


Comparison of prognosis predictive value of 4 disease severity scoring systems in patients with acute respiratory failure in intensive care unit

A STROBE report

Wen-Cheng Huang, MD^a, Hong-Jian Xie, MD^b, Hong-Tao Fan, MD^a, Mei-Hao Yan, MD^a, Yuan-Cheng Hong, MD^{a,*}

Abstract

Various disease severity scoring systems were currently used in critically ill patients with acute respiratory failure, while their performances were not well investigated.

The study aimed to investigate the difference in prognosis predictive value of 4 different disease severity scoring systems in patients with acute respiratory failure.

With a retrospective cohort study design, adult patients admitted to intensive care unit (ICU) with acute respiratory failure were screened and relevant data were extracted from an open-access American intensive care database to calculate the following disease severity scores on ICU admission: acute physiology score (APS) III, Sequential Organ Failure Assessment score (SOFA), quick SOFA (qSOFA), and Oxford Acute Severity of Illness Score (OASIS). Hospital mortality was chosen as the primary outcome. Multivariable logistic regression analyses were performed to analyze the association of each scoring system with the outcome. Receiver operating characteristic curve analyses were conducted to evaluate the prognosis predictive performance of each scoring system.

A total of 4828 patients with acute respiratory failure were enrolled with a hospital mortality rate of 16.78%. APS III (odds ratio [OR] 1.03, 95% confidence interval [CI] 1.02–1.03), SOFA (OR 1.15, 95% CI 1.12–1.18), qSOFA (OR 1.26, 95% CI 1.11–1.42), and OASIS (OR 1.06, 95% CI 1.05–1.08) were all significantly associated with hospital mortality after adjustment for age and comorbidities. Receiver operating characteristic analyses showed that APS III had the highest area under the curve (AUC) (0.703, 95% CI 0.683–0.722), and SOFA and OASIS shared similar predictive performance (area under the curve 0.653 [95% CI 0.631–0.675] and 0.664 [95% CI 0.644–0.685], respectively), while qSOFA had the worst predictive performance for predicting hospital mortality (0.553, 95% CI 0.535–0.572).

These results suggested the prognosis predictive value varied among the 4 different disease severity scores for patients admitted to ICU with acute respiratory failure.

Abbreviations: APACHE = Acute Physiology and Chronic Health Evaluation, APS = acute physiology score, ARDS = acute respiratory distress syndrome, AUC = area under the curve, CI = confidence interval, ICD = International Classification of Diseases, ICU = intensive care unit, OASIS = Oxford Acute Severity of Illness Score, OR = odds ratio, qSOFA = quick SOFA, ROC = receiver operating characteristic, SOFA = Sequential Organ Failure Assessment.

Keywords: APACHE, critical care, organ dysfunction scores, respiratory insufficiency, severity of illness index

Editor: Khaled Saleh.

WCH and HJX contributed equally to this work.

The access to the database has been approved by the institutional review boards of both Beth Israel Deaconess Medical Center and Massachusetts Institute of Technology Affiliates after completing the CITI (Collaborative Institutional Training Initiative) "Data or Specimens Only Research" course. No informed consent was required on the de-identified patients.

Patient consent for publication is not applicable.

The authors have no funding and conflicts of interest to disclose.

The data that support the findings of this study are available from a third party, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are available from the authors upon reasonable request and with permission of the third party.

^aDepartment of Respiratory Medicine, The 910th Hospital of People's Liberation Army, Quanzhou, Fujian, People's Republic of China, ^bDepartment of Respiratory Medicine, Quanzhou Guangqian Hospital, Quanzhou, Fujian, People's Republic of China.

*Correspondence: Yuan-Cheng Hong, Department of Respiratory Medicine, The 910th Hospital of People's Liberation Army, No. 180 Huayuan Road, Fengze District, Quanzhou 362000, Fujian, People's Republic of China (e-mail: qzhongyuancheng@126.com).

Copyright © 2021 the Author(s). Published by Wolters Kluwer Health, Inc.

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial License 4.0 (CCBY-NC), where it is permissible to download, share, remix, transform, and buildup the work provided it is properly cited. The work cannot be used commercially without permission from the journal.

How to cite this article: Huang WC, Xie HJ, Fan HT, Yan MH, Hong YC. Comparison of prognosis predictive value of 4 disease severity scoring systems in patients with acute respiratory failure in intensive care unit: a STROBE report. *Medicine* 2021;100:39(e27380).

