Racial Differences in Accuracy of Predictive Models for High-Flow Nasal Cannula Failure in COVID-19

Supplementary Materials

Philip Yang, MD, MSc, Ismail A. Gregory, MD, Chad Robichaux, MPH, Andre L. Holder, MD, MSc, Greg S. Martin, MD, MSc, Annette M. Esper, MD, MSc, Rishikesan Kamaleswaran, PhD, Judy W. Gichoya, MD, Sivasubramanium V. Bhavani, MD

Table of Contents

eTable 1. Full list of study inclusion and exclusion criteria.

Inclusion Criteria	• Age ≥18 years
	• Admitted to one of four Emory University hospitals listed below between March 2020 and April 2022
	  ○ Emory University Hospital
	  ○ Emory University Hospital Midtown
	  ○ Emory Saint Joseph's Hospital
	  ○ Emory Johns Creek Hospital
	• Diagnosis of COVID-19
	  ○ Positive PCR test result for SARS-CoV-2, and/or
	  ○ ICD-10 billing code for COVID-19
	• Received HFNC therapy
	  ○ Initiated within the first 24 hours of admission
	  ○ For total HFNC therapy duration of ≥6 hours

Exclusion Criteria	• Age <18 years • Admission date outside of the date range specified above • Hospital length of stay <24 hours • Unknown final disposition due to transfer to another hospital • HFNC therapy started more than 24 hours after admission • Total HFNC therapy duration <6 hours

Abbreviations: PCR = polymerase chain reaction, ICD = International Classification of Diseases, HFNC = high-flow nasal cannula

eTable 2. Full list of clinical variables included in the development of predictive models and the number of missing values in these variables in HFNC failure and non-failure patients.

Variable	Number with missing values	
	Failure n=317	**Non-failure n=667**
Age	0	0
Sex	0	0
Elixhauser comorbidities	5	1
Congestive heart failure		
Cardiac arrhythmias		
Valvular disease		
Pulmonary circulation disorders		
Hypertension, uncomplicated		
Hypertension, complicated		
Chronic pulmonary disease		
Diabetes, uncomplicated		
Diabetes, complicated		
Renal failure		
Liver disease		
AIDS		
Lymphoma		
Metastatic cancer		
Solid tumor without metastasis		
Rheumatoid arthritis or CTD		
Coagulopathy		
Obesity		
Fluid and electrolyte disorders		
Vital signs		
HR, max/min/median/mean	0	0
HR, sd	11	6
SBP, max/min/median/mean	0	3
SBP, sd	16	14
DBP, max/min/median/mean	0	3
DBP, sd	16	14
MAP by cuff, max/min/median/mean	19	28
MAP by cuff, sd	56	82
Respiratory rate, max/min/median/mean	0	0
Respiratory rate, sd	16	13
Pulse oximetry, max/min/median/mean	0	0
Pulse oximetry, sd	10	6
Temperature, max/min/median/mean	0	0
Temperature, sd	9	4
Labs		
Sodium, max/min	0	0
Potassium, max/min	0	0
Bicarbonate, max/min	0	0

Blood urea nitrogen, max	0	0
Creatinine, max	0	0
White blood cell count, max/min	0	0
Hemoglobin, min	0	0
Hematocrit, min	0	0
Platelet, min	0	0
Aspartate transferase, max	0	7
Alanine transferase, max	0	7
Alkaline phosphatase, max	0	7
Total bilirubin, max	0	7
C-reactive protein, max	35	97
D-dimer, max	18	63
Arterial blood gas		
pH, max/min	28	171
PaO2, max/min	47	221
PaCO2, max/min	47	220
Treatments		
Dexamethasone	0	0
Remdesivir	0	0
Bolus intravenous fluids	82	185
Norepinephrine	0	0
Epinephrine	0	0
Phenylephrine	0	0
Vasopressin	0	0
Dobutamine	0	0
Dopamine	0	0

Abbreviations: AIDS = acquired immunodeficiency syndrome, CTD = connective tissue disease, HR = heart rate, max = maximum, min = minimum, sd = standard deviation, SBP = systolic blood pressure, DBP = diastolic blood pressure, MAP = mean arterial pressure.

