



medicina *intensiva*

<http://www.medintensiva.org/en/>



ORIGINAL ARTICLE

The value of local validation of a predictive model. A nomogram for predicting failure of non-invasive ventilation in patients with SARS-COV-2 pneumonia

Héctor Hernández Garcés^{a,*}, Alberto Belenguer Muncharaz^a,
Francisco Bernal Julián^b, Irina Herмосilla Semikina^b, Luis Tormo Rodríguez^b,
Estefanía Granero Gasamans^b, Clara Viana Marco^b, Rafael Zaragoza Crespo^a

^a Intensive Care Unit, Hospital Universitario Doctor Peset, Av Gaspar Aguilar 90, 46017 Valencia, Spain

^b Intensive Care Unit, Hospital General Universitario de Castellón, Av Benicassim 128, Castellón, Spain

Received 17 September 2024; accepted 10 November 2024

KEYWORDS

Acute respiratory failure;
ARDS;
COVID-19;
Orotracheal intubation

Abstract

Objective: We aimed to determine predictors of non-invasive ventilation (NIV) failure and validate a nomogram to identify patients at risk of NIV failure.

Design: Observational, analytical study of a retrospective cohort from a single center, compared with an external cohort (March 2020 to August 2021).

Setting: Two intensive care units (ICUs).

Patients: Patients with pneumonia due to severe acute respiratory syndrome (SARS-CoV-2) and NIV > 24 h (154 and 229 in each cohort).

Interventions: The training cohort identified NIV failure predictors. A nomogram, created via logistic regression, underwent validation with the Hosmer-Lemeshow (HL), calibration curve and test and area under the curve (AUC). Its external validity was tested using AUC.

Main variables of interest: Demographics, comorbidities, severity scores, NIV settings, vital signs, blood gases, and oxygenation at the start and 24 h after NIV, NIV failure.

Results: NIV failure was 37.6% and 18% in the training and validation cohorts, respectively. Risk factors for NIV failure included age, obesity, sequential organ failure assessment (SOFA) score at admission, and heart rate (HR) and heart rate, acidosis, consciousness, oxygenation, respiratory rate (HACOR) 24 h post-NIV. The model's HL test result was 0.861, with an AUC of 0.89 (confidence interval [CI] 0.839–0.942); validation AUC was 0.547 (CI 0.449–0.645).

Conclusions: A predictive model using age, obesity, SOFA score, HR, and HACOR at 24 h predicts NIV failure in our COVID-19 patients but may not apply to other ICUs.

© 2025 Published by Elsevier España, S.L.U.

DOI of original article: <https://doi.org/10.1016/j.medin.2025.502148>

* Corresponding author.

E-mail address: hektorhernandez84@gmail.com (H. Hernández Garcés).

<https://doi.org/10.1016/j.medicine.2025.502148>

2173-5727/© 2025 Published by Elsevier España, S.L.U.

Please cite this article as: H. Hernández Garcés, A. Belenguer Muncharaz, F. Bernal Julián et al., The value of local validation of a predictive model. A nomogram for predicting failure of non-invasive ventilation in patients with SARS-COV-2 pneumonia, *Medicina Intensiva*, <https://doi.org/10.1016/j.medicine.2025.502148>

PALABRAS CLAVE

Fracaso respiratorio agudo;
SDRA;
COVID-19;
Intubación orotraqueal

El valor de la validación local de un modelo predictivo. Nomograma para la predicción del fracaso de ventilación no invasiva en pacientes con neumonía por SARS-COV-2

Resumen

Objetivo: Determinar los predictores de fracaso de ventilación no invasiva (VNI) y validar un nomograma para identificar el riesgo de fracaso de VNI.

Diseño: Estudio observacional, analítico de una cohorte retrospectiva de un centro, comparada con una cohorte externa (marzo 2020 a agosto 2021).

Ámbito: Dos unidades de cuidados intensivos (UCI).

Pacientes: Pacientes con neumonía por síndrome respiratorio agudo grave (SARS-COV-2) y VNI>24 h (154 y 229 en cada cohorte).

Intervenciones: Regresión logística para la detección de factores de riesgo de fracaso de VNI en una cohorte de entrenamiento, y elaboración de nomograma para identificar el riesgo de fracaso de VNI. Validación mediante el test de Hosmer-Lemeshow (HL), curva de calibración y área bajo la curva (AUC). Validación externa mediante el AUC.

Variables de interés principales: Demográficas, comorbilidades, scores de severidad, configuración de VNI, constantes vitales, gasométricas y oxigenación al inicio 24 h de VNI, fracaso de VNI.

Resultados: El fracaso de VNI fue del 37,6% y 18% en la cohorte de entrenamiento y validación respectivamente. Los factores relacionados con el fracaso de VNI fueron edad, obesidad, sequential organ failure assessment (SOFA) al ingreso y, frecuencia cardiaca (FC) y acidosis, consciencia, oxigenación y frecuencia respiratoria (HACOR) 24 h post-VNI. El test HL del modelo resultó de 0.861, con AUC 0.89 (intervalo confianza [IC] 0.839–0.942). AUC de validación externa 0.547 (CI 0.449–0.645)

Conclusiones: Un modelo predictivo utilizando edad, obesidad, SOFA score y, FC y HACOR a las 24 h predice el fallo de VNI en nuestros pacientes con COVID-19, pero podría no aplicarse a otras UCIs.

© 2025 Publicado por Elsevier España, S.L.U.

Introduction

Coronavirus disease 2019 (COVID-19) is caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). More than 670 million people suffer from this disease worldwide, of whom nearly seven million are expected to die.¹ Intensive care units (ICUs) have been used to manage patients with severe forms of the disease. Initial data from China suggested that between 5% and 32% of infected patients required admission to the ICU,^{2,3} and 95% received ventilatory support.⁴

In cases of severe SARS-CoV-2 pneumonia (CAP), the use of non-invasive ventilation (NIV) was initially discouraged in various recommendations.^{5–8} In contrast, the National Health Service (NHS) of the United Kingdom proposed the use of NIV with continuous positive airway pressure (CPAP) as the first mode of ventilatory support for patients with COVID-19 pneumonia.⁹ Despite this controversy, several studies on patients with hypoxemia have reported a reduction in orotracheal intubation (OTI) after NIV. These results support the use of NIV during the pandemic.^{10,11} However, late OTI after NIV failure is associated with increased mortality.¹² Therefore, it seems reasonable to identify the correct time at which patients would benefit from safe intubation. However, the parameters to identify patients with high risk of NIV failure is unknown.

Predictive modeling of NIV failure would help in decision-making. This would help avoid excessive exposure to NIV

and delay in OTI and its prognostic consequences. The main objective of this study was to develop and validate a predictive model for the risk of NIV failure in patients with severe SARS-CoV-2 CAP admitted to the ICU.

