## ORIGINAL

# Assessment of heart rate, acidosis, consciousness, oxygenation, and respiratory rate to predict noninvasive ventilation failure in hypoxemic patients

Jun Duan[*], Xiaoli Han, Linfu Bai, Lintong Zhou and Shicong Huang

### Abstract

**Purpose:**  To develop and validate a scale using variables easily obtained at the bedside for prediction of failure of noninvasive ventilation (NIV) in hypoxemic patients.

**Methods:**  The test cohort comprised 449 patients with hypoxemia who were receiving NIV. This cohort was used to develop a scale that considers heart rate, acidosis, consciousness, oxygenation, and respiratory rate (referred to as the HACOR scale) to predict NIV failure, defined as need for intubation after NIV intervention. The highest possible score was 25 points. To validate the scale, a separate group of 358 hypoxemic patients were enrolled in the validation cohort.

**Results:**  The failure rate of NIV was 47.8 and 39.4% in the test and validation cohorts, respectively. In the test cohort, patients with NIV failure had higher HACOR scores at initiation and after 1, 12, 24, and 48 h of NIV than those with successful NIV. At 1 h of NIV the area under the receiver operating characteristic curve was 0.88, showing good predictive power for NIV failure. Using 5 points as the cutoff value, the sensitivity, specificity, positive predictive value, negative predictive value, and diagnostic accuracy for NIV failure were 72.6, 90.2, 87.2, 78.1, and 81.8%, respectively. These results were confirmed in the validation cohort. Moreover, the diagnostic accuracy for NIV failure exceeded 80% in subgroups classified by diagnosis, age, or disease severity and also at 1, 12, 24, and 48 h of NIV. Among patients with NIV failure with a HACOR score of >5 at 1 h of NIV, hospital mortality was lower in those who received intubation at ≤12 h of NIV than in those intubated later [58/88 (66%) vs. 138/175 (79%); $p = 0.03$).

*Correspondence:  duanjun412589@163.com
Department of Respiratory Medicine, The First Affiliated Hospital of Chongqing Medical University, Youyi Road 1, Yuzhong District, Chongqing 400016, People's Republic of China

**Take-home message:** The newly developed HACOR scale takes into account heart rate, acidosis, consciousness, oxygenation, and respiratory rate to predict failure of noninvasive ventilation in patients with hypoxemic respiratory failure. The HACOR scale can therefore be used to make a rapid and easy assessment of noninvasive ventilation failure or success because these variables are easily obtained by simple bedside measurements. Patients with a higher HACOR score are more likely to experience noninvasive ventilation failure. With a score of 5 as cutoff value, the diagnostic accuracy of the HACOR scale was good in different subgroups classified by diagnosis, age, or disease severity, and at different time points. In patients at high risk of noninvasive ventilation failure, early intubation may decrease hospital mortality.

Springer

**Conclusions:** The HACOR scale variables are easily obtained at the bedside. The scale appears to be an effective way of predicting NIV failure in hypoxemic patients. Early intubation in high-risk patients may reduce hospital mortality.

**Keywords:** Noninvasive ventilation, Hypoxemic respiratory failure, Intubation, Predictors, Sensitivity, Specificity

## Introduction

The use of noninvasive ventilation (NIV) in critically ill patients has dramatically increased [1] as it significantly reduces the work of breathing in patients with acute respiratory failure, thereby reducing the need for intubation [2, 3]. Although NIV is frequently used in patients with hypoxemic respiratory failure, its failure rate remains high (25–59%) [4–9], indicating that not all patients benefit from this treatment.

Previous studies have reported that patients who experience NIV failure have a higher heart rate, lower pH, lower Glasgow Coma Scale (GCS) score, lower oxygenation, and higher respiratory rate than those who experience successful NIV [5, 10–16]. These variables can be used to predict NIV failure. However, the predictive power of NIV failure is low when based only on a single variable. Hypothesizing that a combination of these variables has to potential to increase the predictive power, we combined several variables that are easily obtained by simple bedside measurements in patients with hypoxemic respiratory failure to develop a scale for the prediction of NIV failure. We then explored how to use this scale to guide the clinical use of NIV.

## Methods

This was a prospective observational study performed in the respiratory intensive care unit (ICU) of a teaching hospital (The First Affiliated Hospital of Chongqing Medical University, Chongqing, China). The ethics committee and institutional review board approved this study (NO. 22013). All patients who were admitted to the ICU for NIV due to hypoxemic respiratory failure were enrolled in the study, but patients were subsequently excluded due to: presence of do-not-intubate orders, presence of chronic obstructive pulmonary disease, requirement for emergency intubation, and NIV intolerance. NIV intolerance was defined as patient refusal for NIV because of discomfort [17]. Informed consent was obtained from patients or their family members.

The decision to initiate NIV (BiPAP Vision or V60; Philips Respironics, Carlsbad, CA) was made by the attending physicians based on the following criteria: clinical presentation of respiratory distress at rest (such as active contraction of the accessory inspiratory muscles or paradoxical abdominal motion), partial pressure of arterial oxygen (PaO$_2$) of <60 mmHg or a PaO$_2$/fraction of inspired oxygen (FiO$_2$) ratio of <300 with supplemental oxygen.

