubations (*P* value for nonlinearity, 0.04; Figure 1). Relative strengths of contributions of individual variables are visually displayed by nomogram in Figure E1. With 200-fold internal bootstrapping, the *R*² corrected for optimism was 0.267, with a 0.9 quantile of absolute error of 0.95.

In the simplified logistic multiple regression model for preprocedural risk factors of severe hypoxemia, independent predictors of severe hypoxemia included hypoxemic respiratory failure as the indication for intubation (odds ratio [OR], 2.70; 95% CI, 1.58–4.60), lower SpO₂ at induction (OR, 0.92 per 1% increase; 95% CI, 0.89–0.96 per 1% increase), younger age (OR, 0.97 per 1-yr increase; 95% CI, 0.95–0.99 per 1-yr increase), higher BMI (OR, 1.03 per 1 kg/m²; 95% CI, 1.00–1.06 per 1 kg/m²), race (OR, 4.58 for white vs. black; 95% CI, 1.97–10.67; OR, 4.47 for other vs. black; 95% CI, 1.19–16.84), and operator with fewer than 100 prior intubations (OR, 2.83; 95% CI, 1.37–5.85) (Table 3). The model had good discrimination (AUC, 0.756), which persisted for the subset of patients with induction SpO₂ greater than 95% (AUC, 0.742). After 200-fold internal bootstrapping, AUC corrected for optimism was 0.732, with a 0.9 quantile of absolute error of 0.03. A nomogram (Figure E2) is available in the online supplement to visually convey relative strength of risk factors.

Results of the sensitivity analyses are presented in Table E3. In the sensitivity analysis excluding intubations by certified registered nurse anesthetists, operator experience remained significant (*P* = 0.02). The nonlinear term for operator experience

Figure 1. Predicted lowest arterial oxygen saturation for the independently associated risk factors adjusted for covariates. Each plot displays the predicted lowest arterial oxygen saturation and 95% confidence interval for the variation of one risk factor. Predictions are adjusted to the median of the remaining model covariates (indication for intubation neither hypoxemic nor hypercapnic, not receiving noninvasive ventilation in the preceding 6 h, highest fraction inspired of oxygen in the prior 6 h of 0.4, arterial oxygen saturation at induction of 99%, age 58 yr, Acute Physiology and Chronic Health Evaluation [APACHE] II score of 22, body mass index of 27.8 kg/m², male sex, white race, no sepsis, and operator experience with 61 prior procedures). SpO₂ = arterial oxygen saturation as measured by continuous pulse oximetry.

was no longer statistically significant ($P = 0.14$), but the partial effects plot was similar (Figure E3), and the model coefficients did not substantively change (Table E3). When substituting septic shock for sepsis in the model, septic shock was an independent predictor, but the other coefficient terms were not significantly changed. When including

the original trial from which the data were derived as a fixed effect, trial of data origin was not independently associated with lowest SpO₂. The mixed effect models including either study site or operator as a random intercept did not change the primary model; the variance explained by either study site or operator was zero.

Development of Risk Prediction Models

When the same variables from the models used for risk factor identification were included in a risk prediction model for lowest SpO₂ using a penalized maximum likelihood estimation algorithm further corrected for optimism with internal bootstrapping, the R^2 was 0.267, with a 0.9 quantile of absolute error of 0.82. The bootstrapped calibration plot is shown in Figure E4. The complete equation for the model is available in the online supplement.

Similarly, when we used a penalized maximum likelihood estimation algorithm on the risk factor model for severe hypoxemia to develop a prediction model, the penalized model corrected for optimism by internal bootstrapping had good discrimination, with an AUC of 0.734 and a 0.9 quantile of absolute error of 0.03. The full formula for the prediction model is available in the online supplement, as is the bootstrap calibration plot (Figure E5).

Finally, simplification of the penalized model for predicting severe hypoxemia by assigning points to each of the significant risk factors yielded a six-point score: age less

Table 3. Multivariable risk factor model for severe hypoxemia

Variable	Odds Ratio (95% CI)	Adjusted P Value
Age, per yr	0.97 (0.95–0.99)	<0.001
Sex, female vs. male	1.10 (0.66–1.83)	0.7
Body mass index, per kg/m ²	1.03 (1.00–1.06)	0.03
Less-experienced operator, <100 intubations vs. more experience	2.83 (1.37–5.85)	0.005
Race		0.002
White vs. black	4.58 (1.97–10.67)	
Other vs. black	4.47 (1.19–16.84)	
Indication, hypoxemia vs. other	2.70 (1.58–4.60)	<0.001
Arterial oxygen saturation at induction, per 1% increase	0.92 (0.89–0.96)	<0.001

Definition of abbreviation: CI = confidence interval.