Patients and methods

Study design and ethical aspects

We conducted an observational, analytical, and retrospective study in the ICU of two hospitals in Spain (Hospital Universitario Doctor Peset de Valencia and Hospital General Universitario de Castellon) between March 2020 and August 2021. This study was approved by the Ethics Committee of Hospital Universitario Doctor Peset of Valencia (code 66/20). Given the characteristics of this study, the requirement for informed consent was waived.

Patients

Patients with a confirmed final diagnosis of COVID-19 and CAP who received NIV for at least 24 h in the ICU were included. A confirmed case was considered following a positive result in the real-time polymerase chain reaction test for coronavirus in a sample obtained from the nasopharyngeal exudate, tracheal aspirate, and/or bronchoaspirate.

CAP was defined as the presence of fever $>38^{\circ}\text{C}$, cough, dyspnea or tachypnea (respiratory rate [RR] >30 bpm), and hypoxemia (transcutaneous oxygen saturation [SatO₂] $<90\%$ or partial pressure of oxygen/inspired fraction of oxygen ratio [PaO₂/FiO₂] <300) and/or chest radiographic infiltrate.¹³ The exclusion criteria were as follows: 1) do-not intubate order, 2) diagnosis at the end of admission different from CAP COVID-19; 3) invasive mechanical ventilation (IMV) or high-flow nasal oxygen therapy (HFNC) as first support, and 4) duration of NIV of <24 h.

At admission, patients were considered for OTI and direct IMV if they were admitted with a low Glasgow coma scale score (GCS) <8 , significant respiratory distress, and episodes of apnea or extreme desaturation. For those evaluated for NIV initiation, the first-choice ventilators were the specific BiPAP V60[®] NIV ventilators (Respironics Inc, Pennsylvania, USA), and the CPAP level considered for the initiation of therapy was 10–15 cmH₂O. FiO₂ was adjusted to achieve a minimum SatO₂ of 92–94%. NIV was administered continuously for the first 24 h, with minimal interruption of oral medication and fluid intake, and was considered successful if OTI was avoided and the patient was discharged to the hospital ward alive and did not require NIV again. Gold-nasal interfaces were used for the most part and were replaced by others depending on the evolution, type of respiratory failure, appearance of skin lesions, and intolerance. If the patient achieved clinical and/or gasometric improvement that allowed FiO₂ and CPAP levels lower than 0.5 and 10 cmH₂O, the ventilator was withdrawn alternating disconnections from the HFNC, which was progressively increased in duration.

The predictive model was developed for a cohort of patients admitted to Hospital Universitario Doctor Peset, whereas the cohort of patients from Hospital General Universitario de Castellón was used for external validation.

Variables collected

Demographic data, comorbidities, setting, origin, previous chronology, previous respiratory support, severity scores, such as the Sequential Organ Failure Assessment (SOFA) score at admission, and initial NIV configuration were collected. NIV parameters, vital signs, gasometric and oxygenation data were recorded at the start of NIV and 24 h after initiation. The respiratory rate-oxygenation index (iROX) and heart rate, acidosis, conscientiousness, oxygenation, and respiratory rate (HACOR) scores were also calculated during this period.

NIV failure was defined as the need to initiate IMV. The OTI decision was left to the discretion of the attending physician according to the clinical and gasometric criteria.

Statistical analysis

For statistical analysis, the normality of the quantitative variables was assessed using the Kolmogorov–Smirnov test. They were expressed as mean \pm standard deviation or median with interquartile range, and their comparison was performed with Student's t-test or Mann–Whitney U test. Qualitative variables were expressed as frequencies (per-

centages) and compared using the chi-square test or Fisher's exact test. A predictive model was constructed using a multivariate binary logistic regression analysis. Initially, univariate analysis was performed on the training cohort to detect variables that had a relevant clinical association with NIV failure. Variables with $p < 0.1$ were finally included in the multivariate logistic regression model. Variables without statistical significance were manually eliminated by using a stepwise backward procedure until the best possible model was obtained. The relationship was expressed as odds ratio (OR) and 95% confidence interval (95% CI) and was calculated by one unit increment of each continuous variable and by the presence of a factor corresponding to a qualitative variable. The final model was expressed by a predictive equation: $f(x) = 1/1 + \exp(\beta_1 \times 1 + \beta_2 \times 2 \dots)$, where β is the regression coefficient and X is the value of the variable that showed association in the logistic regression. The model was evaluated using the Hosmer–Lemeshow goodness-of-fit test. To analyse the agreement between model predictions and real observations in the training cohort, bootstrap of 1000 resamples were set and calibration curve was created. Predictive capacity was evaluated using a receiver operating curve (ROC). Subsequently, a nomogram was constructed for decision support, using the training cohort as a reference, including the total score and percentage risk of failure. For external validation, an ROC curve was constructed by calculating the area under the curve (AUC) and 95% CI, taking as the state variable the actual NIV failure of the validation cohort and, as a contrast variable, the probability of NIV failure at 24 h, calculated using our nomogram for each of the patients in the validation cohort. The SPSS[®] 20.0.0 software package (Chicago, IL, USA) and RStudio software (version 1.2.5019) were used, and a p-value < 0.05 was considered statistically significant.

Results

A total of 154 patients were included in the training cohort during the study period. Fig. 1 shows the flowchart of these patients. The NIV failure rate in the study population was 37.6% (N = 58).

In the training cohort, the failure group presented with higher age, SOFA score, percentage of patients with obesity, days from symptom onset to ICU admission, and days from hospital admission to ICU admission (Table 1). In the NIV setting, no differences were observed between interfaces (Table 2). Upon ICU admission, the failure group required a higher FiO₂ and presented more deteriorated SatO₂/FiO₂, PaO₂/FiO₂, and iROX values than did the success group (Table 2); moreover, at 24 h, they presented with a higher heart rate (HR), greater need for FiO₂, and lower values of SatO₂, SatO₂/FiO₂, PaO₂/FiO₂, iROX, and HACOR scores (Table 2).