eTable 3. List of R packages utilized for data analysis.

boot
caret
caTools
DescTools
doParallel
foreign
gbm
gdata
ggplot2
ggpubr
givitiR
glmnet
gtools
imputeMissings
lattice
magrittr
mice
MLmetrics
pdp
predtools
pROC
PRROC
qwraps2
randomForest
RANN
readr
reshape2
rms
ROCR
SHAPforxgboost
tidymodels
tidyverse
xgboost

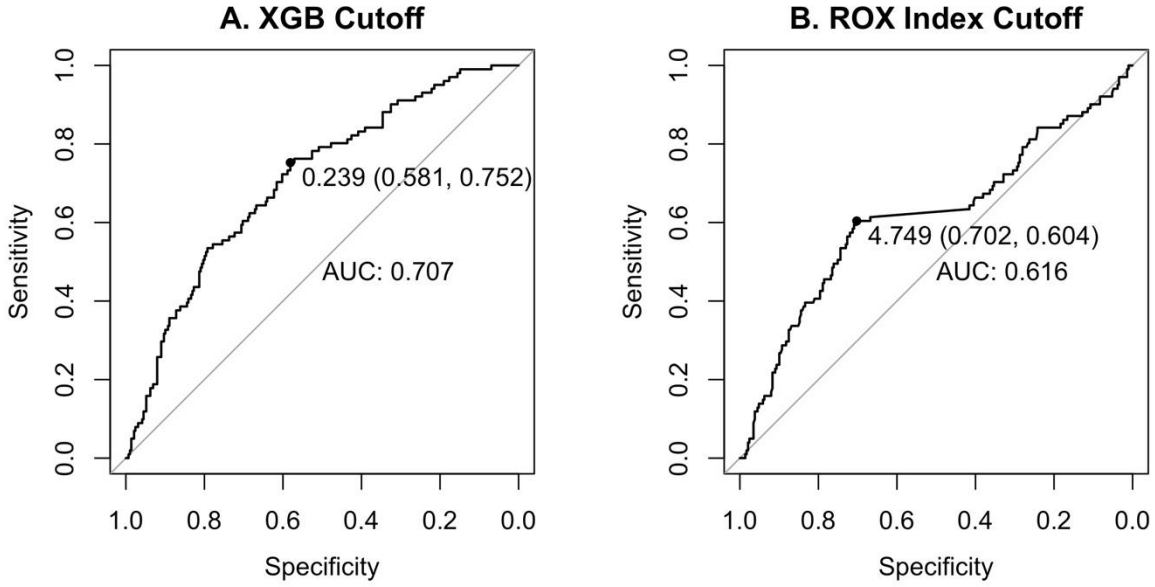
eTable 4. Expanded table of clinical characteristics compared between failure and non-failure patients.