The table displays the odds ratios for development of severe hypoxemia for each variable in the logistic regression model for the comparison denoted in parentheses.

Table 4. Calculation worksheet for AT RISK score

Risk Factors	Points
A: Age < 50 yr	1
T: Trainee-operator with <100 prior intubations	1
R: Race other than black	1
I: Indication is hypoxemic respiratory failure	1
S: Sp _{O₂} at induction < 94	1
K: kg/m ² (BMI) > 35	1
Total	6

Definition of abbreviations: BMI = body mass index; Sp_{O₂} = arterial oxygen saturation as measured by continuous pulse oximetry. Coded 0 to 6, with 6 indicating higher risk of severe hypoxemia.

than 50 years, operator with fewer than 100 prior intubations, race other than black, hypoxemic respiratory failure as indication for intubation, induction Sp_{O₂} less than 94, and BMI greater than 35. Calculation of the Age, Trainee, Race, Indication, Sp_{O₂}, kg/m² score (AT RISK score) is summarized in Table 4. The AUC of the AT RISK score was 0.714 (95% CI, 0.653–0.778 by bootstrap). Test characteristics and probability of desaturation at various cutoffs are summarized in Table 5, and the calibration of the AT RISK score is depicted in Figure 2. Less than 2% of patients with an AT RISK score less than 2 will experience severe hypoxemia (negative predictive value, 98.1%; 95% CI, 93.2–100%), whereas nearly half of patients with an AT RISK score greater than 3 will experience severe hypoxemia (positive predictive value, 47.4%; 95% CI, 37.1–59.1%).

Discussion

This retrospective analysis of data from two randomized trials identified preprocedural

demographic and physiologic risk factors for lower nadir Sp_{O₂} during tracheal intubation in the ICU. Oxygen saturation at induction and acute hypoxemic respiratory failure as the indication for intubation were risk factors most strongly associated with lower oxygen saturation. For operators with fewer than 100 previous intubations, increasing operator experience was associated with higher oxygen saturation. In addition, we found that clinically available preinduction risk factors could be used to predict severe hypoxemia using multivariable modeling or a simplified six-point score.

Hypoxemia is the most common complication of intubation and the most closely associated with cardiac arrest (2, 6, 7). Although several prior studies have examined risk factors for failure to intubate on the first laryngoscopy attempt (4, 5, 12, 20) or the effect of intraprocedural factors such as laryngoscopy duration on the development of hypoxemia (3–7), few prior data inform which preprocedural factors predict oxygen saturation during intubation in the ICU. Smischney and colleagues recently performed a retrospective case-control study of critical care intubations in the ICU in a cohort of 420 patients (13). They reported that provider training level, acute respiratory failure as indication for intubation, and lower baseline Sp_{O₂} levels were associated with hypoxemia, defined as Sp_{O₂} less than 90% within 30 minutes. Results from our prospectively collected cohort, using a more specific interval for airway complications of within 2 minutes recorded by an independent observer, build on and extend these findings. In addition to confirming provider experience, indication for intubation, and Sp_{O₂} at induction as risk factors for hypoxemia, we identified BMI, race, and younger age as risk factors and translated these risk factors into predictive models for lowest Sp_{O₂} and hypoxemia and a simplified six-point score that could be used

for rapid bedside assessment of risk of hypoxemia.

The strongest independent predictor of severe hypoxemia in our study was Sp_{O₂} at induction. A higher Sp_{O₂} at induction was associated with better oxygenation during the procedure across the entire spectrum of values, including for patients with Sp_{O₂} at induction greater than 95%. This finding may contrast with prior recommendations to target a minimum specific Sp_{O₂} threshold for preoxygenation (21), instead suggesting that, even for patients with an Sp_{O₂} greater than 95%, higher may be better, although the benefits of extending the preoxygenation period must be balanced against the risk for possible deterioration in some patients (22).

