

## INTRODUCTION & STUDY AIM

- Postoperative respiratory failure (PRF) is a key surgical complication associated with morbidity, mortality, and functional decline
- PRF is an important quality metric
- PRF incidence after non-cardiac surgery is 1.5-4.0%
- This incidence has remained unchanged in over two decades [1-9]
- Existing tools exhibit poor accuracy and performance [10-12]
- One reason may be over simplistic modeling of non-linear, physiologic parameters and less informative comorbidity information
- There have been calls to employ machine learning (ML) for modeling [13, 14]
- **Study Aim:** Develop a best-in-class PRF prediction tool routine clinical variables combined with ML modeling

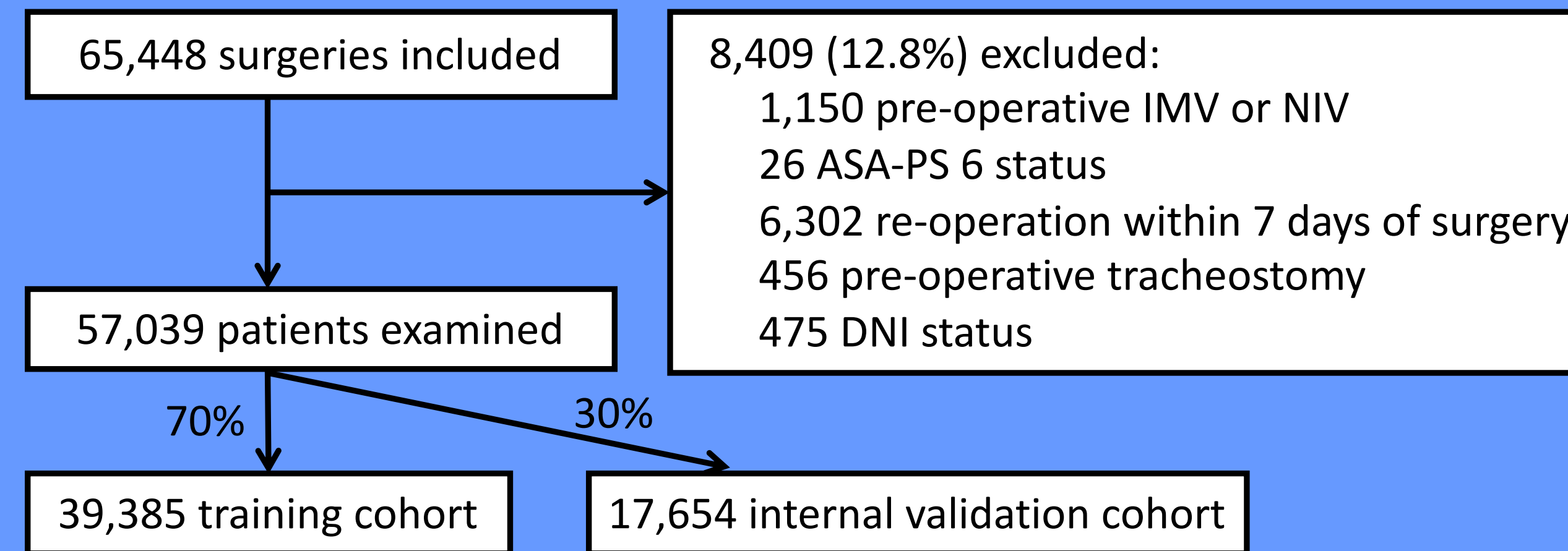
## METHODS

- **Design:** Single center retrospective cohort study at Montefiore Medical Center
- **Inclusion criteria:** ≥18-years-old, same-day or inpatient non-labor analgesia procedures from 1/2018-6/2021.
- **Exclusion criteria:** as illustrated in **Figure 1**.
- **novel RESPIRED composite outcome:** (1) unplanned advanced REspiratory Support including non-invasive ventilation; (2) Prolonged Intubation after the OR for ≥24-hours; (3) Reintubation; and/or (4) Early Death within 7 days of surgery
- 25 routine pre-operative clinical variables were selected *a priori* based on a literature review of PRF risk factors and abstracted from electronic records
- Missingness was imputed as mean or mode and confirmed with Multiple Imputations by Chained Equations; 15 missingness indicator variables generated
- **Model development:** Cohort was randomly split into training (70%) and internal validation (30%) cohorts. All accuracy and performance tests in validation cohort.
- Random Forest algorithm was used to predict outcome in the training cohort based on an ensemble of 1000 random classification decision trees
- AUROC, Brier score, and Area Under Precision-Recall Curve (AUPRC) were used for accuracy testing
- Youden's index and the F1 score were used to stratify score cut-offs
- Sensitivity, PPV, and FPR for performance evaluation
- RESPIRED was compared with two other PRF models: Assess Respiratory Risk in Surgical Patients in Catalonia (ARISCAT) and Score for Prediction of Postoperative Respiratory Complications (SPORC-1)