The NIV was managed by attending physicians, respiratory therapists, and nurses according to previously published methods [18]. A face mask (ZS-MZA Face Mask; Shanghai Zhongshan Medical Technology Co., Shanghai, China) was the first choice for NIV treatment. Selection of the mask was based on the patient's facial type. The straps of the mask were kept as tight as possible while remaining comfortable to the patient. Patients were placed in a semirecumbent position to avoid aspiration, assuming there was no contraindication to this position. The positive-end expiratory pressure was maintained at 4–8 cmH$_2$O. Inspiratory pressure was initially set at 10 cmH$_2$O (above zero) and then increased in increments of 2 cmH$_2$O to achieve the best control of dyspnea and tolerance of the patient. If a patient did not tolerate 10 cmH$_2$O of inspiratory pressure, the latter was allowed to decrease to 8 even further to 6 cmH$_2$O. The fractional concentration of oxygen was set to achieve peripheral oxygen saturation of >92%. At the beginning of treatment, continuous use of NIV was encouraged. Once the patient recovered from respiratory failure, NIV was used intermittently until the patient could be completely weaned from it.

NIV failure was defined as requirement of intubation after NIV intervention based on the following criteria: respiratory or cardiac arrest, failure to maintain a PaO$_2$/FiO$_2$ of >100, development of conditions necessitating intubation to protect the airway (coma or seizure disorders) or to manage copious tracheal secretions, inability to correct dyspnea, lack of improvement of signs of respiratory muscle fatigue, and hemodynamic instability without response to fluids and vasoactive agents [12, 13].

The main objective was to develop and validate a scale to predict NIV failure in patients with hypoxemic respiratory failure. The secondary objective was to report outcomes in high-risk NIV failure patients who underwent intubation at different time points.

### Statistical analysis

Data were analyzed using statistical software (SPSS 17.0; IBM Corp., Armonk, NY). Data are reported as the mean and standard deviation (SD) or as the median and interquartile range, as appropriate. Normally distributed continuous variables were analyzed using the unpaired

Student's *t* test. Non-normally distributed continuous variables were analyzed using the Mann–Whitney *U* test. Categorical variables were analyzed using the Chi-squared test or Fisher's exact test. The ability to predict NIV failure was determined using the area under the receiver operating characteristic curve (AUC). A *p* value of <0.05 was considered to be statistically significant.

We developed the risk model as follows. First we used univariate analysis to identify variables associated with NIV failure collected at 1 h of NIV in test cohort. Second, variables with a *p* value of <0.2 in the univariate analysis were entered in a stepwise multivariate logistic regression analysis to identify independent risk factors associated with NIV failure. The probability of stepwise was 0.05 for entry and 0.1 for removal. We then obtained a regression model. We evaluated the final model for goodness of fit using the Hosmer–Lemeshow test ($p > 0.05$). Third, we used the method suggested by Sullivan et al. to create the risk scale [19]. We classified the variables in the final model to clinically meaningful categories and recorded the midpoint in each category. For each variable, we set a category with the lowest risk for NIV failure as the within-group reference and assigned zero points; we then calculated the weight in each category (the difference between reference category multiplied by the *β* regression coefficient per unit increase). Finally, we assigned 1 point to the category with the lowest weight and set this weight as the between-groups reference. The value of that weight in the other category divided by the between-groups reference was then calculated and this value rounded off to the nearest integer as the assigned points. The risk scale for NIV failure was the sum of the points. We performed 1000 bootstrap samples to estimate the odds ratio (OR) and 95% confidence interval (CI) of NIV failure per 1-point increment for internal validation. The cutoff value was determined based on the positive likelihood ratio of NIV failure of >5.

The sample size was calculated by Buderer's formula [20]. As there was no standard method to diagnose NIV failure, we could not obtain the known sensitivity and specificity. Based on clinical experience, we estimated that the risk scale for NIV failure reached 70% of sensitivity and 90% of specificity. The average prevalence of NIV failure was 43.5% in previous studies [4–8]. We chose the $\alpha = 0.05$ and maximum marginal error of estimate = 5%. Thus, a minimal sample size of 742 cases was required.

## Results

We enrolled 449 patients in the test cohort from June 2011 to June 2014, and another 358 patients in the validation cohort from July 2014 to June 2016 (Table 1). Patients comprising the test cohort had higher APACHE II (Acute Physiology and Chronic Health Evaluation II) scores at start of NIV treatment that did the patients in the validation cohort (mean ± SD: 18 ± 5 vs. 16 ± 5; $p < 0.01$). This difference resulted in a higher NIV failure rate in the test cohort (47.8 vs. 39.4%; $p = 0.02$). Taking both cohorts into consideration, at 1, 12, 24 and 48 h after initiation of NIV, 807, 667, 555 and 438 patients, respectively, were still receiving NIV.

At 1 h of NIV, multivariable logistic regression analysis showed that heart rate, acidosis (assessed by pH), consciousness (assessed by GCS), oxygenation, and respiratory rate were independent risk factors for NIV failure in the test cohort [Electronic Supplementary Material (ESM) Table 1]. We used these five variables to develop a risk scale, which we named HACOR, to predict NIV failure. The categories, weights, and assignment of points in each variable are summarized in ESM Table 2. The HACOR scale ranges from 0 to 25 points. At 1 h of NIV, bootstrap analysis showed that the OR of NIV failure was 1.73 (95% CI 1.58–1.95) per 1-point increment in test cohort.

