Eur Respir J 2005; 25: 348–355
DOI: 10.1183/09031936.05.00085304

A chart of failure risk for noninvasive ventilation in patients with COPD exacerbation

M. Confalonieri*, G. Garuti#, M.S. Cattaruzza¶, J.F. Osborn¶, M. Antonelli+, G. Conti+, M. Kodric*, O. Resta§, S. Marchese*f, C. Gregoretti and A. Rossi, on behalf of the Italian noninvasive positive pressure ventilation (NPPV) study group##**

ABSTRACT: Knowing the likelihood of failure of noninvasive positive pressure ventilation (NPPV) in patients with exacerbation of chronic obstructive pulmonary disease (COPD) could indicate the best choice between NPPV and endotracheal intubation instituted earlier. For this purpose, two risk charts were designed (at admission and after 2 h of NPPV) that included all relevant measurable clinical prognostic indicators derived from a population representing the patients seen routinely in clinical practice.

Risk stratification of NPPV failure was assessed in 1,033 consecutive patients admitted to experienced hospital units, including two intensive care units, six respiratory intermediate care units, and five general wards. NPPV was successful in 797 patients.

Patients with a Glasgow Coma Score <11, acute physiology and chronic health evaluation (APACHE) II ⩾29, respiratory rate ⩾30 breaths·min^{-1} and pH at admission <7.25 have a predicted risk of failure >70%. A pH <7.25 after 2 h greatly increases the risk (>90%). The risk charts were validated on an independent group of 145 consecutive COPD patients treated with NPPV due to an acute ventilatory failure episode. To identify patients with a probability of failure >50%, the sensitivity and specificity were 33% and 96.7% on admission and 52.9% and 94.1% after 2 h of NPPV, respectively.

The prediction chart, based on data from the current study, can function as a simple tool to predict the risk of failure of noninvasive positive pressure ventilation and thus improve clinical management of patients tailoring medical intervention.

KEYWORDS: Chronic obstructive pulmonary disease, endotracheal intubation, noninvasive ventilation, respiratory failure

Noninvasive positive pressure ventilation (NPPV), in patients with exacerbations of chronic obstructive pulmonary disease (COPD) and respiratory acidosis, reduces the intubation rate and mortality [1–9]. Operated by well trained teams, NPPV is effective and safe in both intensive care settings [10] and general respiratory wards [8]. A randomised, clinical trial showed that NPPV also reduces mortality in COPD patients in the intensive care unit (ICU) within the inclusion criteria for intubation [6]. Nevertheless, two recent consensus guidelines on NPPV in acute respiratory failure (ARF) recommend that NPPV should not be used as a substitute for endotracheal intubation and invasive ventilation when the latter is clearly more appropriate [11, 12].

The likelihood of failure of NPPV is crucial in deciding if and when to apply this ventilatory technique. However, attempts to predict the success of NPPV have only been made in single-centre studies that included a limited number of patients [13–15]. Furthermore, results from randomised, controlled studies could be affected significantly by the research setting, *e.g.* strict selection criteria of patients or a unique care location. This limits the possibility of generalising the results to the daily clinical practice of general hospitals [16].

The purpose of the current study was to assess the risk of NPPV failure in a large unselected population admitted to different hospital units with expertise in the application of NPPV and to build a risk chart of failure of NPPV to be used in hospitals.

AFFILIATIONS
*S.C. Pneumologia, Ospedali Riuniti di Trieste, Trieste,
#Unità di Pneumologia, Ospedale di Mirandola, Modena,
¶Dipartimento di Sanità Pubblica, Università ''La Sapienza'', and
+Istituto di Anestesia e Rianimazione, Università del S.Cuore, Roma,
§Istituto di Malattie dell'Apparato Respiratorio, Università di Bari, Bari,
fDivisione di Pneumologia, Ospedale Civico, Palermo,
**Servizio di Anestesia e Rianimazione, CTO, Torino, and
##Divisione di Pneumologia, Ospedali Riuniti, Bergamo, Italy.

CORRESPONDENCE
M. Confalonieri
Ospedale di Cattinara
Strada di Fiume 447
I-34100 Trieste
Italy
Fax: 39 0403994668
E-mail: mconfalonieri@qubisoft.it