In total, 229 patients who met the inclusion criteria during the study period were included in the validation cohort. NIV failure occurred in 42(18%) patients. A comparison of the two cohorts revealed that the validation cohort had a lower SOFA score, shorter time from hospital admission to the initiation of NIV, less use of non-invasive respiratory therapy (NIRT) prior to ICU admission, and a higher pro-

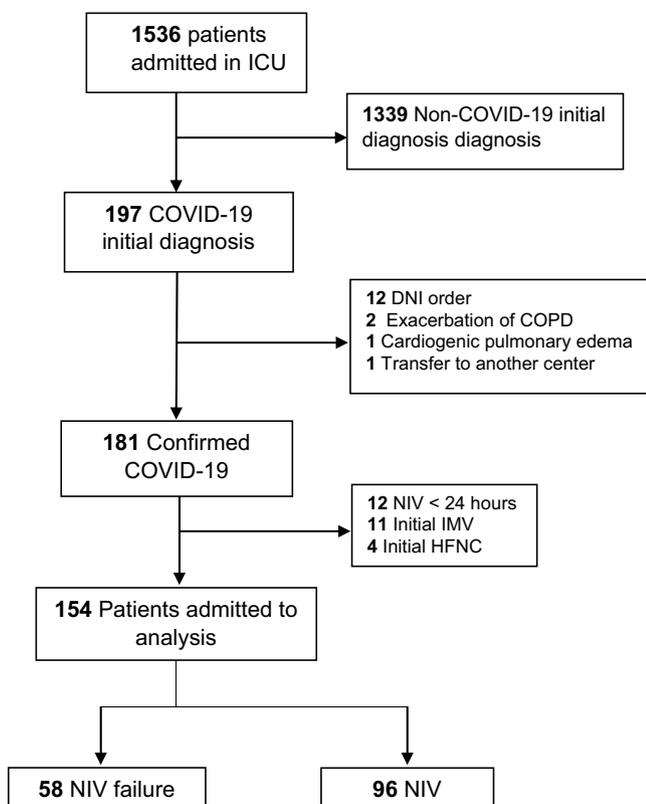


Figure 1 Flowchart of inclusion and exclusion of participants. COPD: chronic obstructive pulmonary disease; DNI: do-not intubate; HFNC: high flow nasal cannula; ICU: intensive care unit; IMV: invasive mechanical ventilation; NIV: non-invasive ventilation.

portion of patients admitted to the emergency department (Table 1). Critical ventilators and helmets were frequently used (Table 2). Regarding oxygenation status, it should be noted that patients in the validation cohort were less hypoxemic, as measured by SatO₂, SatO₂/FiO₂, PaO₂/FiO₂, iROX, and HACOR both on admission and at 24 h (Table 2). ICU mortality in the training and validation cohorts occurred in 39(25%) and 25(11%) patients, respectively, and was higher in the NIV failure group than in the NIV success group in both cohorts ([68% vs. 0%; p < 0.0001] and [60% vs. 0%; p < 0.0001, respectively]) (Table 2).

Multivariate logistic regression analysis revealed that the factors independently associated with NIV failure were higher age, high SOFA score at admission, obesity, elevated HR at 24 h, and high HACOR score at 24 h (Appendix A; Table A1). The model explained 58.7% (Nagelkerke R-squared) of the variation in failure outcomes. A goodness-of-fit test was performed to compare the observed and expected data with the model, demonstrating that there were no differences, with a Hosmer–Lemeshow test result of chi-square = 1.97 and p = 0.961. Calibration curve shows a good fit between the actual and predicted probability shows the curve calibration of our predictive model (Appendix A, Fig. A2).

Based on the final model, a nomogram was constructed using the weights of each model variable associated with NIV

failure (Fig. 2). The probability of NIV failure was calculated using equation (1):

$$(1)y = 1/(1 + e^{-f(x)})$$

$$F(x) = -11.855 + (0.041 \times \text{age}) + (0.747 \times \text{obesity}) \\ + (0.708 \times \text{SOFA}) + (0.038 \times \text{HR24 h}) \\ + (0.653 \times \text{HACOR24 h}).$$

The Supplementary Material (Appendix A, Fig. A1) describes how the nomogram can be used to determine the probability of NIV failure.

The predictive ability of the model was satisfactory in the training cohort, with an AUC of 0.89 (95% CI: 0.839–0.942, p < 0.0001) (Fig. 3A), which was superior to that of each of the model variables separately (Table 3). However, the AUC of the validation cohort was 0.547 (0.449–0.645, p = 0.346) (Fig. 3B).

Discussion

In this study, a nomogram was developed to predict the failure of NIV in patients with COVID-19. It is based on age, obesity, SOFA score at admission, and HR and HACOR scores at 24 h. Patients who failed NIV had higher mortality; therefore, it is important to recognize an early failure of NIV and select those who would benefit from escalating ventilatory support. To simplify and quickly apply the model, analytical variables were not included. The predictive capacity of the model was assessed with the Hosmer–Lemeshow test, the calibration curve and by calculating the AUC on an ROC curve, which reinforces the internal strength of our model.

The training cohort presented an NIV failure rate of 37.6%, which is within the range of 29%–53% reported in different studies.^{14–16} However, the rate of failure in the validation cohort was 18%, which was lower than that reported for patients with COVID-19 in the ICU. This difference could be attributed to the lower severity level in the validation cohort and patients being admitted at earlier disease stages, thus requiring less prior NIRT. Moreover, majority of patients in the validation cohort were administered NIV via a Helmet, an interface linked to lower NIV failure and mortality rates compared with that of the oronasal mask.¹⁷

A retrospective observational study conducted across 23 Chinese hospitals, including non-ICU patients treated with NIV or HFNC, developed a predictive model for NIRT failure risk.¹⁸ This model incorporated age, comorbidity count, oxygenation index, the GCS score, and vasopressor usage; however, unlike ours, it only included variables at ICU admission. The value of our model lies in its utility for patients who have completed 24 h of support with NIV, suggesting an evaluation after a longer trial period than initially recommended in the early stages of the pandemic.⁵ We believe that this aspect fills a gap in the available literature.

Age determines the prognosis of these patients¹⁹ and is a risk factor for NIV failure in both patients with COVID-19^{15,20–22} and acute respiratory distress syndrome (ARDS).^{23,24}

Table 1 Characteristics of patients with successful or failed NIV and group differences.

	Training cohort				Validation cohort				
	Overall (n = 154)	NIV success (n = 96)	NIV failure (n = 58)	p value ¹	Overall (n = 229)	NIV success (n = 187)	NIV failure (n = 42)	p value ¹	p value ²
Age, years	63 (52–70)	60 (49–67)	65 (60–71)	0,001	61 (52–67)	60 (51–66)	67 (58–70)	0,004	0,175
Female gender	43 (28)	25 (26)	18 (31)	0,447	58 (25)	52 (28)	6 (14)	0,069	0,468
SOFA score	4 (3–4)	3 (2–4)	4 (3–5)	<0,0001	3 (3–4)	3 (3–4)	3 (3–4)	0,311	<0,0001
Hypertension	66 (43)	37 (38)	29 (50)	0,141	110 (48)	84 (45)	26 (62)	0,046	0,297
Diabetes	35 (23)	20 (21)	15 (26)	0,25	50 (22)	37 (20)	13 (31)	0,113	0,791
Heart disease	11 (7)	4 (4)	7 (12)	0,136	16 (7)	11 (6)	5 (12)	0,168	0,94
COPD	10 (6)	6 (6)	4 (7)	0,886	12 (5)	8 (4)	4 (9)	0,167	0,443
Obesity	72 (47)	40 (42)	32 (55)	0,049	96 (42)	77 (41)	19 (45)	0,63	0,614
Duration of symptom onset to ICU admission, days	9 (7–11)	10 (7–12)	8 (6–10)	0,0004	9 (7–11)	9 (7–11)	7 (5–9)	0,008	0,847
Duration of hospital admission to ICU admission, days	2 (0–4)	2 (0–5)	1 (0–3)	0,025	1 (0–3)	2 (0–3)	1 (0–2)	0,111	0,166
Duration of hospital admission to NIV onset, hours	51 (22–116)	59 (22–124)	43 (21–97)	0,183	47 (18–85)	48 (22–89)	26 (0–63)	0,061	0,027
NIRS pre-ICU admission Source	107 (69)	64 (67)	43 (74)	0,838	30 (13)	19 (10)	11 (26)	0,005	<0,0001
Emergency area	16 (10)	6 (6)	10 (14)	0,078	43 (19)	33 (18)	10 (24)	0,165	0,023
Hospital ward	150 (97)	90 (94)	60 (86)	0,126	165 (72)	140 (75)	25 (59)	0,083	0,008
Other hospital	0 (0)	0 (0)	0 (0)	–	21 (9)	14 (7)	7 (17)	0,062	0,018