A summary of the NIV failure rate in relation to different ranges of HACOR scores is show in Fig. 1. In both the test and validation cohort, patients with NIV failure had a higher HACOR score at NIV initiation and at 1, 12, 24 and 48 h of NIV than those with NIV success (Table 2; Fig. 2). Before intubation, the highest value of the HACOR score was reached in patients with NIV failure. Further, from the initiation of NIV treatment to 48 h of NIV the HACOR score barely improved in the patients with NIV failure, but it dramatically improved in those patients with NIV success.

The predictive power of NIV failure diagnosed by HACOR score is summarized in Tables 3 and 4 and ESM Figs. 1 and 2. The AUC was 0.88 and 0.91 in the test and validation cohorts, respectively, at 1 h of NIV, which shows the good diagnostic power of the HACOR score for NIV failure. Using a HACOR score of 5 as the cutoff value, the diagnostic accuracy for NIV failure assessed at 1 h of NIV was 81.8 and 86.0% in the test and validation cohorts, respectively. In subgroups classified by diagnosis, age, or disease severity, the diagnostic accuracy also exceeded 80% at 1 h of NIV; furthermore, the diagnostic accuracy also exceeded 80% when the HACOR score was assessed at 1, 12, 24 or 48 h of NIV (Table 4).

There were 505 and 302 patients with a HACOR score of ≤5 and >5 at 1 h of NIV, respectively. Among those with a HACOR score of ≤5, the NIV failure rate was 18.4% and hospital mortality was 21.6%. In comparison, among patients with a HACOR score of >5, the NIV failure rate was 87.1% and hospital mortality was 65.2%. Among the NIV failure patients, 88 patients were intubated within 12 h of NIV initiation (early intubation) and 175 patients were intubated after 12 h of NIV initiation

**Table 1** Demographics of patients with noninvasive ventilation failure and success

| Demographics | Test cohort | | $p^a$ | Validation cohort | | $p^a$ | $p^b$ |
|---|---|---|---|---|---|---|---|
| | NIV failure ($N = 215$) | NIV success ($N = 234$) | | NIV failure ($N = 141$) | NIV success ($N = 217$) | | |
| Age (years) | 66 ± 17 | 65 ± 17 | 0.51 | 67 ± 17 | 65 ± 17 | 0.38 | 0.86 |
| Male gender (%) | 161 (75%) | 153 (65%) | 0.03 | 99 (70%) | 156 (72%) | 0.81 | 0.70 |
| Diagnosis | | | | | | | |
| Pneumonia | 104 (48%) | 141 (60%) | 0.01 | 74 (53%) | 132 (61%) | 0.13 | 0.43 |
| ARDS | 61 (28%) | 24 (10%) | <0.01 | 27 (19%) | 18 (8%) | <0.01 | 0.02 |
| Pulmonary cancer | 30 (14%) | 16 (7%) | 0.02 | 21 (15%) | 16 (7%) | 0.03 | >0.99 |
| Pulmonary embolism | 6 (3%) | 15 (6%) | 0.08 | 3 (2%) | 12 (6%) | 0.18 | 0.86 |
| Heart failure | 2 (1%) | 13 (6%) | <0.01 | 4 (3%) | 16 (7%) | 0.10 | 0.16 |
| Others | 12 (6%) | 25 (11%) | 0.25 | 12 (9%) | 23 (11%) | 0.59 | 0.46 |
| Data collected at NIV initiation | | | | | | | |
| APACHE II score | 19 ± 6 | 16 ± 5 | <0.01 | 18 ± 5 | 14 ± 4 | <0.01 | <0.01 |
| Systolic blood pressure (mmHg) | 131 ± 25 | 132 ± 25 | 0.65 | 133 ± 28 | 134 ± 30 | 0.77 | 0.20 |
| Diastolic blood pressure (mmHg) | 79 ± 17 | 80 ± 15 | 0.35 | 79 ± 17 | 80 ± 17 | 0.85 | 0.85 |
| Heart rate (beats/min) | 124 ± 24 | 110 ± 24 | <0.01 | 120 ± 23 | 113 ± 24 | <0.01 | 0.67 |
| Respiratory rate (breaths/min) | 34 ± 8 | 30 ± 7 | <0.01 | 34 ± 7 | 31 ± 7 | <0.01 | 0.85 |
| pH | 7.40 ± 0.11 | 7.44 ± 0.08 | <0.01 | 7.42 ± 0.10 | 7.43 ± 0.08 | 0.09 | 0.53 |
| $PaCO_2$ (mmHg) | 38 ± 17 | 38 ± 13 | 0.77 | 37 ± 13 | 37 ± 12 | 0.57 | 0.19 |
| $PaO_2/FiO_2$ | 137 ± 65 | 179 ± 83 | <0.01 | 146 ± 68 | 165 ± 63 | <0.01 | 0.79 |
| GCS | 14.4 ± 1.7 | 14.8 ± 0.8 | <0.01 | 14.3 ± 1.6 | 14.8 ± 0.6 | <0.01 | 0.84 |
| Data collected at 1 h of NIV | | | | | | | |
| Systolic blood pressure (mmHg) | 122 ± 22 | 122 ± 20 | 0.75 | 128 ± 26 | 125 ± 23 | 0.36 | <0.01 |
| Diastolic blood pressure (mmHg) | 75 ± 14 | 75 ± 12 | 0.91 | 76 ± 16 | 75 ± 13 | 0.30 | 0.48 |
| Heart rate (beats/min) | 117 ± 24 | 102 ± 22 | <0.01 | 118 ± 25 | 103 ± 21 | <0.01 | 0.96 |
| Respiratory rate (breaths/min) | 33 ± 10 | 27 ± 7 | <0.01 | 33 ± 9 | 27 ± 7 | <0.01 | 0.85 |
| Tidal volume (ml) | 477 ± 184 | 466 ± 180 | 0.53 | 450 ± 167 | 428 ± 156 | 0.21 | <0.01 |
| Minute ventilation (L/min) | 16.8 ± 10.1 | 12.6 ± 6.1 | <0.01 | 14.4 ± 8.1 | 11.4 ± 5.1 | <0.01 | <0.01 |
| pH | 7.39 ± 0.12 | 7.45 ± 0.06 | <0.01 | 7.40 ± 0.12 | 7.44 ± 0.06 | <0.01 | 0.30 |
| $PaCO_2$ (mmHg) | 40 ± 18 | 38 ± 12 | 0.09 | 39 ± 16 | 36 ± 10 | 0.03 | 0.16 |
| $PaO_2/FiO_2$ | 133 ± 69 | 222 ± 80 | <0.01 | 131 ± 60 | 217 ± 75 | <0.01 | 0.48 |
| GCS | 14.1 ± 1.9 | 14.9 ± 0.5 | <0.01 | 13.9 ± 2.0 | 14.9 ± 0.4 | <0.01 | 0.65 |
| Patients remaining on NIV at: | | | | | | | |
| 1 h of NIV | 215 | 234 | – | 141 | 217 | – | 0.02 |
| 12 h of NIV | 151 | 196 | – | 112 | 208 | – | 0.03 |
| 24 h of NIV | 111 | 169 | – | 81 | 194 | – | 0.01 |
| 48 h of NIV | 83 | 129 | – | 58 | 168 | – | <0.01 |