Received:
July 19 2004
Accepted after revision:
November 08 2004

European Respiratory Journal
Print ISSN 0903-1936
Online ISSN 1399-3003

METHODS

Patients

From December 1998 to March 2000 prospective data was collected from 1,033 consecutive patients affected by COPD exacerbation and respiratory acidosis that were treated by NPPV in addition to standard medical therapy (oxygen supplementation, systemic corticosteroids, inhaled bronchodilators, antibiotics, and diuretics if needed). These patients were admitted to two ICU, five general wards and six respiratory intermediate intensive care units (RIICU) where NPPV is the first-line intervention for such patients [17]. The admission criteria to different units were based on the need for an appropriate level of care, but the personnel in each unit were well trained in the use of NPPV, with $\geqslant$3 yrs experience. The various units used different types of ventilators, either intensive care ventilators (Servo 300; Siemens Elema, Solna, Sweden. Puritan Bennet 7200; Puritain Bennett Co., Overland Park KS, USA) or less sophisticated ventilators (BiPAP Respironics; Respironics Inc., Murrysville, PA, USA. O'nyx Mallinckrodt; Mallinckrodt Inc, Minneapolis, MN, USA. Helia Saem; Saime, Savigny-le-Temple, France) *via* face and/or nasal mask. Every patient who received NPPV because of hypercapnic respiratory failure and respiratory acidosis due to exacerbation of COPD was included in the current observational study. Post-operative patients and patients receiving NPPV for weaning were excluded. The definition of COPD exacerbation was in accordance with that of the American Thoracic Society [18]. NPPV failure was defined as the need for endotracheal intubation. Pre-determined criteria for endotracheal intubation included: 1) worsening of pH and carbon dioxide tension in arterial blood (P_{a,CO_2}) in spite of correct NPPV administration (*e.g.* pH $\leqslant$0.04 and $P_{a,CO_2} \geqslant 0.8$ kPa ($\geqslant$6 mmHg)); 2) the need to protect the airways (coma or seizure disorders) or to manage copious secretions; 3) haemodynamic instability (heart rate $<$50 beats·min^{-1} with loss of alertness, and/or systolic blood pressure $<$70 mmHg); and 4) agitation and inability to tolerate the mask.

The following data were collected for every patient. 1) General demographic information (age, sex, weight and height) and clinical data. Severity of illness was assessed by acute physiology and chronic health evaluation (APACHE) II score [19], and neurological status was assessed by the Glasgow Coma Scale (GCS) [20]. 2) Data relative to the institution of NPPV, including cause of exacerbation, arterial blood gases (before beginning NPPV, after 2 h and at discharge), respiratory rate (RR), cardiac frequency, length of stay in hospital, and total hours of ventilation. 3) Data on comorbidities, previous use of home mechanical ventilation and long-term oxygen therapy.

Variable definition and statistical analysis

The outcome variable was defined as failure of NPPV due to invasive ventilation or death.

After descriptive analysis, univariate analysis was performed using the unpaired t-test to examine differences in continuous variables. Multivariable logistic regression was used to analyse the independent effect of each variable on the outcome (treatment failure) and the strategy used to determine which variables should be included in the multivariable analysis was based on whether the variable was of direct interest or merely a confounder. Variables of interest were included, whilst confounders that showed little or no effect, or variables related to a very small number of patients, were eliminated. After the elimination of these confounders, factors of interest were eliminated if their effect was not statistically significant at the 5% level (except pH between 7.25–7.29 at admission due to the interest of pneumologists for this category of pH and to allow comparability between the two charts). This strategy may, of course, be considered rather arbitrary, as statistical significance at the 5% level does not necessarily imply that the variable will be a useful predictor. Also, variables such as body weight, which for many patients was not measured, were excluded as these missing values reduce the effective sample size, and, if they were not available for a new patient, it would not help in the prediction of the outcome of treatment. However, this procedure was considered more efficient in the search for true prognostic variables than an automatic step up or step down elimination procedure. All tests and p-values are two-tailed.

Two charts of failure risk were built from the final predictive models obtained using logistic regression; they refer to the proportion predicted to fail with NPPV treatment at admission and after 2 h of NPPV. The proportions predicted to fail were calculated using the odds ratios (OR) derived from the models. The predicted risks of failure were graphically depicted in risk charts with changing colours showing the increasing level of risk, similar to the risk charts recently proposed by the European Societies of Cardiovascular Disease Prevention [21–25]. To graphically graduate risk intensity, four different colours were used according to the quartile of predicted risk probabilities. The aim was for the predicted risk of failure to be used as a rule to decide whether or not to intubate. To explore this possibility, arbitrary probabilities of failure equal to 25%, 50% and 75% were selected as possible cut-off values and sensitivity, specificity and percentage of correctly classified patients were calculated for each cut-off value.

The chart of risk for failure of NPPV was evaluated in an independent sample of 145 consecutive patients admitted during a 3-month period to three units (one general ward, one RIICU, and one ICU) to be treated with NPPV, due to an acute episode of COPD exacerbation.

Accuracy and generalisability

Accuracy, defined as the degree to which predictions match outcomes, and generalisability, defined as the ability to provide accurate predictions in different samples of patients [26], represent two important aspects of prognostic assessment and are essential if prognostic evaluation is aimed to improve clinical management of patients. Inaccuracy can be attributable either to errors in calibration (the predicted probabilities may result in a too high or too low) or to errors in discrimination (relative ranking of individuals may be out of order).

Concerning accuracy, the calibration of the model was assessed by evaluating expected to observed outcomes across deciles of risk, and discrimination was measured as the area under the receiver operating characteristic (ROC) curve (C statistic) [27].