Another clinically modifiable variable predictive of lowest Sp_{O₂} and severe hypoxemia in our study was the operator's prior intubating experience. For operators who had performed fewer than 100 prior intubations, a "dose-response" relationship appeared to exist, in which increasing operator experience was associated with higher oxygen saturations during the intubation. For operators who had performed 100 or more prior intubations, there was not a clear association between experience and oxygenation outcomes.

Previous studies of trainees learning to perform tracheal intubation report an inconsistent relationship between operator experience and complications of intubation. In a 253-procedure French cohort of tracheal intubations, inexperienced trainees had fewer complications, a finding explained as a dual-operator advantage because less-experienced operators were closely supervised (3, 23). By contrast, nearly all intubations in our cohort were supervised by attending physicians, regardless of operator experience. Other studies have reported associations between operators at more advanced stages of training and fewer laryngoscopy attempts

Table 5. Test characteristics of the AT RISK score

AT RISK	Sensitivity	Specificity	Negative Predictive Value	Positive Predictive Value
≥1 (n = 418)	100 (100–100)	2.4 (0.9–4.2)	100 (100–100)	22.2 (18.5–26.2)
≥2 (n = 373)	98.9 (96.4–100)	15.6 (11.3–19.5)	98.1 (93.2–100)	24.7 (20.4–28.8)
≥3 (n = 223)	76.3 (68.1–85.4)	54.4 (49.2–59.8)	89.2 (84.5–93.5)	31.8 (26.7–37.6)
≥4 (n = 78)	39.8 (30.9–50.6)	87.7 (84.4–90.8)	83.9 (79.5–87.8)	47.4 (37.1–59.1)
≥5 (n = 9)	5.4 (1.1–10.8)	99.7 (97.6–100)	78.9 (75–82.8)	55.6 (25–100)

Definition of abbreviations: AT RISK = Age, Trainee, Race, Indication, Sp_{O₂}, kg/m²; Sp_{O₂} = arterial oxygen saturation as measured by continuous pulse oximetry. Test characteristics in the cohort expressed by % (95% confidence intervals) at different thresholds of AT RISK scores. The number of patients in the cohort with each score was: AT RISK 0, n = 8; AT RISK 1, n = 45; AT RISK 2, n = 150; AT RISK 3, n = 145; AT RISK 4, n = 69; AT RISK 5, n = 8; and AT RISK 6, n = 1.

Figure 2. Calibration of Age, Trainee, Race, Indication, SpO_2 , kg/m^2 (AT RISK) score with severe hypoxemia. Height of bar depicts percentage of patients with severe hypoxemia by AT RISK score. SpO_2 = arterial oxygen saturation as measured by continuous pulse oximetry.

and complications (2, 4, 5, 12, 24), including hypoxemia in the Smischney study (13). The number of procedures needed to achieve competency is a critical question for procedural training, but few emergency airway management studies have evaluated the relationship between the operator's absolute procedural volume and outcomes (25). Studies of anesthesia resident learning curves for tracheal intubation have suggested a median minimal procedural volume of about 50 tracheal intubations to achieve an acceptable intubation success rate, but the range for individual operators was large (26–28). Whether the same number is necessary or sufficient for tracheal intubations outside of the operating room is unknown, and to our knowledge has not been previously examined with respect to peri-intubation complications. Although our analysis suggests a leveling of learning beyond 100 procedures' experience, caution is warranted before assigning a specific volume of procedural experience necessary to achieve competency for intubations in the critically ill. For example, our model predicts that for an operator with prior experience of 50 procedures, the average patient would have a lowest SpO_2 of about 90%, a safe and acceptable level (29), and the predicted mean difference in lowest oxygen saturation between operator experience of 40 and 85 intubations is 2%. Furthermore, the current study was designed *a priori* to predict lowest SpO_2 and severe hypoxemia rather than to

assess procedural competency. When considering procedural competency, other operator-related characteristics may be important, and investigation into additional endpoints beyond lowest SpO_2 would be warranted.