Characteristic	Failure (N = 317)	Non-failure (N = 667)	p
Age, median (IQR)	65 (55-76)	61 (50-71)	<0.01
Sex, n (%)			
Female	147 (46.4%)	297 (44.5%)	0.63
Male	170 (53.6%)	370 (55.5%)	
Race, n (%)			
Asian	14 (4.4%)	29 (4.4%)	0.89
Black	163 (51.4%)	337 (50.5%)	
Other	32 (10.1%)	79 (11.8%)	
White	108 (34.1%)	222 (33.3%)	
Comorbidities, n (%)			
Congestive heart failure	63/312 (20.2%)	135/666 (20.3%)	1.00
HTN, uncomplicated	117/312 (37.5%)	253/666 (38.0%)	0.94
HTN, complicated	121/312 (38.8%)	186/666 (27.9%)	<0.01
Chronic pulmonary disease	71/312 (22.8%)	141/666 (21.2%)	0.62
Diabetes, uncomplicated	30/312 (9.6%)	66/666 (9.9%)	1.00
Diabetes, complicated	117/312 (37.5%)	189/666 (28.4%)	<0.01
Renal failure	95/312 (30.5%)	134/666 (20.1%)	<0.01
Liver disease	11/312 (3.5%)	33/666 (5.0%)	0.41
AIDS	3/312 (1.0%)	5/666 (0.8%)	0.72
Lymphoma	7/312 (2.2%)	9/666 (1.4%)	0.29
Metastatic cancer	7/312 (2.2%)	12/666 (1.8%)	0.63
Solid tumor without mets	9/312 (2.9%)	27/666 (4.1%)	0.47
Vital signs, mean (sd)			
Heart rate, max	103.71 ± 17.54	102.64 ± 18.48	0.38
SBP, min	113.80 ± 19.31	113.64 ± 18.82	0.90
DBP, min	62.88 ± 10.96	63.89 ± 11.64	0.19
MAP by cuff, min	82.53 ± 15.15	84.59 ± 12.57	0.04
Respiratory rate, max	31.08 ± 7.75	28.67 ± 10.67	<0.01
Pulse oximetry, min	85.11 ± 9.46	87.73 ± 8.32	<0.01
Temperature, max	37.77 ± 0.85	37.61 ± 0.82	0.01
Temperature, min	36.35 ± 0.67	36.37 ± 0.55	0.68
Laboratory values, mean (sd)			
BUN, max	32.46 ± 24.44	26.10 ± 19.89	<0.01
Creatinine, max	1.86 ± 2.03	1.50 ± 1.65	0.01
WBC, max	10.52 ± 10.93	10.02 ± 6.20	0.45
WBC, min	8.82 ± 9.38	8.34 ± 5.43	0.39
Platelets, min	204.27 ± 81.94	230.61 ± 88.60	<0.01
C-reactive protein, max	170.04 ± 90.59	151.50 ± 86.39	<0.01
D-dimer, max	5,586 ± 13,045	5,791 ± 13,902	0.83
Arterial blood gas in the first 24 hours, n; mean (sd)			
pH, max	289; 7.43 ± 0.06	496; 7.44 ± 0.05	0.12

pH, min	289; 7.39 ± 0.09	496; 7.41 ± 0.07	<0.01
PaO2, max	270; 100.46 ± 57.75	446; 94.55 ± 54.24	0.18
PaO2, min	270; 67.87 ± 29.00	446; 72.40 ± 27.82	0.04
PaCO2, max	270; 38.99 ± 11.41	447; 37.14 ± 8.84	0.02
PaCO2, min	270; 33.64 ± 7.58	447; 34.34 ± 7.67	0.24
Treatments received			
Bolus IV fluids, median (IQR)	750 (500-1,250)	750 (250-1,250)	0.11
Dexamethasone, n (%)	241 (76.0%)	539 (80.8%)	0.10
Remdesivir, n (%)	143 (45.1%)	362 (54.3%)	0.01
Norepinephrine, n (%)	64 (20.2%)	38 (5.7%)	<0.01
Vasopressin, n (%)	10 (3.2%)	12 (1.8%)	0.25
Outcomes			
NIV, n (%)	105 (33.1%)	99 (14.8%)	<0.01
IMV, n (%)	288 (90.9%)	116 (17.4%)	<0.01
Mortality, n (%)	141 (44.5%)	49 (7.4%)	<0.01
HFNC duration, median (IQR)	179 (62-379)	151 (79-249)	0.04
LOS, median (IQR)	409 (213-712)	221 (160-354)	<0.01
Worst ROX index in the first 24 hours, mean (sd)	5.54 ± 4.36	7.11 ± 4.80	<0.01
SOFA score at the time of HFNC initiation, mean (sd)	7.21 ± 2.69	4.27 ± 2.47	<0.01

Abbreviations: IQR = interquartile range, HTN = hypertension, AIDS = acquired immunodeficiency syndrome, SBP = systolic blood pressure, DBP = diastolic blood pressure, MAP = mean arterial pressure, BUN = blood urea nitrogen, WBC = white blood cell count, IV = intravenous, NIV = non-invasive ventilation, IMV = invasive mechanical ventilation, HFNC = high-flow nasal cannula, LOS = length of stay, SOFA = sequential organ failure assessment.

eFigure 1. Determining the optimal cut-offs in the eXtreme Gradient Boosting (XGB) model and the ROX index.