Data are presented as the mean ± SD or as a number with/without the percentage in parenthesis, as appropriate
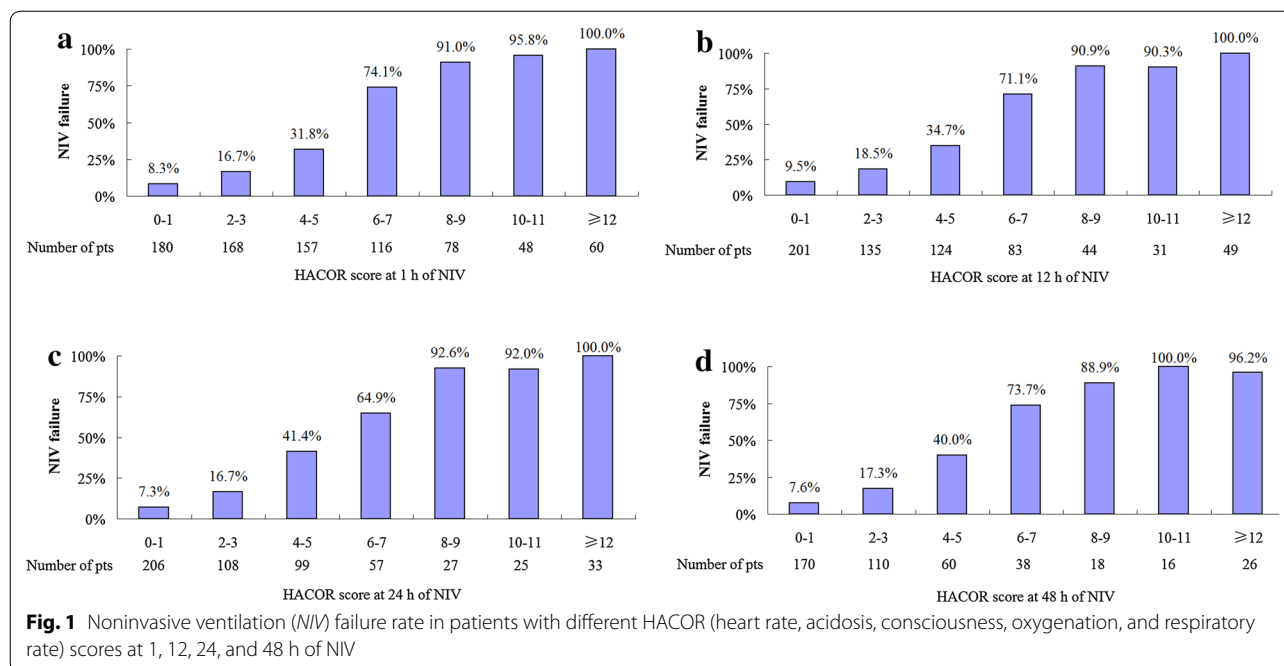
*NIV* Noninvasive ventilation, *ARDS* acute respiratory distress syndrome, *APACHE II* Acute Physiology and Chronic Health Evaluation II, *PaO₂* partial pressure of arterial oxygen, *FiO₂* fraction of inspired oxygen, *GCS* Glasgow Coma Scale

[a] *p* for difference between NIV failure vs. success

[b] *p* for difference between test cohort vs. validation cohort

(late intubation) (Table 5). Patients who were intubated early had a higher HACOR score at NIV initiation and 1 h of NIV than those who were intubated late, but the earlier had lower hospital mortality [58/88 (66%) vs. 138/175 (79%); $p = 0.03$]. Compared with late intubation, the crude OR of early intubation for death in hospital

**Fig. 1** Noninvasive ventilation (*NIV*) failure rate in patients with different HACOR (heart rate, acidosis, consciousness, oxygenation, and respiratory rate) scores at 1, 12, 24, and 48 h of NIV