Our models predicting absolute lowest SpO_2 and severe hypoxemia also have several implications for clinical trials. First, since nearly half of the patients who developed severe hypoxemia were intubated for reasons other than hypoxemic or hypercapnic respiratory failure, investigators attempting to enrich the study population for those at risk for hypoxemia by selectively including only patients with hypoxemic respiratory failure should be cautious. Doing so would exclude almost half the patients who experience severe hypoxemia during tracheal intubation in the ICU. Second, there is increasing recognition of the need for clinical trials to address the potential for heterogeneity of treatment effect by baseline risk of the primary outcome, because an intervention may benefit those at high risk for the outcome and may be ineffective (or even harmful) for those at low risk (14). By providing a detailed formula for predicting lowest SpO_2 and severe hypoxemia during intubation, the current study will allow future randomized trials targeting these outcomes to calculate each individual patient's baseline risk of the outcome and assess for heterogeneity of treatment effect by risk of the primary outcome.

Alternatively, researchers could quickly calculate the AT RISK score before enrollment to enrich the trial population for patients likely to experience the outcome.

Our analysis has a number of strengths. All data were collected prospectively and systematically, avoiding many of the biases of retrospective research. Moreover, an independent observer, whose only role at the procedure was to complete the trial data collection sheet, including vital sign data, recorded the primary endpoint. Prespecification of the covariates for the multiple regression model reduces the likelihood of overfitting and improves the likelihood of validation of the model in new populations. The clinical trials from which this cohort was derived were conducted in a pragmatic fashion, enrolling nearly all eligible intubations of critically ill adults in the participating units, which increases the generalizability of our models. Penalized maximum likelihood estimation and internal bootstrapping were performed to avoid overfitting and provide a more conservative prediction model, and the stability of the parameter estimates in the risk factor assessments through multiple sensitivity analyses provides evidence for robustness of their independent risk estimates.

Our investigation has some limitations. The sample size was fixed by the enrollment of the previously conducted trials, thereby limiting the number of covariates that could be analyzed. This analysis was retrospective, so we were limited to variables collected as part of the original prospective trials. Oxygen saturation was only available in the study datasets for the 2 minutes after endotracheal intubation, precluding prediction of complications at later time points. The relationship between lowest oxygen saturation, severe hypoxemia, and longer-term, patient-centered outcomes remains uncertain. Race as an independent predictor of lowest SpO_2 is without physiologic rationale and may relate to our incomplete knowledge of comorbidities that confound the race variable. The association of older age with less risk of hypoxemia appears counterintuitive. Clinicians' thresholds for intubating younger and older patients may differ, with intubation of younger patients deferred until less physiologic reserve remains. Alternatively, the relationship may be explained by residual confounding. Furthermore, the inclusion of variables reported by previous

investigators to be predictive of airway difficulty, such as the presence of obstructive apnea syndrome or higher Mallampati scores (5), may have improved our predictions, but they were not collected routinely in all patients. Our data may not generalize to patients who were excluded from our trials, namely those undergoing tracheal intubation for cardiopulmonary arrest or those for whom treating clinicians

believed tracheal intubation was required too emergently to allow for randomization. Finally, the models derived in this study will require prospective validation in another study population.

Lowest oxygen saturation and severe hypoxemia during tracheal intubation in the ICU can be accurately predicted using routinely available preprocedure clinical data, with SpO₂ at induction and hypoxemic

respiratory failure as the indication for intubation being the strongest predictors. A simple bedside risk score may risk-stratify patients at risk for hypoxemia during intubation to help target delivery of preventative interventions and facilitate prognostic enrichment in clinical trials. ■