Receiver-operator characteristic (ROC) curves of A. eXtreme Gradient Boosting (XGB) model and B. ROX index. The point marked on the ROC curves indicate the optimal cut-off value, with specificity and sensitivity values at that cut-off in parentheses. Area under the ROC curve (AUC) are indicated separately.

eTable 5. eXtreme Gradient Boosting (XGB) and ROX index model performance metrics for predicting high-flow nasal cannula failure.

	Sensitivity (True positive rate)	**Specificity (True negative rate)**	**False positive rate**	**False negative rate**
XGB model-predicted probability of HFNC failure (Optimal cutoff = 0.239)				
Overall cohort	0.753	0.581	0.419	0.248
Black patients	0.723	0.568	0.432	0.277
White patients	0.828	0.582	0.418	0.172
ROX index score (Optimal cutoff = 4.749)				
Overall cohort	0.604	0.702	0.298	0.396
Black patients	0.615	0.710	0.290	0.385
White patients	0.655	0.704	0.296	0.345

eTable 6. Stratified analyses of predictive models for high-flow nasal cannula failure by race.

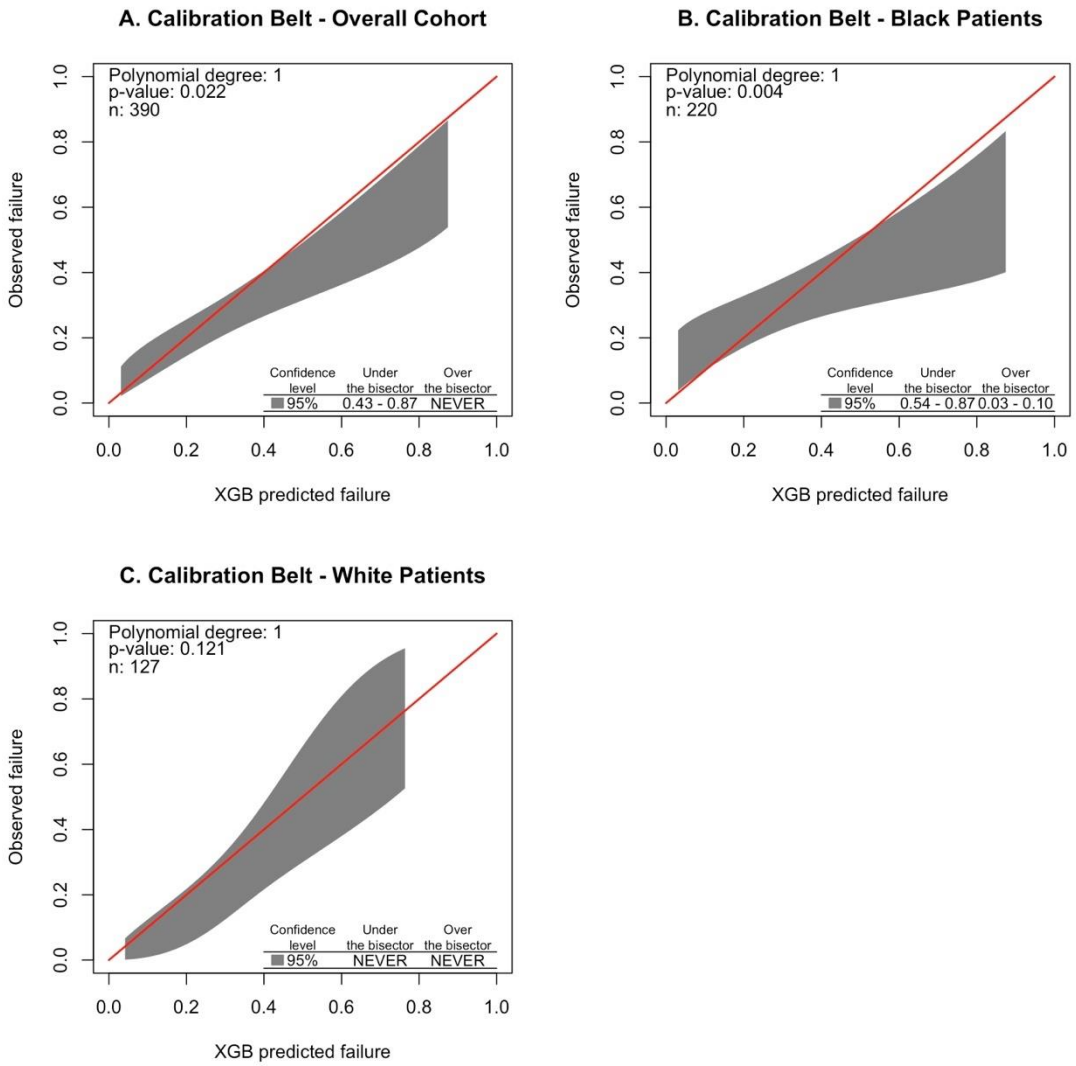
Model	AUROC (95%CI), overall cohort	AUROC (95%CI), Black patients	AUROC (95%CI), White patients	p-value comparing Black vs. White patients
XGB	0.707 (0.650-0.765)	0.663 (0.586-0.740)	0.808 (0.717-0.900)	0.02
LR	0.673 (0.612-0.735)	0.662 (0.583-0.742)	0.709 (0.601-0.817)	0.50
SVM	0.657 (0.597-0.717)	0.626 (0.546-0.705)	0.700 (0.593-0.806)	0.28
ROX index	0.616 (0.546-0.685)	0.613 (0.523-0.702)	0.691 (0.571-0.811)	0.31
KNN	0.526 (0.461-0.592)	0.509 (0.424-0.593)	0.572 (0.449-0.696)	0.40