**Table 2  The HACOR score at different time points**

| NIV time points | Test cohort | | p | Validation cohort | | p |
|---|---|---|---|---|---|---|
| | NIV failure | NIV success | | NIV failure | NIV success | |
| Time points | | | | | | |
| NIV initiation | 7.5 ± 3.5 | 4.5 ± 3.0 | <0.01 | 7.3 ± 3.5 | 4.9 ± 3.0 | <0.01 |
| 1 h of NIV | 7.9 ± 4.3 | 2.5 ± 2.2 | <0.01 | 8.2 ± 4.3 | 2.5 ± 2.1 | <0.01 |
| 12 h of NIV | 7.9 ± 5.3 | 2.4 ± 2.4 | <0.01 | 7.7 ± 4.4 | 2.0 ± 2.0 | <0.01 |
| 24 h of NIV | 7.8 ± 5.5 | 1.9 ± 2.3 | <0.01 | 7.4 ± 4.6 | 1.7 ± 1.9 | <0.01 |
| 48 h of NIV | 7.1 ± 5.8 | 2.0 ± 2.5 | <0.01 | 8.2 ± 5.1 | 1.3 ± 1.6 | <0.01 |
| Before intubation | 11.8 ± 5.2 | – | – | 10.2 ± 4.6 | – | – |
| Difference between time points | | | | | | |
| H0 − H1 | 0 (−2 to −2) | 2 (0–4) | <0.01 | −1 (−3 to −1) | 2 (0–4) | <0.01 |
| H0 − H12 | 0 (−3 to −3) | 2 (0–4) | <0.01 | 0 (−2 to −1.8) | 3 (1–5) | <0.01 |
| H0 − H24 | 0 (−4 to −3) | 3 (0–5) | <0.01 | 0 (−2.5 to −2) | 4 (2–5) | <0.01 |
| H0 − H48 | 1 (−3 to −5) | 3 (0–5) | <0.01 | 0 (−3 to −2) | 4 (2–6) | <0.01 |
| H0 − intubation | −3 (−8 to −0) | – | – | −2 (−5 to –0) | – | – |

Values in table are presented as the mean ± SD or as the median with the interquartile range (IQR) in parenthesis, as appropriate

*HACOR* Heart rate, acidosis, consciousness, oxygenation, and respiratory rate, *H0* initiation of NIV, *H1* 1 h of NIV, *H12* 12 h of NIV, *H24* 24 h of NIV, *H48* 48 h of NIV

was 0.52 (95% CI 0.29–0.92). We adjusted the OR by age, APACHE II score, diagnosis and HACOR scores and still found that early intubation was a protective factor for death in hospital (adjusted OR 0.52, 95% CI 0.27–0.99).

## Discussion

We developed a novel scale, called HACOR, for prediction of NIV failure in patients with hypoxemic respiratory failure. This scale takes into account heart rate, acidosis, consciousness, oxygenation, and respiratory rate, variables which are easily obtained by simple bedside measurements. Thus, the HACOR scale is a rapid and convenient tool to assess and predict NIV failure. We also showed that a HACOR score of 5 as cutoff value has good diagnostic accuracy for NIV failure even when the scale was assessed in different subgroups classified by diagnosis, age, or disease severity, or at different time points.

**Fig. 2** The HACOR score in patients with NIV failure and success from initiation to 48 h of NIV
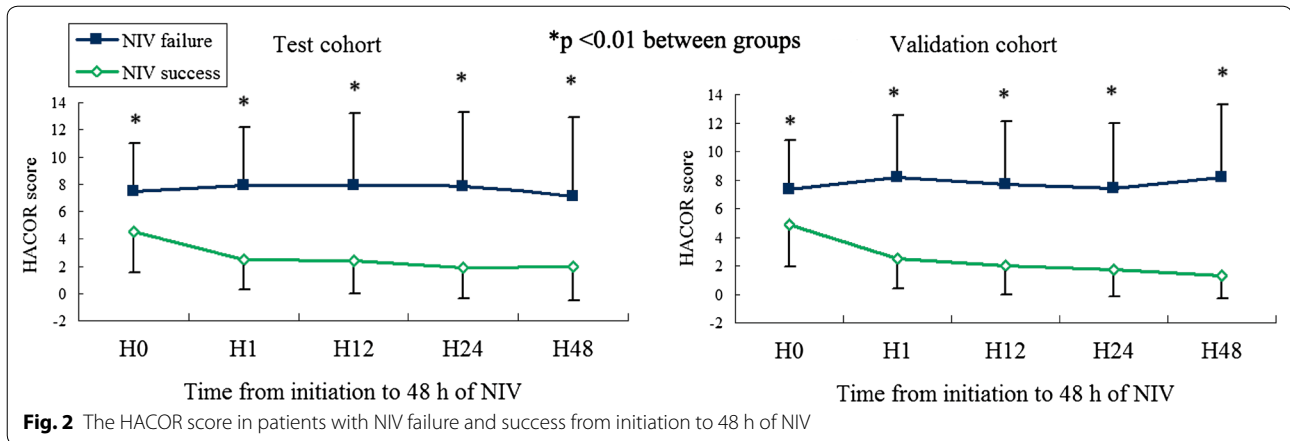
**Table 3** Predictive power of noninvasive ventilation failure diagnosed by the HACOR score assessed at 1 h of NIV