Concerning generalisability, the predictive performance of the model was assessed through external validation, using an independent sample and comparing expected to observed

TABLE 1 Main characteristics of patients at the start of noninvasive positive pressure ventilation

	General ward	RIICU	ICU	All units
Patients	176 (17.0)	779 (75.4)	78 (7.6)	1033 (100)
Males	127 (72.2)	515 (66.4)	40 (72.7)	682 (66.0)
Age yrs	68.1±9.8	69.7±8.9	71.5±7.9	69.5±9.0
APACHE II score	18.2±4.5	20.3±6.1	25.7±6.2	20.3±5.9
Glasgow Coma Scale	14.3±1.5	13.0±2.4	12.3±3.0	13.2±2.3
Previous LTOT	99 (56.2)	329 (42.2)	18 (23.0)	446 (43.1)
Previous home MV	30 (17.0)	90 (11.5)	18 (23.0)	138 (13.4)
Hospitalisations·yr^{-1}	1.8±2.3	1.6±1.8	2.7±1.1	1.6±1.9
Chronic comorbidity	113 (64.2)	344 (44.2)	64 (82.1)	521 (50.4)
P_{a,O_2} mmHg	57.8±19.4	54.3±15.7	50.0±21.2	54.5±16.9
P_{a,CO_2} mmHg	75.9±16.3	80.7±16.9	88.0±18.3	80.4±16.9
pH	7.31±0.08	7.28±0.06	7.22±0.06	7.28±0.07
$P_{a,O_2}/F_{i,O_2}$	201.9±43.6	186.9±46.8	127.1±58.1	180.3±48.2
RR breaths·min^{-1}	28.9±5.6	29.0±6.3	33.9±5.7	29.4±6.1
Pneumonia	32 (18.3)	93 (12.0)	5 (6.4)	130 (12.6)

Data are presented as mean±SD or n (%). RIICU: respiratory intermediate intensive care units; ICU: intensive care unit; APACHE: acute physiology and chronic health evaluation; LTOT: long-term oxygen therapy; MV: mechanical ventilation; P_{a,O_2}: arterial oxygen tension; P_{a,CO_2}: carbon dioxide arterial tension; F_{i,O_2}: inspiratory oxygen fraction; RR: respiratory rate. 1 mmHg=0.133 kPa.

outcomes across deciles of risk, and in totals and measuring the area under the ROC curve [27].

RESULTS

A total of 1,033 patients were recruited in the 13 centres. All the patients were treated with NPPV for the management of ARF due to COPD exacerbation. The main demographic and physiological characteristics at the start of NPPV are shown in table 1.

Table 2 summarises the outcome data in the study population. NPPV was performed successfully in 797 patients (77.1%) until the normalisation of arterial blood gases. Among the 236 patients who failed, 185 patients were intubated (17.9% of the total and 78.4% of those for whom NPPV failed) and 51 died without intubation due to a previous ''do-not-intubate'' order. Among the 185 intubated patients, 94 patients successfully completed the treatment and 91 patients died. In total, 142 patients died (13.7% of the total and 60.2% of those for whom NPPV had failed). The intubation rates varied among the centres from 0 (zero intubated out of 72) to 72.2% (39 intubated out of 54). However, the rates were highest in the ICU at 50% and lower in general wards (13.6%) and in RIICU (15.7%).

Table 3 shows the means of variables which may be associated with a successful outcome. Successful patients were younger,

TABLE 2 Patient outcome (all units)

NPPV failure	236 (22.8)
Endotracheal intubation	185 (17.9)
Total deaths	142 (13.7)
Deaths among intubated patients	91 (49.2)
Hospital LOS days	15.2±9.4
ICU LOS days	13.0±7.9
MV days	8.0±6.7
NPPV h	59.9±59.3

Data are presented as n (%) or mean±SD. NPPV: noninvasive positive pressure ventilation; LOS: length of stay; ICU: intensive care unit; MV: mechanical ventilation.

TABLE 3 Univariate comparison between patients that succeeded or failed after treatment with noninvasive positive pressure ventilation

	Success	Failure	p-value
Subjects n	797	236	
Age	69.1±9.1	71.0±8.5	0.0041
Glasgow Coma Scale	13.7±1.9	11.6±3.0	<0.0001
APACHE II score	18.9±5.1	25.2±6.7	<0.0001
ABG at admission			
P_{a,O_2} mmHg	54.7±16.3	53.9±19.0	0.5331
P_{a,CO_2} mmHg	78.8±16.2	86.0±19.0	<0.0001
pH	7.29±0.06	7.25±0.08	<0.0001
RR	28.7±5.8	32.0±7.3	<0.0001
$P_{a,O_2}/F_{i,O_2}$	189.7±47.2	155.1±57.8	<0.0001
ABG after 2 h			
P_{a,O_2} mmHg	63.4±11.9	62.8±21.5	0.5972
P_{a,CO_2} mmHg	69.0±14.1	76.8±18.7	<0.0001
pH	7.34±0.05	7.27±0.1	<0.0001

Data are presented as mean±SD. APACHE: acute physiology and chronic health evaluation; ABG: arterial blood gases; P_{a,O_2}: arterial oxygen tension; P_{a,CO_2}: carbon dioxide arterial tension; RR: respiratory rate; F_{i,O_2}: inspiratory oxygen fraction.