Data summarised as number (%) or median (interquartile range).

COPD: chronic obstructive pulmonary disease; ICU: intensive care unit; NIRS: non-invasive respiratory support; NIV: non-invasive ventilation; SOFA score: sequential organ failure assessment score.

¹ p value for difference between NIV failure vs succes group.

² p value for difference training vs validation group.

Table 2 Vital signs, arterial blood gas parameters, treatment variables at baseline and 24 h after NIV initiation, and main variables in patients with successful or failed NIV in both cohorts.

	Training cohort				Validation cohort				
	Overall (n = 154)	NIV success (n = 96)	NIV failure (n = 58)	p value ¹	Overall (n = 229)	NIV success (n = 187)	NIV failure (n = 42)	P value ¹	p value ²
NIV initiation									
Type of ventilator									
Critic ventilator	2 (1)	0 (0)	2 (3)	0,129	125 (55)	106 (57)	19 (45)	,178	<0,0001
Specific NIV ventilator	152 (99)	96 (100)	56 (97)	0,798	104 (45)	81 (43)	23 (55)	,073	<0,0001
Interface									
Helmet	4 (3)	2 (2)	2 (3)	0,570	223 (97)	183 (98)	40 (95)	,471	<0,0001
Orofacial mask	112 (73)	68 (71)	44 (79)	0,110	4 (2)	2 (1)	2 (5)	,04	<0,0001
Total face mask	37 (24)	26 (27)	11 (18)	0,512	2 (1)	2 (1)	0 (0)	,540	<0,0001
PEEP, cm H2O	12 (11–14)	12 (12–14)	13 (12–15)	0,095	10 (10–12)	10 (10–12)	12 (10–13)	,289	<0,0001
FiO2	0,75 (0,6–1)	0,7 (0,6–0,8)	0,8 (0,7–1)	0,001	0,6 (0,5–0,7)	0,6 (0,5–0,7)	0,6 (0,5–0,65)	,518	<0,0001
HR, beats per min	86 ± 17	85 (75–97)	88 (79–98)	0,122	81 ± 17	81 ± 18	80 ± 15	,838	0,003
SP, mmHg	135 ± 18	129 ± 18	131 ± 21	0,626	135 ± 20	135 ± 19	132 ± 21	,435	0,848
RR, breaths per min	30 (24–35)	31 (26–36)	30 (25–33)	0,490	26 (22–30)	26 (22–30)	25 (23–30)	,538	<0,0001
pH	7,46 (7,43–7,48)	7,45 (7,43–7,48)	7,45 (7,41–7,48)	0,347	7,46 (7,42–7,5)	7,47 (7,43–7,5)	7,45 (7,39–7,51)	,324	0,268
PaCO2, mmHg	36 (33–39)	35 (33–39)	36 (33–39)	0,975	37 (32–41)	37 (32–41)	36 (34–41)	,741	0,104
SatO2, %	94 (91–96)	94 (91–96)	93 (90–96)	0,275	96 (93–98)	96 (93–97)	94 (89–98)	,131	0,006
SatO2/FiO2 ratio	130 ± 34	133 ± 23	116 ± 27	0,001	165 ± 35	166 ± 35	159 ± 34	,291	<0,0001
PaO2/FiO2 ratio, mmHg	95 (69–118)	101 (79–121)	84 (64–104)	<0,0001	125 (99–156)	124 (99–152)	129 (97–196)	,605	<0,0001
iROX	4,1 (3,3–5,5)	4,4 (3,5–5,4)	3,7 (2,9–5,5)	0,027	6,2 (4,9–7,6)	6,3 (4,9–7,8)	6,28 (5–7,4)	,722	<0,0001
HACOR	6 (5–7)	6 (5–7)	6 (5–6)	0,196	5 (4–6)	5 (4–6)	5 (4–6)	,367	<0,0001
Data 24 h of NIV									
FiO2	0,6 (0,5–0,7)	0,6 (0,5–0,7)	0,75 (0,55–0,9)	<0,0001	0,6 (0,5–0,7)	0,6 (0,5–0,7)	0,6 (0,5–0,65)	,758	0,075
HR, beats per min	74 ± 15	70 ± 9	80 ± 7	<0,0001	69 ± 14	69 ± 15	70 ± 12	,708	0,002
SP, mmHg	130 ± 16	136 ± 25	134 ± 20	0,889	133 ± 20	133 ± 21	133 ± 15	,995	0,098
RR, breaths per min	23 (20–27)	24 (20–28)	25 (21–28)	0,316	23 (19–26)	23 (19–25)	24 (20–27)	,121	0,474
pH	7,45 (7,42–7,47)	7,45 (7,43–7,47)	7,43 (7,42–7,45)	0,13	7,44 (7,44–7,47)	7,44 (7,41–7,47)	7,44 (7,38–7,47)	,882	0,025
PaCO2, mmHg	38 (35–42)	38 (35–41)	39 (35–43)	0,138	40 (36–43)	40 (36–43)	40 (34–44)	,638	0,012
SatO2, %	96 (94–97)	96 (95–98)	94 (92–96)	<0,0001	98 (96–99)	98 (97–99)	97 (95–99)	,013	<0,0001

Table 2 (Continued)