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

References

- Martin LD, Mhyre JM, Shanks AM, Tremper KK, Kheterpal S. 3,423 emergency tracheal intubations at a university hospital: airway outcomes and complications. *Anesthesiology* 2011;114:42–48.
- Simpson GD, Ross MJ, McKeown DW, Ray DC. Tracheal intubation in the critically ill: a multi-centre national study of practice and complications. *Br J Anaesth* 2012;108:792–799.
- Jaber S, Amraoui J, Lefrant J-Y, Arich C, Cohendy R, Landreau L, et al. Clinical practice and risk factors for immediate complications of endotracheal intubation in the intensive care unit: a prospective, multiple-center study. *Crit Care Med* 2006;34:2355–2361.
- Griesdale DEG, Bosma TL, Kurth T, Isac G, Chittock DR. Complications of endotracheal intubation in the critically ill. *Intensive Care Med* 2008;34:1835–1842.
- De Jong A, Molinari N, Terzi N, Mongardon N, Arnal J-M, Guitton C, et al.; AzuRéa Network for the Frida-Réa Study Group. Early identification of patients at risk for difficult intubation in the intensive care unit: development and validation of the MACOCHA score in a multicenter cohort study. *Am J Respir Crit Care Med* 2013;187:832–839.
- Mort TC. The incidence and risk factors for cardiac arrest during emergency tracheal intubation: a justification for incorporating the ASA Guidelines in the remote location. *J Clin Anesth* 2004;16:508–516.
- De Jong A, Rolle A, Molinari N, Paugam-Burtz C, Constantin J-M, Lefrant J-Y, et al. Cardiac arrest and mortality related to intubation procedure in critically ill adult patients: a multicenter cohort study. *Crit Care Med* 2018;46:532–539.
- 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.
- Janz DR, Semler MW, Lentz RJ, Matthews DT, Assad TR, Norman BC, et al.; Facilitating Endotracheal intubation by Laryngoscopy technique and apneic Oxygenation Within the ICU Investigators and the Pragmatic Critical Care Research Group. Randomized trial of video laryngoscopy for endotracheal intubation of critically ill adults. *Crit Care Med* 2016;44:1980–1987.
- 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.
- Janz DR, Semler MW, Joffe AM, Casey JD, Lentz RJ, deBoisblanc BP, et al.; Check-UP Investigators; Pragmatic Critical Care Research Group. A multicenter randomized trial of a checklist for endotracheal intubation of critically ill adults. *Chest* 2018;153:816–824.
- Mort TC. Emergency tracheal intubation: complications associated with repeated laryngoscopic attempts. *Anesth Analg* 2004;99:607–613.
- Smischney NJ, Seisa MO, Heise KJ, Wiegand RA, Busack KD, Loftsgard TO, et al. Predictors of arterial desaturation during intubation: a nested case-control study of airway management-part I. *J Thorac Dis* 2017;9:3996–4005.
- 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.
- McKown AC, Casey JD, Russell DW, Joffe AM, Janz DR, Rice TW, et al. Predictors of hypoxemia during endotracheal intubation of critically ill adults [abstract]. *Am J Respir Crit Care Med* 2018;197:A1116.
- Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–381.
- Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. New York: Springer Science & Business Media; 2008.
- Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;15:361–387.
- Harrell FE. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis, 2nd ed. New York: Springer; 2015.
- Li J, Murphy-Lavoie H, Bugas C, Martinez J, Preston C. Complications of emergency intubation with and without paralysis. *Am J Emerg Med* 1999;17:141–143.
- Pourmand A, Robinson C, Dorwart K, O'Connell F. Pre-oxygenation: implications in emergency airway management. *Am J Emerg Med* 2017;35:1177–1183.
- Mort TC, Waberski BH, Clive J. Extending the preoxygenation period from 4 to 8 mins in critically ill patients undergoing emergency intubation. *Crit Care Med* 2009;37:68–71.
- Jaber S, Jung B, Corne P, Sebbane M, Muller L, Chanques G, et al. An intervention to decrease complications related to endotracheal intubation in the intensive care unit: a prospective, multiple-center study. *Intensive Care Med* 2010;36:248–255.
- Schwartz DE, Matthey MA, Cohen NH. Death and other complications of emergency airway management in critically ill adults: a prospective investigation of 297 tracheal intubations. *Anesthesiology* 1995;82:367–376.
- Roux D, Reignier J, Thiery G, Boyer A, Hayon J, Souweine B, et al. Acquiring procedural skills in ICUs: a prospective multicenter study. *Crit Care Med* 2014;42:886–895.
- Konrad C, Schüpfer G, Wietlisbach M, Gerber H. Learning manual skills in anesthesiology: is there a recommended number of cases for anesthetic procedures? *Anesth Analg* 1998;86:635–639.
- de Oliveira Filho GR. The construction of learning curves for basic skills in anesthetic procedures: an application for the cumulative sum method. *Anesth Analg* 2002;95:411–416.
- Komatsu R, Kasuya Y, Yogo H, Sessler DI, Mascha E, Yang D, et al. Learning curves for bag-and-mask ventilation and orotracheal intubation: an application of the cumulative sum method. *Anesthesiology* 2010;112:1525–1531.
- O'Driscoll BR, Howard LS, Davison AG; British Thoracic Society. BTS guideline for emergency oxygen use in adult patients. *Thorax* 2008;63:vi1–vi68.