TABLE 4 Logistic regression model used to build the risk chart

Variables	Admission			After 2 h		
	OR	p-value	95% CI	OR	p-value	95% CI
APACHE II score ⩾29	3.30	0.0001	1.81–6.01	4.79	0.0001	2.20–10.44
GCS 12–14	2.29	0.0008	1.41–3.72	1.93	0.0493	1.00–3.72
GCS ⩽11	4.40	<0.0001	2.59–7.49	5.16	<0.0001	2.54–10.50
pH <7.25	1.97	0.0046	1.23–3.15	21.02	<0.0001	10.07–43.87
pH 7.25–7.29	1.08	0.7511	0.68–1.72	2.92	0.0004	1.62–5.28
RR 30–34	1.83	0.0086	1.17–2.88	2.67	0.0021	1.43–4.99
RR ⩾35	2.66	<0.0001	1.66–4.25	4.95	<0.0001	2.64–9.29

OR: odds ratio; CI: confidence intervals; APACHE: acute physiology and chronic health evaluation; GCS: Glasgow Coma Scale; RR: respiratory rate.

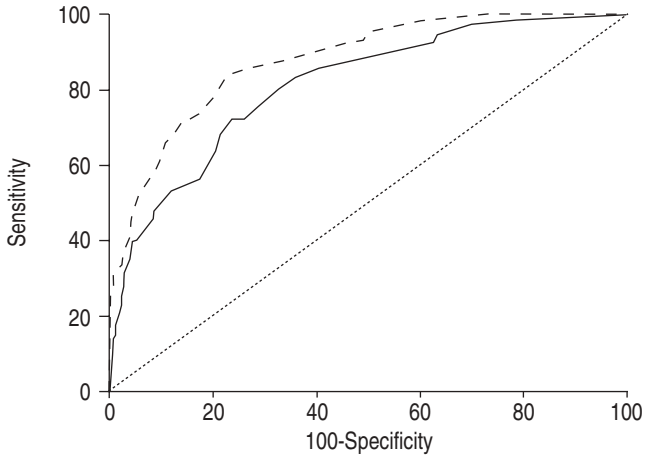

FIGURE 1. Receiver operating characteristic curves at admission (——) and after 2 h (- - -) of noninvasive positive pressure ventilation. ·····: line corresponds to a test which is useless as a discriminator between successes and failures.

less severe cases and had better values of P_{a,CO_2}, pH, RR, and oxygen arterial tension/inspiratory oxygen fraction.

Table 4 shows the final multivariable models including all the statistically significant predictors of NPPV failure. The variables measured on admission that were found to significantly increase the probability of NPPV failure were: APACHE II score ⩾29 (OR=3.30, p<0.001), GCS between 12–14 (OR=2.29, p<0.001) or GCS ⩽11 (OR=4.40, p<0.001), pH <7.25 (OR=1.97 p<0.05), and RR between 30–34 breaths·min^{-1} (OR=1.83, p<0.05) or >35 breaths·min^{-1} (OR=2.66, p<0.001). To allow comparison between admission and after 2 h, it was only possible to examine 594 patients out of 1,033 due to missing values, either for early failure or a lack of blood gas assessment. After 2 h of NPPV, the main factor influencing the outcome was the pH value: if pH <7.25 the OR for failure is 21.02 (p<0.0001), whereas if pH after 2 h is between 7.25–7.30, the OR is 2.92 (p<0.005). Also, all the other variables remained statistically significant: APACHE II score ⩾29 (OR=4.79, p<0.001), GCS between 12–14 (OR=1.93, p<0.05) or GCS ⩽11 (OR=5.16, p<0.001), RR between 30–34 breaths·min^{-1} (OR=2.67, p<0.001) or >35 breaths·min^{-1} (OR=4.95, p<0.001). These models were then used to predict the risk of NPPV failure. Accuracy of the predictions was good as there were no statistically significant differences between expected and observed failures when compared across deciles of risk (at admission: Chi-squared=3.46, degrees of freedom (dof)=8, p=0.9023; after 2 h: Chi-squared=1.33, dof=9, p=0.9982). The final models had a high predictive capability, as illustrated by the values of the C statistic: 0.81 at admission (95% confidence interval (CI): 0.78– 0.84) and 0.88 after 2 h (95% CI: 0.85–0.90). The C statistic, which is equal to the area under the ROC curve, is a measure of the ability of the model to discriminate between patients who failed or succeeded with the NPPV treatment. The comparison of the two ROC curves was statistically significant (p<0.001), implying that the ROC curve after 2 h of NPPV is more efficient for prediction (fig. 1).