	Training cohort				Validation cohort				
	Overall (n = 154)	NIV success (n = 96)	NIV failure (n = 58)	p value ¹	Overall (n = 229)	NIV success (n = 187)	NIV failure (n = 42)	P value ¹	p value ²
NIV initiation									
SatO ₂ /FiO ₂ ratio	158 (132–190)	163 (140–196)	128 (100–170)	<0,0001	163 (142–196)	164 (141–196)	163 (149–194)	,580	0,004
PaO ₂ /FiO ₂ ratio, mmHg	131 (108–170)	151 (122–196)	104 (77–132)	<0,0001	154 (128–196)	155 (130–200)	150 (113–184)	,136	<0,0001
iROX	4 (3–6)	7,9 (6,1–9,8)	5,3 (4,3–8,7)	<0,0001	7,42742 (5,8–9,7)	7,5 (5,9–9,8)	7,1 (5–8,9)	,093	0,011
HACOR	4 (3–6)	4 (2–5)	6 (4–6)	<0,0001	4 (2–5)	4 (2–5)	4 (2–5)	,528	0,006
NIV and main outcomes									
Duration of NIV, hours	96 (59–140)	103 (60–157)	94 (54–115)	0,225	96 (55–144)	96 (54–144)	99 (66–162)	0,211	0,972
Causes of NIV failure									
Inability to correct	–	–	43 (74)	–	–	–	33 (79)	–	0,362
Hypoxemia	–	–	9 (16)	–	–	–	6 (14)	–	0,083
Intolerance	–	–	3 (5)	–	–	–	2 (5)	–	0,159
Shock	–	–	3 (5)	–	–	–	1 (2)	–	0,057
Duration of IMV, days	9 (5–15)	–	10 (6–15)	–	9 (4–16)	–	9 (5–16)	–	0,984
ICU stay, days	8 (5–14)	6 (4–9)	15 (11–22)	<0,0001	6 (4–10)	5 (4–8)	19 (12–26)	<0,0001	0,224
Hospital stay, days	20 (14–27)	20 (14–26)	22 (14–30)	0,649	18 (14–25)	18 (14–24)	26 (15–38)	0,001	0,005
ICU mortality	39 (25)	0 (0)	39 (68)	<0,0001	25 (11)	0 (0)	25 (60)	<0,0001	<0,0001

Data summarised as number (%), mean ± standard deviation or median (interquartile range).

FiO₂: fraction of inspired oxygen; HR: heart rate; HACOR: heart rate, acidosis, consciousness, oxygenation and respiratory rate score; ICU: intensive care unit; IMV: invasive mechanical ventilation; iROX: respiratory rate-oxygenation index; NIV: non-invasive ventilation; PaCO₂: partial pressure of arterial carbon dioxide; PaO₂: partial pressure of arterial oxygen; PEEP: positive end-expiratory pressure; RR: respiratory rate; SatO₂: blood oxygen saturation; SP: systolic pressure.

¹ p value for difference between NIV failure vs success group.

² p value for difference training vs validation group.

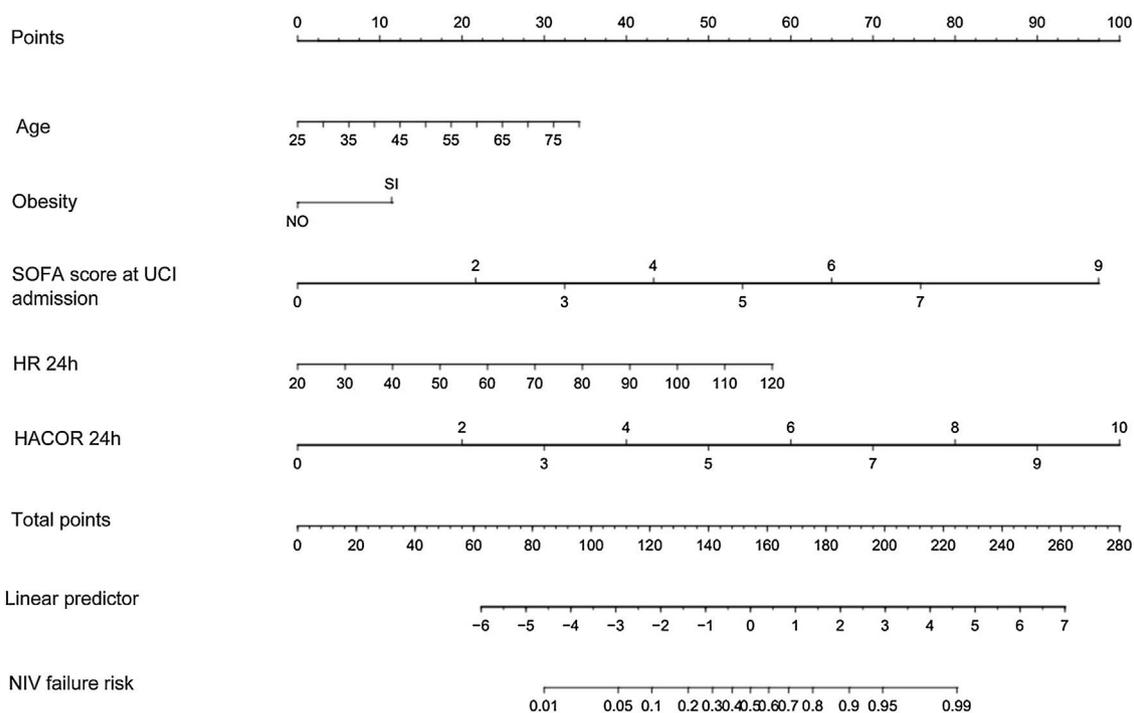


Figure 2 Nomogram to predict NIV failure in patients with SARS-CoV-2 pneumonia. (HR: heart rate; HACOR: heart rate, acidosis, conscienciousness, oxygenation and respiratory rate score; ICU: intensive care unit; NIV: non-invasive ventilation; SOFA: sequential organ failure assessment score).