| Variables | AUC (95% CI) | Cutoff point | SE (%) | SP (%) | PPV (%) | NPV (%) | Diagnostic accuracy (%) | LR+ | LR− |
|---|---|---|---|---|---|---|---|---|---|
| Test cohort (N = 449) | 0.88 (0.84–0.91) | >5 | 72.6 | 90.2 | 87.2 | 78.1 | 81.8 | 7.38 | 0.30 |
| Validation cohort (N = 358) | 0.91 (0.88–0.94) | >5 | 75.9 | 92.6 | 87.0 | 85.5 | 86.0 | 10.29 | 0.26 |
| Subgroups classified by diagnosis | | | | | | | | | |
| Pneumonia (N = 451) | 0.89 (0.85–0.91) | >5 | 73.0 | 90.8 | 83.9 | 83.8 | 83.8 | 7.98 | 0.30 |
| ARDS (N = 130) | 0.89 (0.82–0.94) | >5 | 71.6 | 92.9 | 95.5 | 60.9 | 78.5 | 10.02 | 0.31 |
| Pulmonary cancer (N = 83) | 0.84 (0.74–0.91) | >5 | 72.5 | 93.7 | 94.9 | 68.2 | 80.7 | 11.61 | 0.29 |
| Pulmonary embolism (N = 36) | 0.85 (0.69–0.95) | >5 | 77.8 | 92.3 | 77.8 | 92.6 | 88.7 | 10.5 | 0.24 |
| Heart failure (N = 35) | 0.99 (0.87–1.00) | >5 | 66.7 | 100 | 100 | 93.5 | 94.3 | -[a] | 0.33 |
| Subgroups classified by age | | | | | | | | | |
| ≤59 years (N = 233) | 0.90 (0.85–0.93) | >5 | 77.1 | 89.8 | 84.1 | 84.8 | 84.6 | 7.54 | 0.26 |
| 60–79 years (N = 404) | 0.90 (0.86–0.92) | >5 | 73.0 | 91.6 | 87.2 | 81.2 | 83.4 | 8.69 | 0.29 |
| ≥80 years (N = 170) | 0.88 (0.82–0.92) | >5 | 72.0 | 93.2 | 90.8 | 78.1 | 83.0 | 10.55 | 0.30 |
| Subgroups classified by APACHE II score | | | | | | | | | |
| ≤14 (N = 284) | 0.87 (0.83–0.91) | >5 | 64.6 | 94.2 | 81.0 | 87.3 | 86.0 | 11.03 | 0.38 |
| 15–19 (N = 299) | 0.88 (0.83–0.91) | >5 | 70.7 | 91.6 | 87.0 | 79.6 | 82.0 | 8.38 | 0.32 |
| ≥20 (N = 224) | 0.90 (0.85–0.94) | >5 | 81.9 | 83.8 | 90.1 | 72.0 | 82.6 | 5.04 | 0.22 |
| Difference of HACOR score between NIV initiation and 1 h of NIV | | | | | | | | | |
| ΔH0–H1 in test cohort | 0.68 (0.64–0.73) | ≤−3 | 23.3 | 96.2 | 84.7 | 57.7 | 61.2 | 6.05 | 0.80 |
| ΔH0–H1 in validation cohort | 0.76 (0.73–0.82) | ≤−3 | 27.0 | 97.2 | 86.4 | 67.2 | 69.6 | 9.75 | 0.75 |

*AUC* Area under the curve of receiver operating characteristics, *CI* confidence interval, *SE* sensitivity, *SP* specificity, *PPV* positive predictive value, *NPV* negative predictive value, *LR+* positive likelihood ratio, *LR−* negative likelihood ratio, *ΔH0–H1* Difference between NIV initiation and 1 h of NIV

-[a] means the value is plus infinity

**Table 4** Predictive power of noninvasive ventilation failure diagnosed by the HACOR score assessed at 1 h, 12 h, 24 h and 48 h of NIV

| NIV time points | AUC (95% CI) | Cutoff point | SE (%) | SP (%) | PPV (%) | NPV (%) | Diagnostic accuracy (%) | LR+ | LR− |
|---|---|---|---|---|---|---|---|---|---|
| 1 h of NIV (N = 807) | 0.89 (0.87–0.91) | >5 | 73.9 | 91.4 | 87.1 | 81.6 | 83.7 | 8.54 | 0.29 |
| 12 h of NIV (N = 667) | 0.87 (0.85–0.90) | >5 | 66.9 | 92.3 | 85.0 | 81.0 | 82.3 | 8.72 | 0.36 |
| 24 h of NIV (N = 555) | 0.88 (0.85–0.90) | >5 | 61.5 | 93.4 | 83.1 | 82.1 | 82.3 | 9.30 | 0.41 |
| 48 h of NIV (N = 438) | 0.87 (0.83–0.90) | >5 | 60.3 | 95.6 | 86.7 | 83.5 | 84.2 | 13.8 | 0.42 |

*HACOR* heart rate, acidosis, consciousness, oxygenation, and respiratory rate, *AUC* area under the curve of receiver operating characteristics, *CI* confidence interval, *SE* sensitivity, *SP* specificity, *PPV* positive predictive value, *NPV* negative predictive value, *LR+* positive likelihood ratio, *LR−* negative likelihood ratio, *NIV* noninvasive ventilation