The risk chart at admission (fig. 2) shows, in the top right corner, the lowest predicted probabilities of NPPV failing. These probabilities increase moving across the table and reach the highest peak in the bottom left corner. In detail, patients with pH >7.30, and APACHE II <29 have the lowest predicted probabilities of NPPV failing (~10%), probabilities which increase when RR increases and GCS decreases. In contrast,

		pH admission <7.25		pH admission 7.25–7.29		pH admission >7.30	
	RR	APACHE ≥29	APACHE <29	APACHE ≥29	APACHE <29	APACHE ≥29	APACHE <29
GCS 15	<30	29	11	18	6	17	6
	30–34	42	18	29	11	27	10
	≥35	52	24	37	15	35	14
GCS 12–14	<30	48	22	33	13	32	12
	30–34	63	34	48	22	46	21
	≥35	71	42	57	29	55	27
GCS ≤11	<30	64	35	49	23	47	21
	30–34	76	49	64	35	62	33
	≥35	82	59	72	44	70	42

FIGURE 2. Failure risk chart of noninvasive positive pressure ventilation at admission (the values in the table correspond to the percentage of patients who fail in each category). ■: 0–24%; ■: 25–49%; ■: 50–74%; ■: 75–100%. RR: respiratory rate; APACHE: acute physiology and chronic health evaluation II score; GCS: Glasgow Coma Scale.

		pH after 2 h <7.25		pH after 2 h 7.25–7.29		pH after 2 h ≥7.30	
	RR	APACHE ≥29	APACHE <29	APACHE ≥29	APACHE <29	APACHE ≥29	APACHE <29
GCS 15	<30	72	35	27	7	11	3
	30–34	88	59	49	17	25	7
	≥35	93	73	64	27	38	11
GCS 12–14	<30	84	51	41	13	19	5
	30–34	93	74	65	28	39	12
	≥35	96	84	78	42	54	20
GCS ≤11	<30	93	74	65	28	39	12
	30–34	97	88	83	51	63	26
	≥35	99	93	90	66	76	40

FIGURE 3. Failure risk chart of noninvasive positive pressure ventilation after 2 h (the values in the table correspond to the percentage of patients who fail in each category). ■: 0–24%; ■: 25–49%; ■: 50–74%; ■: 75–100%. RR: respiratory rate; APACHE: acute physiology and chronic health evaluation II score; GCS: Glasgow Coma Scale.

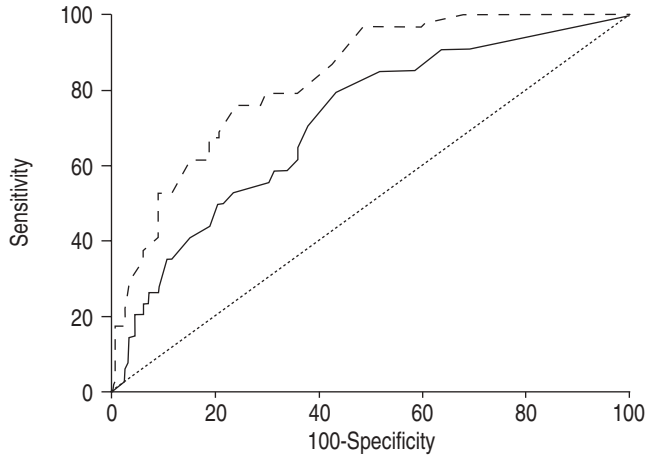

FIGURE 4. Receiver operating characteristic curves at admission (——) and after 2 h (- - -) of noninvasive positive pressure ventilation using the predicted failure percentages of the risk charts, applied to the 145 patients that were prospectively collected (independent sample). ·····: line corresponds to a test which is useless as a discriminator between successes and failures.

patients with pH <7.25, APACHE II ⩾29 and GCS <11 have a predicted risk of failure >70%.

The risk chart after 2 h of NPPV is interpreted similarly (fig. 3), but the predicted percentages of failure increase and are ~90% when pH is between 7.25–7.29, APACHE II is ⩾29 and GCS ⩾11. If pH is <7.25 after 2 h of ventilation, this greatly increases the risk of failure, especially if APACHE II is ⩾29.

The generalisability of these models was assessed through an external validation using an independent sample of 145 patients, and the results were encouraging. There were no statistically significant differences between expected and observed failures when compared across deciles of risk and in totals at admission (Chi-squared=2.96, dof=8, p=0.9371) and after 2 h (Chi-squared=3.53, dof=9, p=0.9397), and the C statistic was 0.71 at admission (95% CI: 0.63–0.79) and 0.83 after 2 h (95% CI: 0.76–0.89). Again, the comparison of the two ROC curves was statistically significant (p<0.001, fig. 4).

To explore the possibility of using the predicted risk of failure as a rule to decide whether or not to intubate, the following probability of failure was arbitrarily chosen: 25%, 50% and 75% as possible cut-off values and the corresponding scenarios are illustrated in table 5. Sensitivity decreases when risk of failure increases, whilst specificity increases when risk of failure increases. The percentage of patients correctly classified is quite good and reaches its maximum at a cut-off value of 50% (82 and 85%, respectively at admission and after 2 h).

TABLE 5 Sensitivity (%), specificity (%) and patients correctly classified (%) at admission and after 2 h

Risk of failure %	Measurement	Admission	After 2 h
25	Sensitivity	68.4 (59.9–76.1)	75.0 (66.9–82.0)
	Specificity	78.4 (74.3–82.1)	81.7 (77.9–85.1)
	Patients correctly classified	76.1	80.2
50	Sensitivity	33.1 (25.3–41.7)	52.9 (44.2–61.5)
	Specificity	96.7 (94.7–98.2)	94.1 (91.6–96.1)
	Patients correctly classified	82.1	84.7
75	Sensitivity	13.2 (8.0–20.1)	28.7 (21.3–37.1)
	Specificity	99.3 (98.1–99.9)	99.1 (97.8–99.8)
	Patients correctly classified	79.6	83.0

Data are presented as a percentage (95% confidence intervals).