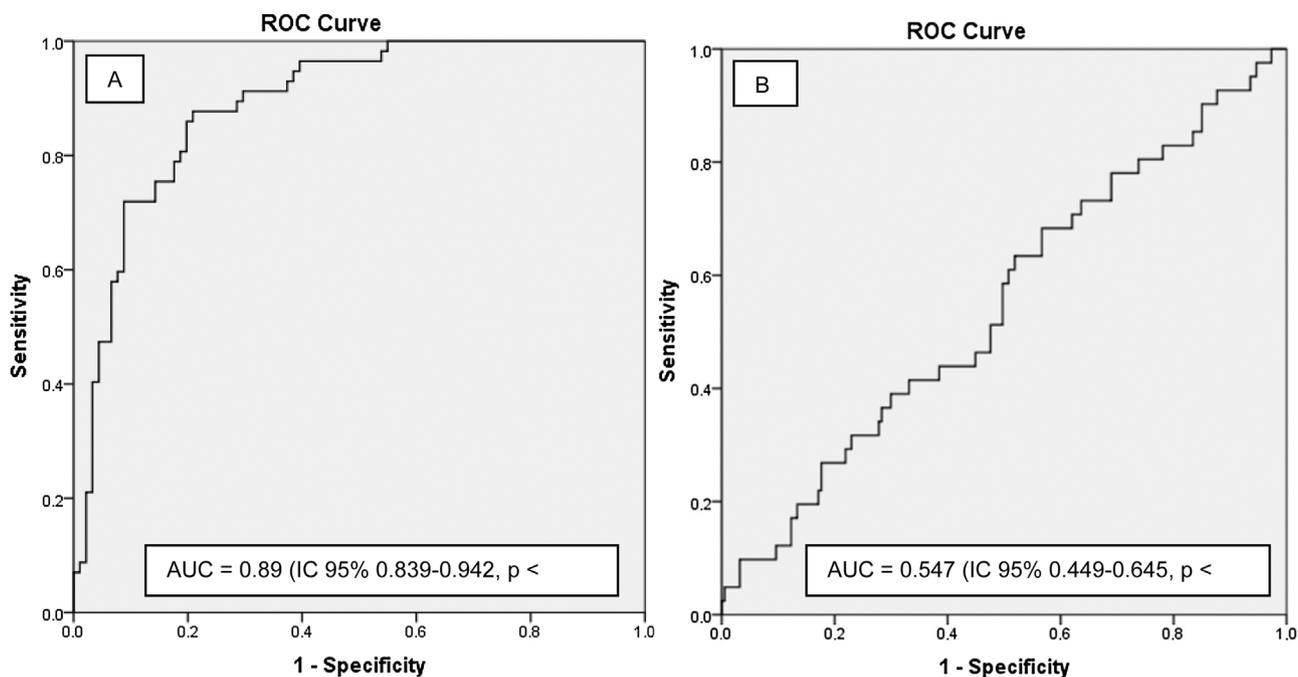


Figure 3 Nomogram ROC curves. 3A Training cohort; 3B Validation cohort.

Obesity is associated with NIV failure in patients with COVID-19²⁰ and influenza.²⁵ It is assumed that patients with obesity tolerate higher levels of positive end-expiratory pressure (PEEP), with an improvement in the expiratory flow limitation experienced by these patients.²⁶ Although we could not measure the transpulmonary pressure (Ptp), we

managed the undetected elevated Ptp, indicating the need for high levels of PEEP. However, there was likely incorrect adherence to the sitting position, related to reducing the diaphragmatic load on the lungs and airflow limitation.²⁶

The SOFA score on admission is a risk factor for NIV failure in patients with ARDS of any etiology or COVID-19.^{21,27}

Table 3 AUC for variables model in the training and validation model.

	Training cohort		Validation cohort	
	AUC (95% CI)	p value	AUC (95% CI)	p value
Age, years	0,705 (0,619–0,791)	<,0001	0,639 (0,544–0,735)	,05
Obesity	0,51 (0,487–0,689)	,048	0,526 (0,428–0,624)	,605
SOFA at admission	0,584 (0,483–0,685)	,039	0,455 (0,365–0,545)	,366
HR (24 h)	0,672 (0,578–0,765)	,001	0,518 (0,426–0,61)	,717
HACOR (24 h)	0,801 (0,724–0,879)	<,0001	0,531 (0,434–0,628)	,535

Data summarised as odds ratio (OR) and 95% confidence interval (CI).

HACOR: heart rate, acidosis, conscienciousness, oxygenation and respiratory rate score; HR: heart rate; SOFA: sequential organ failure assessment score.

The training cohort was conditioned by the respiratory component since most patients had no other organ failure. This parameter allowed us to consider that patients with only respiratory failure would be candidates for initiating support with NIV, as suggested for patients with influenza.²⁵

The fourth risk factor identified was HR at 24 h. The nomogram by Liu et al. showed the same result as ours in the univariate analysis, without confirmation in the multivariate analysis.¹⁸ We hypothesized that patients with a higher HR have higher tissue O₂ demand and consumption.

The last factor included in our model was the HACOR score, which incorporated HR, GCS, RR, PaO₂/FiO₂, and Ph.²⁸ Its usefulness has been demonstrated in the prediction of NIV failure (cutoff point of 5) in patients with acute respiratory failure.^{29–31} The novelty of the results from the training cohort is that the HACOR score was identified as a risk factor at a more advanced stage (24 h). It can be concluded that the HACOR score is a reference scale for the detection of possible NIV failure, extending its usefulness up to 24 h after initiation.

Regarding the equipment used in the training cohort, no risk factors were identified in the univariate analysis. It is important to focus on two issues. First, the small number of patients in whom a helmet was used in our series, despite the results of the study by Grieco et al.¹⁴ Considering that patients undergo prolonged periods of NIV and that the helmet is better tolerated with fewer lesions on the bridge of the nose,³² it serves as a viable starting point. In favor of this position is the low failure rate of the validation cohort, in which helmets were widely used. Second, not all data were available to analyze the use of active humidification as a risk factor. This is listed as optional during NIV in the humidification guidelines for IMV and NIV, discouraging the use of heat-humidity exchangers during NIV.³³ The use of a heated humidifier (HH) improves adaptation and secretion clearance, compared with a heat and moisture exchanger.³⁴ In addition, Esquinas et al. found that 50% of the patients who were labeled as having difficult OTI after NIV failure did not receive HH, resulting in thick and dry secretions which were a probable reason for failure.³⁵ Although this was not a part of our analysis, it could be suggested that HH use be started at the beginning of NIV support.

The nomogram did not have a good predictive capacity in the validation cohort. This result is striking given that both cohorts were managed in units with a clear preference for the use of NIRT as initial support. There are different reasons why an external validation process can be

difficult: differences between participating clinicians, populations included, and interventions performed,³⁶ an issue that seems to be observed in our training and validation cohorts (lower severity and greater use of helmets). This reflection should be more profound, but it highlights the importance of validating the proposed predictive model at a local level.

We think it would be questionable the usefulness of external validation of a predictive model for NIV failure. The decision for OTI is complex, arbitrary, individualized, intuitive, and the criteria for patients with hypoxemia have not been clearly established, with great variability between different studies.³⁷ This leads us to believe that the same patient could receive OTI at different times (and even avoid OTI), depending on the department to which the patient is admitted. Furthermore, in a review of 14 studies, it was observed that the protocol for the initiation of IMV after NIV did not affect the intubation rate,³⁸ questioning the usefulness of establishing criteria for NIV failure in hypoxemic patients. In addition, only a small percentage of patients received OTI within 3 h of meeting the previously established NIV failure criteria,³⁹ and even when these criteria were met, a high percentage of patients avoided OTI.⁴⁰ The criteria for OTI likely differed between the two cohorts. This supports the idea that, in establishing a predictive model, it is possible that this would not be a representation of patient status but rather reflect the decision of the responsible clinicians.

Limitations of the study

Our study has some important limitations. First, given the retrospective nature of the study and the workload during the pandemic, some important variables were not collected and analyzed for inclusion in the model. Esophageal pressure changes were not monitored, and the tidal volume developed during NIV, which could be related to the occurrence of self-inflicted lung injury in patients, was not collected. Second, the training cohort was from a single center and included a relatively small number of patients. The validity of the model should be evaluated in a larger population. It is important to note, our results preclude conclusions about the benefit of early IOT in patients under NIV. The study was conducted during the pandemic, when the availability of ICU beds varied over time, which may have influenced the criteria for ICU admission and the selection of patients

to receive OTI. Finally, our nomogram would only be useful for assessing NIV failure at 24 h. It may not be used for later stages of the disease.