Abbreviations: AUROC = area under the receiver-operator characteristic curve, CI = confidence interval, XGB = eXtreme Gradient Boosting, SVM = support vector machines, LR = logistic gression, KNN = k-nearest neighbor

eTable 7. Stratified analyses of the eXtreme Gradient Boosting (XGB) model for high-flow nasal cannula failure by sex and age group.

Stratified analyses	AUROC (95% CI)	p-value
By sex		
Female (n=186)	0.694 (0.610-0.778)	0.68
Male (n=204)	0.717 (0.639-0.795)	
By age group		
<65 years old (n=232)	0.690 (0.612-0.769)	0.43
≥65 years old (n=158)	0.739 (0.656-0.822)	

eFigure 2. Calibration belts for eXtreme Gradient Boosting (XGB) model predictions for high-flow nasal cannula failure.



In each panel, the red bisector line indicates the ideal, perfect correlation between prediction and observation, and the gray ribbon plots the model predicted risk of failure (on x-axis) against the actual observation (on y-axis). If the gray ribbon is under the bisector, then the predicted risk was higher than the observed value, indicating that the model overestimated failure; if the gray ribbon is over the bisector, then the predicted risk was lower than the observed value, indicating that the model underestimated failure. P-value <0.05 indicates miscalibration in the model. At the bottom right of each panel, the ranges of predicted risk of failure for which observed values deviated significantly from the bisector in each subgroup are reported. (A) In the overall cohort, model overestimated failure (i.e. gray ribbon was under the bisector) for patients whose predicted risk of failure was 43-87%. (B) In Black patients, model overestimated failure (i.e. gray ribbon was under the bisector) for patients with predicted risk of failure between 54-87%, and underestimated failure (i.e. gray ribbon was over the bisector) for those with predicted risk of failure between 3-10%. (C) In White patients, the model neither over- nor under-estimate the risk of failure.

eTable 8. Expanded table of clinical characteristics compared between Black and White patients.