DISCUSSION

Prediction of success/failure of NPPV in the "real world"

The present study evaluated a large unselected group of patients with acute exacerbation of COPD and respiratory acidosis treated with noninvasive ventilation in different care settings (ICUs, respiratory intermediate units and general wards; all were experienced care settings). The current authors used these data to assess the risk of NPPV failure and built two risk charts for admission and after 2 h of ventilation. These charts could be used to determine the possibility of failure or success of NPPV in patients with acute decompensate COPD and to help the clinical decision-making process. In particular, given that the Gold Initiative for Chronic Obstructive Lung Disease (GOLD) guidelines [28] suggested an initial trial of NPPV for most patients anyway, the decision to continue or not continue with NPPV can be greatly helped by the risk chart of failure tailored after 2 h of noninvasive ventilation.

The mortality rate in the current study is similar to that observed in the most quoted controlled trials on NPPV in carefully selected patients with acute exacerbation of COPD [1–3, 5]. The present study showed also that outside a specialised research setting, in the everyday clinical environment, NPPV can be effectively used in all COPD patients, such as those with pneumonia or with a lower GCS who might otherwise be excluded in a restricted clinical trial.

A series of randomised trials on noninvasive ventilation in acute exacerbations of COPD was published showing the effectiveness of NPPV, but in single ward settings. Nevertheless, the possibility to predict at hospital admission the chance of success of NPPV in patients with exacerbation of COPD should be based on a patient population with a variable degree of severity of illness. ANTON *et al*. [15] studied only 44 patients in a single centre and derived a regression equation using six parameters, including the baseline forced expiratory volume in one second and P_{a,CO_2} in a stable condition, that may not be available at all or not obtainable from older records at the time of admission. In contrast to ANTON *et al*. [15], the present authors preferred not to use change variables as they are more likely to indicate the rapidity with which the patient improves (and consequently the length of stay in the hospital) rather than the failure/success of NPPV. Moreover, given that, in practice, the two values are often not available for every patient, the inclusion of changes in the risk chart would have reduced its usefulness for prediction, as well as increasing the sampling error due to the reduction in sample size. Indeed, there was already a reduction in sample size from admission to 2 h in the present study, and to have included the changes in the levels would have caused the sample size at 2 h to have been even further reduced.

Data from a multicentre study performed in British respiratory general wards showed that if pH and/or P_{a,CO_2} improved after 1–4 h, successful NPPV was probable [9].

In the large population of unselected patients in the present study, on admission, severity of illness and neurological impairment, respectively (assessed by APACHE II score and GCS), were found to influence significantly the probability of NPPV failure together with RR and pH. Clear differences were found between the patients who succeeded and those who failed, with respect to the predictor variables observed at admission and after 2 h of treatment. In particular, after 2 h of NPPV, the main factor influencing the outcome was the pH value. Since pH seems to be a very important variable in predicting failure of NPPV and there is much discussion about which is the correct cut off to choose, the current authors evaluated in both models a pH <7.25 and a pH $\geqslant 7.25$ but <7.30, even if the latter was not statistically significant at admission. However, given that the decision to intubate is based on clinical judgement, and this judgement is also probably based on the values of predictor variables, there is a danger of falling into a circular argument. For example, patients with pH <7.2 tend to be judged to be in need of intubation and, therefore, pH <7.2 significantly predicts intubation. However, in the current analysis the cut-off values of the predictor variables were decided *a priori*, which should reduce this effect.

The presence of pneumonia, at admission and after 2 h, did not significantly influence the probability of NPPV failure in the patients of the present study. AMBROSINO *et al*. [13] reported a high failure rate in nine COPD patients with pneumonia in a retrospective single centre study, but this was not confirmed by either a recent meta-analysis [8] or a prospective, randomised controlled study on NPPV in community-acquired pneumonia [29]. In agreement with the current study, previous studies have reported that severity of illness scores [2, 14, 30] and pH [1, 2, 5, 31, 32] at enrolment are independent risk factors for NPPV failure in patients with exacerbations of COPD. A reduction in RR with NPPV has been shown in a number of studies, with larger falls generally being associated with a successful outcome [2, 3, 13, 14]. Clearly, it is not only clinical status that can improve the outcome of NPPV *e.g.* intolerance of the mask can lead to failure of NPPV [10]. However, this cannot be predicted for individual patients before starting NPPV; the patient should try NPPV for a minimum of 1–2 h, adjusting the mask and changing the interface. It should be noted that the present study population included a fairly high percentage of COPD patients, mechanically ventilated at home. This is more typical for Italy than other European Countries, and it could have some impact on the generalisability of results as these patients are "experts" at NPPV.