Our nomogram is a simple tool that can be used to predict the risk of NIV failure in patients with COVID-19. These patients may benefit from close monitoring. However, extrapolation of our findings to patients managed in other ICUs may not be useful. The study findings open avenues for further research and adaptation in diverse clinical environments.

CRedit authorship contribution statement

Héctor Hernández: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing-original draft, Writing-review & editing, Visualization, Project administration. Alberto Belenguer: Conceptualization, Methodology, Investigation, Data curation, Writing-review & editing, Formal analysis. Francisco Bernal: Investigation, Writing-original draft, Data curation. Irina Hermsilla: Investigation, Writing-original draft, Data curation. Lluís Tormo: Investigation, Writing-original draft, Data curation. Estefanía Granero: Investigation, Writing-original draft, Data curation. Clara Viana: Investigation, Writing-original draft, Data curation. Rafael Zaragoza: Conceptualization, Methodology, Writing-review & editing, Visualization, Supervision.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.j.medine.2025.502148>.

Declaration of competing interest

None.

References

1. COVID-19 map. Johns Hopkins Coronavirus Resource Center n.d. <https://coronavirus.jhu.edu/map.html> (accessed April 4, 2024).
2. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 2020;395:497–506, [http://dx.doi.org/10.1016/S0140-6736\(20\)30183-5](http://dx.doi.org/10.1016/S0140-6736(20)30183-5).
3. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med*. 2020;382:1708–20, <http://dx.doi.org/10.1056/NEJMoa2002032>.
4. Wunsch H. Mechanical ventilation in COVID-19: interpreting the current epidemiology. *Am J Respir Crit Care Med*. 2020;202:1–4, <http://dx.doi.org/10.1164/rccm.202004-1385ED>.
5. Government of Canada. In: *Clinical Management of Patients with COVID-19-2nd Interim Guidance Clinical Management of Patients with COVID-19-2nd Interim Guidance*; 2020 <https://open.canada.ca/data/en/info/cf69ab87-c86e-4a6e-be12-e193cfe98383> (accessed April 4, 2024).
6. WHO. In: *COVID-19 Clinical management: living guidance*; 2021 <https://www.who.int/publications/i/item/WHO-2019-nCoV-clinical-2021-1> (accessed April 4, 2024).
7. Alhazzani W, Møller MH, Arabi YM, Loeb M, Gong MN, Fan E, et al. Surviving Sepsis Campaign: guidelines on the management of critically ill adults with Coronavirus Disease 2019 (COVID-19). *Intensive Care Med*. 2020;46:854–87, <http://dx.doi.org/10.1007/s00134-020-06022-5>.
8. Vidal-Cortés P, Díaz Santos E, Aguilar Alonso E, Amezcua Menéndez R, Ballesteros M, Bodí MA, et al. Recommendations for the management of critically ill patients with COVID-19 in Intensive Care Units. *Med Intensiva (Engl Ed)*. 2022;46:81–9, <http://dx.doi.org/10.1016/j.medine.2021.11.019>.
9. NHS England and NHS improvement, Guidance for the role and use of non-invasive respiratory support in adult patients with COVID-19 (confirmed or suspected). <https://www.nice.org.uk/guidance/ng159>. (accessed April 4, 2024).
10. Ferrer M, Esquinas A, Leon M, Gonzalez G, Alarcon A, Torres A. Noninvasive ventilation in severe hypoxemic respiratory failure: a randomized clinical trial. *Am J Respir Crit Care Med*. 2003;168:1438–44, <http://dx.doi.org/10.1164/rccm.200301-0720C>.
11. Carrillo A, Gonzalez-Diaz G, Ferrer M, Martinez-Quintana ME, Lopez-Martinez A, Llamas N, et al. Non-invasive ventilation in community-acquired pneumonia and severe acute respiratory failure. *Intensive Care M*. 2012;38:458–66, <http://dx.doi.org/10.1007/s00134-012-2475-6>.
12. González J, Benítez ID, de Gonzalo-Calvo D, Torres G, de Batlle J, Gómez S, et al. Impact of time to intubation on mortality and pulmonary sequelae in critically ill patients with COVID-19: a prospective cohort study. *Crit Care*. 2022;26:18, <http://dx.doi.org/10.1186/s13054-021-03882-1>.
13. Manejo clínico de pacientes con enfermedad por el nuevo coronavirus (COVID-19). https://www.mscbs.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCovChina/documentos/Protocolo_manejo_clinico_ah_COVID-19.pdf (accessed April 4, 2024).
14. Grieco DL, Menga LS, Cesarano M, Rosà T, Spadaro S, Bitondo MM, et al. Effect of helmet noninvasive ventilation vs high-flow nasal oxygen on days free of respiratory support in patients with COVID-19 and moderate to severe hypoxemic respiratory failure: The HENIVOT randomized clinical trial. *JAMA*. 2021;325:1731–43, <http://dx.doi.org/10.1001/jama.2021.4682>.
15. Chacko B, Thomas L, Sharma R, Yadav B, Jeyaseelan L, Arul AO, et al. Noninvasive ventilation in the management of respiratory failure due to COVID-19 infection: experience from a resource-limited setting. *Mayo Clin Proc*. 2022;97:31–45, <http://dx.doi.org/10.1016/j.mayocp.2021.10.002>.
16. Walker J, Dolly S, Ng L, Prior-Ong M, Sabapathy K. The role of CPAP as a potential bridge to invasive ventilation and as a ceiling-of-care for patients hospitalized with Covid-19-An observational study. *PLoS One*. 2020;15:e0244857, <http://dx.doi.org/10.1371/journal.pone.0244857>.
17. Chaudhuri D, Jinah R, Burns KEA, Angriman F, Ferreyro BL, Munshi L, et al. Helmet noninvasive ventilation compared to facemask noninvasive ventilation and high-flow nasal cannula in acute respiratory failure: a systematic review and meta-analysis. *Eur Respir J*. 2022;59:2101269, <http://dx.doi.org/10.1183/13993003.01269-2021>.
18. Liu L, Xie J, Wu W, Chen H, Li S, He H, et al. A simple nomogram for predicting failure of non-invasive respiratory strategies in adults with COVID-19: a retrospective multicentre study. *Lancet Digit Health*. 2021;3:e166–74, [http://dx.doi.org/10.1016/S2589-7500\(20\)30316-2](http://dx.doi.org/10.1016/S2589-7500(20)30316-2).