Characteristic	Black patients (N = 500)	White patients (N = 330)	p
Age, median (IQR)	61 (51-71)	67 (57-76)	<0.01
Sex, n (%)			
Female	261 (52.2%)	127 (38.5%)	<0.01
Male	239 (47.8%)	203 (61.5%)	
Comorbidities, n (%)			
Congestive heart failure	116/496 (23.4%)	67/329 (20.4%)	0.35
HTN, uncomplicated	190/496 (38.3%)	122/329 (37.1%)	0.77
HTN, complicated	194/496 (39.1%)	88/329 (26.8%)	<0.01
Chronic pulmonary disease	120/496 (24.2%)	71/329 (21.6%)	0.40
Diabetes, uncomplicated	57/496 (11.5%)	27/329 (8.2%)	0.16
Diabetes, complicated	189/496 (38.1%)	67/329 (20.4%)	<0.01
Renal failure	154/496 (31.1%)	57/329 (17.3%)	<0.01
Liver disease	19/496 (3.8%)	13/329 (4.0%)	1.00
AIDS	6/496 (1.2%)	2/329 (0.6%)	0.49
Lymphoma	8/496 (1.6%)	5/329 (1.5%)	1.00
Metastatic cancer	8/496 (1.6%)	10/329 (3.0%)	0.22
Solid tumor without mets	17/496 (3.4%)	18/329 (5.5%)	0.16
Vital signs, mean (sd)			
Heart rate, max	104.82 ± 18.56	99.83 ± 17.54	<0.01
SBP, min	114.80 ± 19.99	112.45 ± 18.57	0.09
DBP, min	64.55 ± 11.94	62.44 ± 10.93	<0.01
MAP by cuff, min	85.75 ± 14.04	82.24 ± 13.26	<0.01
Respiratory rate, max	30.02 ± 11.60	28.12 ± 7.56	<0.01
Pulse oximetry, min	87.74 ± 8.36	86.15 ± 9.24	0.01
Temperature, max	37.61 ± 0.79	37.65 ± 0.86	0.44
Temperature, min	36.34 ± 0.63	36.33 ± 0.57	0.84
Laboratory values, mean (sd)			
BUN, max	29.93 ± 24.02	27.95 ± 17.41	0.17
Creatinine, max	1.94 ± 2.20	1.33 ± 1.11	<0.01
WBC, max	10.27 ± 9.97	10.23 ± 5.89	0.95
WBC, min	8.68 ± 8.76	8.36 ± 4.73	0.50
Platelets, min	230.29 ± 86.85	204.69 ± 82.16	<0.01
C-reactive protein, max	156.49 ± 81.07	151.58 ± 90.10	0.46
D-dimer, max	7,316 ± 16,235	3,488 ± 7,662	<0.01
Arterial blood gas in the first 24 hours, n; mean (sd)			
pH, max	394; 7.43 ± 0.06	263; 7.45 ± 0.06	<0.01
pH, min	394; 7.39 ± 0.08	263; 7.41 ± 0.08	<0.01
PaO2, max	366; 99.74 ± 56.97	231; 91.60 ± 51.34	0.07
PaO2, min	366; 71.97 ± 29.21	231; 69.70 ± 28.37	0.35
PaCO2, max	366; 39.24 ± 10.85	231; 36.88 ± 9.55	0.01
PaCO2, min	366; 35.27 ± 7.65	231; 33.43 ± 8.43	0.01

Missing ABG data, n (%)			
pH, min/max	106 (21.2%)	67 (20.3%)	0.82
PaO2, min/max	134 (26.8%)	99 (30.0%)	0.35
PaCO2, min/max	134 (26.8%)	99 (30.0%)	0.35
Treatments received			
Bolus IV fluids, median (IQR)	750 (500-1,250)	750 (375-1,250)	0.44
Dexamethasone, n (%)	385 (77.0%)	273 (82.7%)	0.06
Remdesivir, n (%)	238 (47.6%)	192 (58.2%)	<0.01
Norepinephrine, n (%)	57 (11.4%)	32 (9.7%)	0.51
Vasopressin, n (%)	12 (2.4%)	7 (2.1%)	1.00
Outcomes			
Failure, n (%)	163 (32.6%)	108 (32.7%)	1.00
NIV, n (%)	127 (25.4%)	65 (19.7%)	0.07
IMV, n (%)	221 (44.2%)	126 (38.2%)	0.10
Mortality, n (%)	86 (17.2%)	77 (23.3%)	0.04
Time to HFNC, median (IQR)	2.9 (0.6-9.5)	4.4 (1.1-12.3)	<0.01
HFNC duration, median (IQR)	147 (67-282)	160 (81-268)	0.63
LOS, median (IQR)	258 (166-503)	255 (161-411)	0.11
Worst ROX index in the first 24 hours, mean (sd)	6.35 ± 4.48	6.39 ± 4.85	0.92
SOFA score at the time of HFNC initiation, mean (sd)	5.41 ± 2.97	5.08 ± 2.75	0.10