Applicability of predictive risk charts

The charts of risk for NPPV failure showed a good generalisability assessed on an independent sample of patients with acute exacerbation of COPD. Calculation of the area under the ROC curves showed a high accuracy of the prediction, and the rate of the patients correctly classified is >75% at each cut-off value arbitrarily chosen. Since prediction improves after 2 h of NPPV and the GOLD guidelines suggest an initial trial of NPPV for most patients anyway, the decision to continue or not continue with NPPV can be greatly helped by a risk chart of failure tailored after 2 h of noninvasive ventilation. This discussion should also focus on which cut-off value is the best choice *i.e.* above which risk of failure it is suggested to intubate. At one extreme, the risk of failure can be considered to be $\geqslant 0\%$, which implies that all the patients will fail with NPPV treatment and will be intubated, with a sensitivity of 100% and a specificity of 0%. At the other extreme, the risk of failure can be considered to be equal to

100%, which implies that all the patients will succeed with NPPV treatment and nobody will be intubated, with a sensitivity of 0% and a specificity of 100%. Clearly, these two options are only hypothetical; more realistic ones are depicted in table 5. The risk of failure equal to 25% appears as a good option, as sensitivity and specificity are both at high levels, particularly after 2 h. However, a risk of failure equal to 25% can be considered as a low risk and several people will be indicated for intubation without really needing it (false positive). If the risk of failure of NPPV equals 75%, fewer people will be indicated for intubation, but some people who really needed it could be missed (false negative). Thus, perhaps a risk of failure equal to 50% could be a useful compromise for whoever wants to start using the predictive chart of risk, keeping in mind that this cut-off value could be changed if clinical experience suggests it.

Furthermore, it would be interesting to assess if both risk charts (admission and after 2 h) are useful in the clinical practice. In this case, the risk charts will be used consecutively, which will maximise specificity, but lower sensitivity. One ends up surer that positive test results represent failure of NPPV, but there is an increased risk of missing people who really need intubation.

Indeed, the GOLD guidelines suggest trying NPPV initially on all patients and only after 2 h of NPPV can a decision then be made whether to continue or not. In this case, only the risk chart of failure tailored after 2 h of noninvasive ventilation can be used, but it should greatly help the decision.

Conclusions

The efficacy of noninvasive positive pressure ventilation in acute exacerbation of chronic obstructive pulmonary disease is so well documented that international guidelines [28] recommend it as the first choice treatment of acute respiratory failure with respiratory acidosis. Nevertheless, given that noninvasive positive pressure ventilation is used in a variety of care settings, it may be important to know the likelihood of failure to decide upon the best choice between noninvasive positive pressure ventilation and endotracheal intubation instituted earlier. The prediction charts are based on data collected by the present authors and take into consideration all relevant clinical prognostic indicators and are derived from a population representing the patients seen routinely in clinical practice. Thus, the current authors think they could greatly help the decision on clinical management of the patient. Using the chart, it is possible to predict ''*a priori*'' the probability of noninvasive positive pressure ventilation failure and reduce the useless and prolonged use of noninvasive positive pressure ventilation in patients with respiratory acidosis due to chronic obstructive pulmonary disease exacerbation.

ACKNOWLEDGEMENTS

The authors would like to thank the Eleonora Lorillard Spencer Cenci Foundation for its contribution to the research.

The following physicians participated in the Italian NPPV study group: A. Bellone, R. Della Porta, A. LoCoco, M. Demsar, G. Gadaleta, M. Moretti, P. Parigi, G. Pinelli, M. Piattella, A. Potena, A. Rossi, F. Turati, A. Vianello, M. Vitacca.