**Table 5** Early versus late intubation in patients with a HACOR score of >5 at 1 h of noninvasive ventilation

| NIV time points and hospital mortality | Intubation at ≤12 h (*N* = 88) | Intubation at >12 h (*N* = 175) | *p* |
|---|---|---|---|
| HACOR score at NIV initiation | 8.9 ± 3.9 | 7.7 ± 3.0 | <0.01 |
| HACOR score at 1 h of NIV | 11.4 ± 4.0 | 8.8 ± 3.1 | <0.01 |
| HACOR score before intubation | 12.2 ± 4.7 | 11.4 ± 5.0 | 0.22 |
| Time from NIV initiation to intubation (h) | 5 (2–10) | 53 (23–132) | <0.01 |
| Hospital mortality | 58 (66%) | 138 (79%) | 0.03 |

Values in table are presented as the mean ± SD, the median with the IQR in parenthesis, or as a number with the percentage in parenthesis, as appropriate

*HACOR* heart rate, acidosis, consciousness, oxygenation, and respiratory rate, *NIV* noninvasive ventilation

In previous studies the average rate of NIV failure was reported to be 43.5% (range 25–59%) in patients with hypoxemic respiratory failure [4–8]. In our study, the failure rate was 47.9 and 39.4% in the test and validation cohorts, respectively. These failure rates in our study thus fall within the range of those reported previously, indicating that the results of our study can be extrapolated to other studies.

Patients with more severe illness are more likely to experience NIV failure [5, 7–14]. In our study, the APACHE II score in the test cohort was much higher than that in the validation cohort, resulting in higher NIV failure rate in the test cohort. Although the NIV failure rate was different in the two cohorts, the HACOR score showed good distinguishing power for NIV failure in both cohorts. Further, in the subgroups classified by APACHE II score, the HACOR score also showed good distinguishing power for NIV failure. These results indicate that the HACOR score can be used in patients with different disease severity.

The power of a single variable to predict NIV failure is low [5, 7, 11, 12]. To the best of our knowledge, our study is the first to report and assess a scale based on multiple variables to predict NIV failure in hypoxemic patients. We used five variables to develop this scale to predict NIV failure in 449 patients and validated it in another 358 patients. We found that the diagnostic accuracy of NIV failure was good in both the test and validation cohorts. We therefore conclude that this scale is a good tool to help clinical practitioners manage NIV in patients with hypoxemic respiratory failure.

Patients with ARDS or cancer are reported to have a high NIV failure rate [13, 21], while those with heart failure are reported to have a very low NIV failure rate [22]. Our study confirmed these results, although in contrast to previous studies, we used the HACOR score to predict NIV failure. Although the NIV failure rate was quite different between patients who had different diagnoses, the HACOR scale achieved a good power to distinguish NIV failure in each subgroup. The NIV failure rate also differs greatly according to patient age [23]. The HACOR scale

also showed good distinguishing power for NIV failure in young, old, or very old patients.

Identifying patients who respond well or badly to NIV is also important. Previous studies have reported that clinical conditions improved in responders but only slightly improved or did not improve in nonresponders [5, 12, 13, 24, 25]. Similar outcomes were also found in our study. The HACOR score was lower in patients who successfully underwent NIV, but it did not improve in patients with NIV failure. Therefore, we suggest that the HACOR score can also be used to assess the efficacy of NIV.

Intubation for invasive mechanical ventilation is associated with many complications, such as diaphragmatic weakness and ventilator-associated pneumonia [26, 27]. Thus, intubation should be avoided if at all possible; however, delayed intubation increases mortality in patients for whom intubation is indicated [28, 29]. Therefore, it is important to be able to identify those patients who require intubation and those who do not. The results of our study demonstrate that a HACOR score of 5 as cutoff value had a good distinguishing power for NIV failure. At 1 h of NIV, 87.1% of patients with a HACOR score of >5 required intubation and 81.6% of patients with HACOR score ≤5 did not require intubation. These values indicate that the risk of NIV failure was high in patients with a HACOR score of >5. Further, the high-risk patients who received early intubation had lower hospital mortality than those who received late intubation. Thus, the HACOR score can also be used to assess the need for intubation. Earlier intubation may benefit patients at high risk for NIV failure.

There are a number of limitations to our study. First, as this study was performed in a respiratory ICU, some patients enrolled in the study were not admitted to our unit, such as those with postoperative respiratory failure and trauma. Therefore, caution is advised when the HACOR scale is used to assess the efficacy of NIV in other patient groups. Second, the sample size was small in some subgroups, such as pulmonary embolism and heart failure and, consequently, the efficacy of the

HACOR scale in these patients may be skewed. In future studies the sample size should larger to improve the diagnostic power. Third, the result that early intubation may reduce mortality is a secondary analysis. As our study was an observational study, this result should be investigated further in randomized controlled trials.

In conclusion, we found that the HACOR scale was able to effectively predict NIV failure in patients with hypoxemic respiratory failure. A higher score indicates a higher chance of NIV failure. Because the scale consists of variables that are easily obtained at bedside, it can be used conveniently to assess the efficacy of NIV. Patients with a HACOR score of >5 had a very high risk of NIV failure. In these high-risk patients, early intubation may reduce hospital mortality.

**Electronic supplementary material**

**Compliance with ethical standards**

**Conflicts of interest**

On behalf of all authors, the corresponding author states that there is no conflict of interest.