19. Wang Y, Lu X, Li Y, Chen H, Chen T, Su N, et al. Clinical course and outcomes of 344 intensive care patients with COVID-19. *Am J Respir Crit Care Med.* 2020;201:1430–4, <http://dx.doi.org/10.1164/rccm.202003-0736LE>.
20. Kofod LM, Nielsen Jeschke K, Kristensen MT, Krogh-Madsen R, Monefeldt Albek C, Hansen EF. COVID-19 and acute respiratory failure treated with CPAP. *Eur Clin Respir J.* 2021;8:1910191, <http://dx.doi.org/10.1080/20018525.2021.1910191>.
21. Menzella F, Barbieri C, Fontana M, Scelfo C, Castagnetti C, Ghidoni G, et al. Effectiveness of noninvasive ventilation in COVID-19 related-acute respiratory distress syndrome. *Clin Respir J.* 2021;15:779–87, <http://dx.doi.org/10.1111/crj.13361>.
22. Faraone A, Beltrame C, Crociani A, Carrai P, Lovicu E, Filetti S, et al. Effectiveness and safety of noninvasive positive pressure ventilation in the treatment of COVID-19-associated acute hypoxemic respiratory failure: a single center, non-ICU setting experience. *Intern Emerg Med.* 2021;16:1183–90, <http://dx.doi.org/10.1007/s11739-020-02562-2>.
23. Antonelli M, Conti G, Moro M, Esquinas A, Gonzalez-Diaz G, Confalonieri M, et al. Predictors of failure of noninvasive positive pressure ventilation in patients with acute hypoxemic respiratory failure: a multi-center study. *Intensive Care Med.* 2001;27:1718–28, <http://dx.doi.org/10.1007/s00134-001-1114-4>.
24. Zhan Q, Sun B, Liang L, Yan X, Zhang L, Yang J, et al. Early use of noninvasive positive pressure ventilation for acute lung injury: a multicenter randomized controlled trial. *Crit Care Med.* 2012;40:455–60, <http://dx.doi.org/10.1097/ccm.0b013e318232d75e>.
25. Masclans JR, Pérez M, Almirall J, Lorente L, Marqués A, Socias L, et al. Early non-invasive ventilation treatment for severe influenza pneumonia. *Clin Microbiol Infect.* 2013;19:249–56, <http://dx.doi.org/10.1111/j.1469-0691.2012.03797.x>.
26. Lemyze M, Mallat J, Duhamel A, Pepy F, Gasan G, Barrailler S, et al. Effects of sitting position and applied positive end-expiratory pressure on respiratory mechanics of critically ill obese patients receiving mechanical ventilation. *Crit Care Med.* 2013;41:2592–9, <http://dx.doi.org/10.1097/CCM.0b013e318298637f>.
27. Boscolo A, Pasin L, Sella N, Pretto C, Tocco M, Tamburini E, et al. Outcomes of COVID-19 patients intubated after failure of non-invasive ventilation: a multicenter observational study. *Sci Rep.* 2021;11:17730, <http://dx.doi.org/10.1038/s41598-021-96762-1>.
28. Carrillo A, Lopez A, Carrillo L, Caldeira V, Guia M, Alonso N, et al. Validity of a clinical scale in predicting the failure of non-invasive ventilation in hypoxemic patients. *J Crit Care.* 2020;60:152–8, <http://dx.doi.org/10.1016/j.jcrc.2020.08.008>.
29. Bellani G, Laffey JG, Pham T, Madotto F, Fan E, Brochard L, et al. Noninvasive ventilation of patients with acute respiratory distress syndrome. Insights from the LUNG SAFE Study. *Am J Respir Crit Care Med.* 2017;195:67–77, <http://dx.doi.org/10.1164/rccm.201606-1306OC>.
30. Carreaux G, Millán-Guilarte T, De Prost N, Razazi K, Abid S, Thille AW, et al. Failure of noninvasive ventilation for de novo acute hypoxemic respiratory failure: Role of tidal volume. *Crit Care Med.* 2016;44:282–90, <http://dx.doi.org/10.1097/CCM.0000000000001379>.
31. Frat JP, Ragot S, Coudroy R, Constantin JM, Girault C, Prat G, et al. Predictors of intubation in patients with acute hypoxemic respiratory failure treated with a noninvasive oxygenation strategy. *Crit Care Med.* 2018;46:208–15, <http://dx.doi.org/10.1097/ccm.0000000000002818>.
32. Antonelli M, Conti G, Pelosi P, Gregoretti C, Pennisi MA, Costa R, et al. New treatment of acute hypoxemic respiratory failure: noninvasive pressure support ventilation delivered by helmet—a pilot controlled trial. *Crit Care Med.* 2002;30:602–8, <http://dx.doi.org/10.1097/00003246-200203000-00019>.
33. Restrepo RD, Walsh BK, American Association for Respiratory Care. Humidification during invasive and noninvasive mechanical ventilation: 2012. *Respir Care.* 2012;57:782–8, <http://dx.doi.org/10.4187/respcare.01766>.
34. Kumar A, Sinha C, Kumar A, Kumar N, Kumar A, Singh K, et al. Inefficient humidification as the cause of non-invasive ventilation failure in COVID-19 patients. *Braz J Anesthesiol.* 2021;71:671–2, <http://dx.doi.org/10.1016/j.bjane.2021.07.021>.
35. Esquinas A, Nava S, Scala R, Carrillo A, González Diaz G, Artacho R. Humidification and difficult endotracheal intubation in failure of noninvasive mechanical ventilation. Preliminary results. *Am J Respir Crit Care Med.* 2008;177:A644.
36. Rothwell PM. External validity of randomised controlled trials: “To whom do the results of this trial apply?”. *Lancet.* 2005;365:82–93, [http://dx.doi.org/10.1016/S0140-6736\(04\)17670-8](http://dx.doi.org/10.1016/S0140-6736(04)17670-8).
37. Tulaimat A. Examining the intubation decision in randomised clinical trials. *Eur Respir J.* 2021;57:2100051, <http://dx.doi.org/10.1183/13993003.00051-2021>.
38. Coudroy R, Hoppe MA, Robert R, Frat JP, Thille AW. Influence of noninvasive ventilation protocol on intubation rates in subjects with de novo respiratory failure. *Respir Care.* 2020;65:525–34, <http://dx.doi.org/10.4187/respcare.07104>.
39. Yarnell CJ, Johnson A, Dam T, Jonkman A, Liu K, Wunsch H, et al. Do thresholds for invasive ventilation in hypoxemic respiratory failure exist? A cohort study. *Am J Respir Crit Care Med.* 2023;207:271–82, <http://dx.doi.org/10.1164/rccm.202206-1092OC>.
40. Hakim R, Watanabe-Tejada L, Sukhal S, Tulaimat A. Acute respiratory failure in randomized trials of noninvasive respiratory support: A systematic review of definitions, patient characteristics, and criteria for intubation. *J Crit Care.* 2020;57:141–7, <http://dx.doi.org/10.1016/j.jcrc.2020.02.018>.