Abbreviations: IQR = interquartile range, HTN = hypertension, AIDS = acquired immunodeficiency syndrome, SBP = systolic blood pressure, DBP = diastolic blood pressure, MAP = mean arterial pressure, BUN = blood urea nitrogen, WBC = white blood cell count, PaO2 = partial pressure arterial oxygen, PaCO2 = partial pressure arterial carbon dioxide, ABG = arterial blood gas, IV = intravenous, NIV = non-invasive ventilation, IMV = invasive mechanical ventilation, HFNC = high-flow nasal cannula, LOS = length of stay, SOFA = sequential organ failure assessment.

eTable 9. Sensitivity analyses.

Sensitivity analysis	AUROC, overall cohort	AUROC, Black patients	AUROC, White patients	Absolute difference in AUROC between Black and White patients	p-value comparing AUROC between Black and White patients
XGB model – modifying model parameters					
Original XGB model validation	0.707	0.663	0.808	0.145	0.02
Exclude SpO2 variables, predicting HFNC failure	0.690	0.638	0.794	0.155	0.01
Exclude SpO2 and RR variables, predicting HFNC failure	0.656	0.617	0.726	0.109	0.08
Exclude all vital sign variables, predicting HFNC failure	0.626	0.581	0.715	0.134	0.03
Include all variables from original XGB model, *predicting mortality*	0.760	0.784	0.749	0.035	0.67
XGB model – modifying strategies for handling missing data					
Missing variables *left as missing*	0.724	0.673	0.833	0.160	<0.01
Original XGB model validation with *median imputation*	0.707	0.663	0.808	0.145	0.02
Predictive mean matching imputation	0.696	0.659	0.779	0.120	0.05
Exclude patients with missing ABG data (n=223)	0.654	0.629	0.669	0.040	0.65
ROX index					
Original ROX index validation w/ *median imputation*	0.616	0.613	0.691	0.079	0.31
Predictive mean matching imputation for ROX index	0.632	0.610	0.745	0.135	0.04
Exclude patients with missing ROX index, predicting HFNC failure (n=318)	0.614	0.605	0.705	0.100	0.18

Abbreviations: AUROC = area under receiver-operator characteristic curve, XGB = eXtreme Gradient Boosting, SpO2 = oxygen saturation, HFNC = high-flow nasal cannula, RR = respiratory rate.

eTable 10. Stratified analyses of predictive models for high-flow nasal cannula failure by race, with high-flow nasal cannula failure defined only by intubation

Model	AUROC (95%CI), overall cohort	AUROC (95%CI), Black patients	AUROC (95%CI), White patients	p-value comparing Black vs. White patients
XGB	0.700 (0.640-0.760)	0.667 (0.587-0.747)	0.767 (0.673-0.862)	0.11
LR	0.657 (0.594-0.721)	0.646 (0.564-0.727)	0.668 (0.553-0.784)	0.75
SVM	0.635 (0.572-0.698)	0.607 (0.526-0.688)	0.660 (0.542-0.778)	0.48
ROX index	0.607 (0.536-0.678)	0.620 (0.530-0.710)	0.656 (0.528-0.784)	0.66
KNN	0.534 (0.467-0.601)	0.520 (0.435-0.605)	0.561 (0.433-0.690)	0.59

Abbreviations: AUROC = area under the receiver-operator characteristic curve, CI = confidence interval, XGB = eXtreme Gradient Boosting, SVM = support vector machines, LR = logistic gression, KNN = k-nearest neighbor