## References

1. Demoule A, Chevret S, Carlucci A et al (2016) Changing use of noninvasive ventilation in critically ill patients: trends over 15 years in francophone countries. Intensive Care Med 42:82–92
2. L'Her E, Deye N, Lellouche F et al (2005) Physiologic effects of noninvasive ventilation during acute lung injury. Am J Respir Crit Care Med 172:1112–1118
3. Peter JV, Moran JL, Phillips-Hughes J et al (2002) Noninvasive ventilation in acute respiratory failure–a meta-analysis update. Crit Care Med 30:555–562
4. Ferrer M, Esquinas A, Leon M et al (2003) Noninvasive ventilation in severe hypoxemic respiratory failure: a randomized clinical trial. Am J Respir Crit Care Med 168:1438–1444
5. Yoshida Y, Takeda S, Akada S et al (2008) Factors predicting successful noninvasive ventilation in acute lung injury. J Anesth 22:201–206
6. Agarwal R, Handa A, Aggarwal AN et al (2009) Outcomes of noninvasive ventilation in acute hypoxemic respiratory failure in a respiratory intensive care unit in north India. Respir Care 54:1679–1687
7. Carron M, Freo U, Zorzi M, et al (2010) Predictors of failure of noninvasive ventilation in patients with severe community-acquired pneumonia. J Crit Care 25:540 e549–514
8. Masclans JR, Perez M, Almirall J et al (2013) Early non-invasive ventilation treatment for severe influenza pneumonia. Clin Microbiol Infect 19:249–256
9. Muriel A, Penuelas O, Frutos-Vivar F et al (2015) Impact of sedation and analgesia during noninvasive positive pressure ventilation on outcome: a marginal structural model causal analysis. Intensive Care Med 41:1586–1600
10. Nicolini A, Ferraioli G, Ferrari-Bravo M et al (2016) Early non-invasive ventilation treatment for respiratory failure due to severe community-acquired pneumonia. Clin Respir J 10:98–103
11. Carrillo A, Gonzalez-Diaz G, Ferrer M et al (2012) Non-invasive ventilation in community-acquired pneumonia and severe acute respiratory failure. Intensive Care Med 38:458–466
12. Antonelli M, Conti G, Esquinas A et al (2007) A multiple-center survey on the use in clinical practice of noninvasive ventilation as a first-line intervention for acute respiratory distress syndrome. Crit Care Med 35:18–25
13. Antonelli M, Conti G, Moro ML et al (2001) Predictors of failure of non-invasive positive pressure ventilation in patients with acute hypoxemic respiratory failure: a multi-center study. Intensive Care Med 27:1718–1728
14. Rana S, Jenad H, Gay PC et al (2006) Failure of non-invasive ventilation in patients with acute lung injury: observational cohort study. Crit Care 10:R79
15. Adda M, Coquet I, Darmon M et al (2008) Predictors of noninvasive ventilation failure in patients with hematologic malignancy and acute respiratory failure. Crit Care Med 36:2766–2772
16. Thille AW, Contou D, Fragnoli C et al (2013) Non-invasive ventilation for acute hypoxemic respiratory failure: intubation rate and risk factors. Crit Care 17:R269
17. Liu J, Duan J, Bai L et al (2016) Noninvasive ventilation intolerance: characteristics, predictors, and outcomes. Respir Care 61:277–284
18. Zhang Z, Duan J (2015) Nosocomial pneumonia in non-invasive ventilation patients: incidence, characteristics, and outcomes. J Hosp Infect 91:153–157
19. Sullivan LM, Massaro JM, D'Agostino RB Sr (2004) Presentation of multivariate data for clinical use: The Framingham Study risk score functions. Stat Med 23:1631–1660
20. Buderer NM (1996) Statistical methodology: I. Incorporating the prevalence of disease into the sample size calculation for sensitivity and specificity. Acad Emerg Med 3:895–900
21. Ferreira JC, Medeiros P Jr, Rego FM et al (2015) Risk factors for noninvasive ventilation failure in cancer patients in the intensive care unit: a retrospective cohort study. J Crit Care 30:1003–1007
22. Vital FM, Ladeira MT, Atallah AN (2013) Non-invasive positive pressure ventilation (CPAP or bilevel NPPV) for cardiogenic pulmonary oedema. Cochrane Database Syst Rev 5:CD005351
23. Ozsancak Ugurlu A, Sidhom SS, Khodabandeh A et al (2016) Use and outcomes of noninvasive ventilation for acute respiratory failure in different age groups. Respir Care 61:36–43
24. Nicolini A, Tonveronachi E, Navalesi P et al (2012) Effectiveness and predictors of success of noninvasive ventilation during H1N1 pandemics: a multicenter study. Minerva Anestesiol 78:1333–1340
25. Jaber S, Chanques G, Sebbane M et al (2006) Noninvasive positive pressure ventilation in patients with respiratory failure due to severe acute pancreatitis. Respiration 73:166–172
26. Jaber S, Petrof BJ, Jung B et al (2011) Rapidly progressive diaphragmatic weakness and injury during mechanical ventilation in humans. Am J Respir Crit Care Med 183:364–371
27. American Thoracic Society, Infectious Diseases Society of America (2005) Guidelines for the management of adults with hospital-acquired, ventilator-associated, and healthcare-associated pneumonia. Am J Respir Crit Care Med 171:388–416
28. Esteban A, Frutos-Vivar F, Ferguson ND et al (2004) Noninvasive positive-pressure ventilation for respiratory failure after extubation. N Engl J Med 350:2452–2460
29. Kikuchi T, Toba S, Sekiguchi Y et al (2011) Protocol-based noninvasive positive pressure ventilation for acute respiratory failure. J Anesth 25:42–49