## ABBREVIATIONS & DEFINITIONS

PRF = post-operative respiratory failure; StEP = Standardized Endpoints for Perioperative Medicine; IMV = invasive mechanical ventilation; NIV = non-invasive ventilation; PPC = post-operative pulmonary complications; ARISCAT = Assess Respiratory Risk in Surgical Patients in Catalonia; ML = machine learning; EHR = electronic health record.; CDS, clinical decision support; OR = operating room; RESPIRED = REspiratory Support including NIV, Prolonged Intubation ≥24-hours, Reintubation, and/or Early Death within 7 days of surgery; SPORC-1 = Score for Prediction of Postoperative Respiratory Complications; AUROC = Area Under Receiver Operating Curve; AUPRC = Area Under Precision-Recall Curve, PPV = positive predictive value; NPV = negative predictive value; LOS = length of stay; ICU = intensive care unit; n/a = not applicable.

**FIGURE 1.** Study cohort inclusion/exclusion criteria

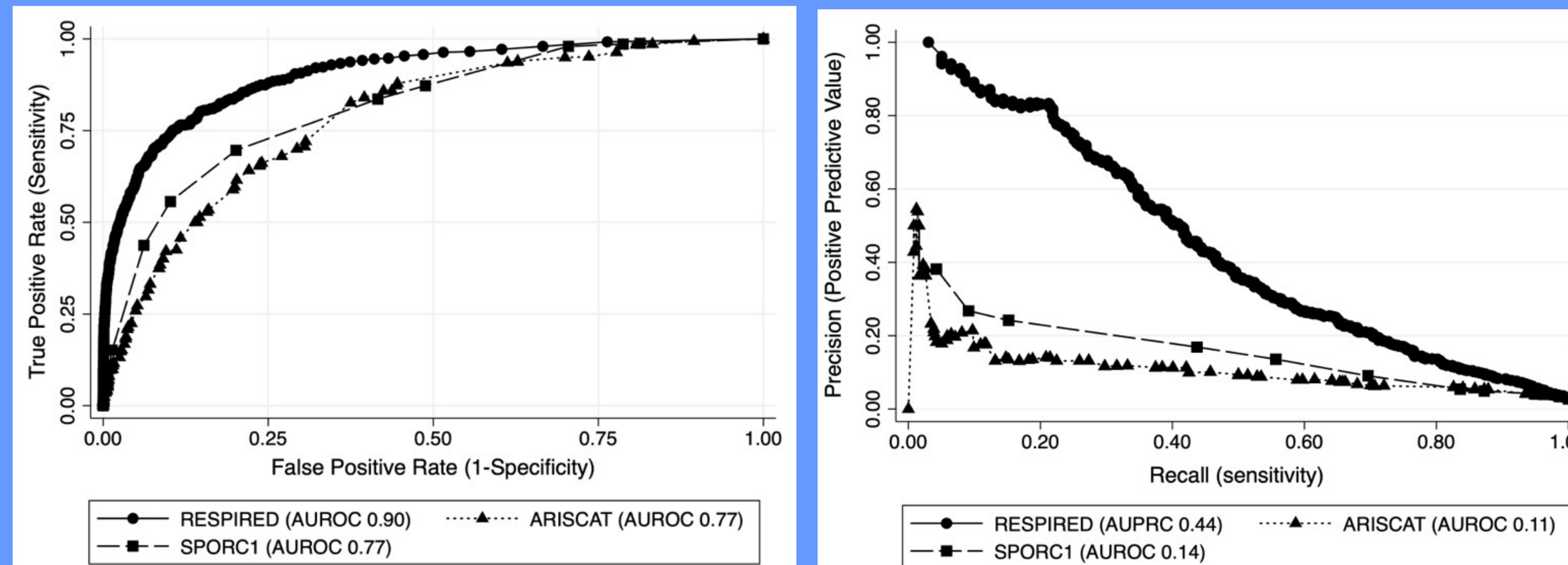


**Table 1.** Subcomponents & secondary outcomes associated with the 7-day post-operative RESPIRED composite outcome

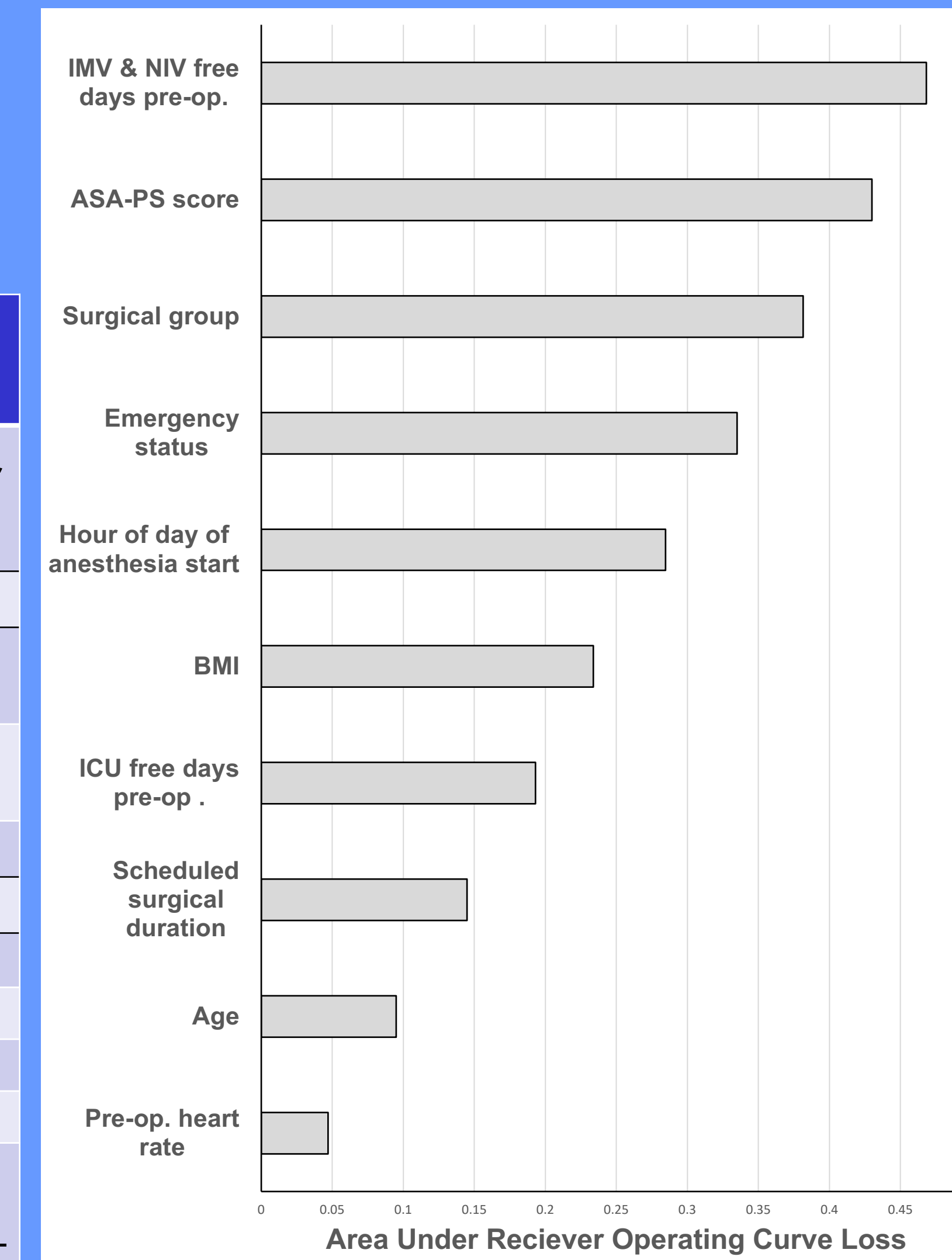
	RESPIRED outcome (n=1,630; 2.9%)	No RESPIRED outcome (n=55,409, 97.1%)	P value for difference
<i>Subcomponents of RESPIRED composite outcome</i>			
≥24-hours continued intubation after surgery, n	860 (52.8%)	0	n/a
Unplanned re-intubation within 7 days, n	642 (39.3%)	0	n/a
Postop 7-day mortality, n	368 (22.6%)	0	n/a
<i>Secondary outcomes associated with RESPIRED composite outcome</i>			
StEP PRF outcome, n	1,456 (~90%)	293 (0.5%)	<0.001
ICU admit within 30 days, n	1,446 (88.7%)	6,512 (11.8%)	<0.001
Postop. hospital LOS, days	14 (7, 25)	2 (1, 5)	<0.0001
Post-p. 30-day mortality, n	541 (33.2%)	419 (0.8%)	<0.001