REFERENCES

1 Bott J, Carroll MP, Conway JH, *et al.* Randomized controlled trial of nasal ventilation in acute ventilatory failure due to chronic obstructive airways disease. *Lancet* 1993; 341: 1555–1557.
2 Brochard L, Mancebo J, Wysocki M, *et al.* Noninvasive ventilation for acute exacerbations of chronic obstructive pulmonary disease. *N Engl J Med* 1995; 333: 817–822.
3 Kramer N, Meyer TJ, Meharg J, Cece RD, Hill NS. Randomized, prospective trial of noninvasive positive pressure ventilation in acute respiratory failure. *Am J Respir Crit Care Med* 1995; 151: 1799–1806.
4 Celikel T, Sungur M, Ceyhan B, Karakurt S. Comparison of noninvasive positive pressure ventilation with standard medical therapy in hypercapnic acute respiratory failure. *Chest* 1998; 114: 1636–1642.
5 Plant PK, Owen JL, Elliott MW. Early use of non-invasive ventilation for acute exacerbations of chronic obstructive pulmonary disease on general respiratory wards: a multicentre randomised controlled trial. *Lancet* 2000; 355: 1931–1935.
6 Conti G, Antonelli M, Navalesi P, *et al.* Noninvasive *versus* conventional mechanical ventilation in patients with chronic obstructive pulmonary disease after failure of medical treatment in the ward: a randomized trial. *Intensive Care Med* 2002; 28: 1701–1707.
7 Peter JV, Moran JL, Phillips-Hughes J, Warn D. Noninvasive ventilation in acute respiratory failure. A meta-analysis update. *Crit Care Med* 2002; 30: 555–562.
8 Lightowler JV, Wedzicha JA, Elliott MW, Ram FSF. Non-invasive positive pressure ventilation to treat respiratory failure resulting from exacerbations of chronic obstructive pulmonary disease: Cochrane systematic review and meta-analysis. *BMJ* 2003; 326: 185–190.
9 Plant PK, Owen JL, Elliott MW. Non-invasive ventilation in acute exacerbations of chronic obstructive pulmonary disease: long term survival and predictors of in-hospital outcome. *Thorax* 2001; 56: 708–712.
10 Carlucci A, Richard JC, Wysocki M, Lepage E, Brochard L. Noninvasive *versus* conventional mechanical ventilation. An epidemiologic survey. *Am J Respir Crit Care Med* 2001; 163: 874–880.
11 American Thoracic Society, European Respiratory Society, European Society of Intensive Care Medicine and Societè de Reanimation de Langue Francaise. International Consensus Conference in Intensive Care Medicine: positive pressure ventilation in acute respiratory failure. *Am J Respir Crit Care Med* 2001; 163: 281–291.
12 British Thoracic Society Standards of Care Committee. Non-invasive ventilation in acute respiratory failure. *Thorax* 2002; 57: 192–211.
13 Ambrosino N, Foglio K, Rubini F, Clini E, Nava S, Vitacca M. Noninvasive mechanical ventilation in acute respiratory failure due to chronic obstructive pulmonary disease: correlates for success. *Thorax* 1995; 50: 755–757.
14 Soo Hoo GW, Santiago S, Williams J. Nasal mechanical ventilation for hypercapnic respiratory failure in chronic obstructive pulmonary disease: determinants for success and failure. *Crit Care Med* 1994; 27: 417–434.
15 Anton A, Guell R, Gomez J, *et al.* Predicting the result of noninvasive ventilation in severe acute axacerbations of

patients with chronic airflow limitation. *Chest* 2000; 117: 823–833.
16 Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies and the hierarchy of research designs. *N Engl J Med* 2000; 342: 1887–1892.
17 Confalonieri M, Gorini M, Ambrosino N, Mollica C, Corrado A. Respiratory intensive care units in Italy: a national census and prospective cohort study. *Thorax* 2001; 56: Suppl. 5 pt 2, 373–378.
18 American Thoracic Society. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1995; 152: S77–S121.
19 Knaus WA, Draper EA, Wagner DP, *et al.* APACHE II: a severity of disease. classification system. *Crit Care Med* 1985; 13: 818–829.
20 Teasdale G, Jennet B. Assessment of coma and impaired consciousness: a practical scale. *Lancet* 1974; 2: 81–84.
21 Marchioli R, Avanzini F, Barzi F, *et al.* Assessment of absolute risk of death after myocardial infarction by use of multiple-risk-factor assessment equation: GISSI-Prevention mortality risk chart. *Eur Heart J* 2001; 22: 2085–2103.
22 Wood D, De Backer G, Faergeman O, Graham I, Mancia G, Pyörälä K, together with members of the Task Force. Prevention of coronary heart disease in clinical practice. Recommendations of the Second Joint Task Force of European and other Societies on coronary prevention. *Eur Heart J* 1998; 19: 1434–1503.
23 Pyörälä K, De Backer G, Graham I, Poole Wilson P, Wood D, on behalf of the Task Force. Prevention of coronary heart disease in clinical practice. Recommendations of the Task Force of the European Society of Cardiology, European Atherosclerosis Society and European Society of Hypertension. *Eur Heart J* 1994; 15: 1300–1331.
24 Anderson KM, Wilson PWF, Odell PM, Kannel WB. An updated coronary risk profile: a statement for health professionals. *Circulation* 1991; 83: 356–362.
25 Conroy RM, Pyörälä K, Fitzgerald AP, *et al.* Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J* 2003; 24: 987–1003.
26 Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med* 1999; 130: 515: 24.
27 Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; 143: 29–36.
28 National Heart, Lung and Blood Institute/World Health Organization Report. Global Initiative for Chronic Obstructive Lung Disease, 2001. NIH Publication No. 2701: 91–92.
29 Confalonieri M, Potena A, Carbone G, Della Porta R, Tolley E, Meduri GU. Acute respiratory failure in patients with severe community-acquired pneumonia: a prospective randomized evaluation of noninvasive ventilation. *Am J Respir Crit Care Med* 1999; 160: 1585–1591.
30 Confalonieri M, Aiolfi S, Gandola L, Scartabellati A, Della Porta R, Parigi P. Severe exacerbations of chronic obstructive pulmonary disease treated with BiPAP® by nasal mask. *Respiration* 1994; 61: 310–316.
31 Meduri GU, Turner RE, Abou-Shala N, Wunderink R, Tolley E. Noninvasive positive pressure ventilation *via* face mask. First-line intervention in patients with acute hypercapnic and hypoxemic respiratory failure. *Chest* 1996; 109: 179–193.
32 Meduri GU, Abou-Shala N, Fox RC, Jones CB, Leeper KV, Wunderink RG. Noninvasive face mask mechanical ventilation in patients with acute hypercapneic respiratory failure. *Chest* 1991; 100: 445–454.