Column values indicate the incidence (percent) or median time (IQR 25-75%) within either the RESPIRED or the non-RESPIRED cohort sub-group. Adverse hospital discharge was defined as new admission to a skilled nursing facility or long-term acute care hospital or inpatient death. Groups were compared using Chi-Squared, Fisher's Exact, or Mann Whitney U tests, where appropriate. Abbreviations not previously specified: IMV, invasive mechanical ventilation; StEP PRF, Standardized Endpoints for Perioperative Medicine-defined post-operative respiratory failure composite outcome.

**FIGURE 3.** Superior accuracy of the RESPIRED model in the validation cohort compared to ARISCAT and SPORC-1 as assessed by (A) AUROC and (B) Area Under Precision-Recall Curve



**FIGURE 2.** Top 10 most important clinical variables for accuracy of the RF-based RESPIRED model as assessed by loss in AUROC with variable elimination



**Table 3.** Superior performance of RESPIRED model compared with competitor ARISCAT and SPORC-1 models

	RESPIRED		ARISCAT		SPORC-1	
	Moderate-Risk Score ≥0.04	High-Risk Score ≥0.29	Moderate-Risk Score ≥26	High-Risk Score ≥52	Moderate-Risk Risk ≥5	High-Risk Risk ≥8
<b>Sensitivity (95% CI)</b>	81% (77-84%)	40% (36-45%)	84% (81-87%)	32% (28-36%)	84% (80-87%)	44% (39-48%)
<b>Specificity (95% CI)</b>	83% (83-84%)	99% (98-99%)	61% (60-61%)	93% (93-94%)	58% (58-59%)	94% (93-94%)
<b>PPV (95% CI)</b>	12% (11-13%)	51% (46-56%)	6% (5-6%)	12% (10-14%)	6% (5-6%)	17% (15-19%)
<b>FPR (95% CI)</b>	17% (16-17%)	1% (1-2%)	39% (39-40%)	7% (6-7%)	42% (41-42%)	6% (7-6%)

Moderate- and high-risk score cut-points for each prediction model were estimated using Youden's index and the F1 score, respectively.

**Table 4.** Good performance of RESPIRED model for prediction of key secondary outcomes

	Expert Consensus StEP PRF Outcome		30-day ICU admission	
	Moderate-Risk	High-Risk	Moderate-Risk	High-Risk
<b>Sensitivity (95% CI)</b>	81% (77-84%)	37% (33-41%)	66% (64-64%)	14% (13-15%)
<b>Specificity (95% CI)</b>	84% (83-84%)	99% (98-99%)	89% (89-90%)	99% (99-99%)
<b>PPV (95% CI)</b>	13% (13-15%)	51% (41-64%)	51% (49-52%)	93% (89-95%)
<b>FPR (95% CI)</b>	16% (16-17%)	1% (1-2%)	11% (10-11%)	1% (0-1%)

This model was calibrated for its intended outcome, RESPIRED, as indicated in **Table 3**.

## CONCLUSIONS & FUTURE WORK

- ML-based RESPIRED prediction score was more accurate and demonstrated better clinical usability than two leading scores, with lower false-positive rate and higher positive-predictive value at a comparable sensitivity
- Performed equally well for an expert consensus definition of PRF
- Also useful for predicting ICU admission
- Confirms utility of ML for adverse surgical outcomes *Limitations:* single center study without external validation
- *Strengths:* ~60,000 surgical patients, ethnically diverse, recent cohort reflecting latest perioperative techniques, and includes nearly all procedures with exception of outpatient surgery
- *Future work:* external validation, prospective implementation, development of related post-operative RESPIRED score

## REFERENCES

- Abbott et al. Br J Anaesth 2018;
- Canet et al. Eur J Anaesthesiol 2015
- Arozullah et al. Ann Surg 2000
- Fernandez-Bustamante et al. JAMA Surg 2017
- Rostin et al. Anaesthesia 2019
- Dasta et al. Crit Care Med 2005 https
- Gershengorn et al. Ann Am Thorac Soc 2015
- AHRQ. Patient Safety Indicators 2021.
- CMS BENCHMARKS 2016.
- Canet J et al. Anesthesiology 2010
- Brueckmann B et al. Anesthesiology 2013
- Dziadzko et al. Crit Care 2018
- Nijbroek et al. Curr Opin Anaesthesiol 2019
- Churpek et al. Crit Care Med 2016