Received: 11 December 2020 / Received in final form: 23 July 2021 / Accepted: 14 September 2021

<http://dx.doi.org/10.1097/MD.00000000000027380>

1. Introduction

Several disease severity scoring systems have been used in clinical practice for prognostic evaluation of intensive care unit (ICU) patients, such as the Acute Physiology and Chronic Health Evaluation (APACHE) system, the Sequential Organ Failure Assessment (SOFA) score, and the quick SOFA (qSOFA) score.^[1,2] Some new disease severity scoring systems are also being developed, such as the Oxford Acute Severity of Illness Score (OASIS).^[3] Numerous disease severity scoring systems on one hand provide more options for predicting the prognosis of ICU patients, but on the other hand, which scoring system should be chosen becomes a new question. For ICU patients, early and accurate prognosis prediction and risk stratification are significant,^[4] so are patients with acute respiratory failure, especially acute respiratory distress syndrome (ARDS).^[5] As we know, acute respiratory failure, one of the common diseases of critically ill patients in ICU^[6] is usually caused by acute pathogenic factors, such as severe lung diseases, trauma, shock, electric shock, acute airway obstruction, which lead to rapid decline in lung function. Since the body's compensation cannot occur in such a short period of time, such patients need to be rescued in time.^[7] Thus, it is crucial to predict the prognosis of acute respiratory failure. However, unlike the SOFA score for sepsis,^[8] it is still unclear whether different disease severity scoring systems have different prognosis predictive value for ICU patients with acute respiratory failure. In order to provide research evidence about scoring system selection for acute respiratory failure, the study analyzed the prognosis predictive power of 4 different disease severity scoring systems in ICU patients with acute respiratory failure.

2. Materials and methods

2.1. Study design and population

This study was a retrospective cohort study. Patients with acute respiratory failure were screened from the Medical Information Mart for Intensive Care III, an open-access database comprising de-identified health-related data associated with over 40,000 patients who stayed in critical care units of the Beth Israel Deaconess Medical Center between 2001 and 2012.^[9] Those who met the following criteria were included and analyzed: age \geq 18 years old; a diagnosis of acute respiratory failure (identified by the International Classification of Diseases [ICD], Ninth Revision, Clinical Modification code 5173, 5185, 51851, 51852, 51853, 51881, 51882, 51883, 51884, 7991, V461, V4611, V4612, V4613, V4614, and V462); first hospital admission in the database; and need of oxygen therapy on the first day of ICU admission. The access to the database has been approved by the institutional review boards of both Beth Israel Deaconess Medical Center and Massachusetts Institute of Technology Affiliates after completing the Collaborative Institutional Training Initiative "Data or Specimens Only Research" course. No informed consent was required on the de-identified patients.

2.2. Exposures, covariates, and outcome

For all included patients, the following data were extracted: age, sex, comorbidities, and clinical outcomes. Variables collected within the first 24 hours after ICU admission were also extracted to calculate the studied scoring systems. Each patient's acute physiology score (APS, extracted from the APACHE III system),

SOFA, qSOFA, and OASIS score on the first 24 hours after ICU admission were calculated according to specific calculation requirements of each scoring system.^[3,8,10] Missing components for the calculation of each scoring system were treated as normal (usually 0). Hospital mortality was selected as the primary outcome. The data extraction processes were performed mainly using structured query language codes from the Medical Information Mart for Intensive Care Code Repository.^[11]

2.3. Statistical analysis

Statistical analysis was performed by Empower(R) (www.empowerstats.com; X&Y solutions, Inc., Boston, MA) and R software, version 3.4.3 (http://www.r-project.org; R Foundation for Statistical Computing, Vienna, Austria). Continuous variables were expressed as mean \pm standard deviation or median (25% quantile–75% quantile), and comparisons between groups were examined by a *t* test or rank sum test. Categorical variables were expressed as numbers and percentages, and comparisons between groups were examined by chi-square test or Fisher exact test. To adjust potential confounders when evaluating the association of each score with the outcome, variables with $P < .1$ in the univariable logistic analysis were considered to be included in the multivariable logistic regression model for multivariable analysis and the odds ratio (OR) was calculated. To evaluate the predictive power of each score for the outcome, the receiver operating characteristic (ROC) curve was plotted and the area under the curve (AUC) was also calculated. A *P* value less than .05 was considered statistically significant.

3. Results

3.1. Basic characteristics of the patients

Four thousand eight hundred twenty-eight patients admitted to ICU with acute respiratory failure were included finally. The mean age of the patients was 63.34 ± 17.21 years old, among which the male accounted for 53.44%. The hospital mortality was 16.78% (810/4828). Compared with the survivors, the non-survivors had a higher median APS III, SOFA, and OASIS score, but the same median qSOFA score (Table 1). The 5 most common comorbidities were coagulopathy (17.15%), renal failure (12.12%), other neurological disease (11.64%), alcohol abuse (10.60%), and hypertension (10.56%).

3.2. Logistic regression analysis

Univariable logistic analysis showed that age, APS III, SOFA, qSOFA, OASIS, sepsis, Elixhauser Comorbidity Index (SID30), congestive heart failure, cardiac arrhythmias, peripheral vascular disorder, paralysis, uncomplicated diabetes, renal failure, liver disease, lymphoma, metastatic cancer, solid tumor, coagulopathy, obesity, fluid and electrolyte disorders, deficiency anemia, alcohol abuse, drug abuse, psychoses, and depression significantly associated with hospital mortality in patients with acute respiratory failure (Table 2).

After adjusted for age, sepsis, and Elixhauser Comorbidity Index (SID30) (model I), APS III, SOFA, qSOFA, and OASIS were all significantly associated with hospital mortality. Results were similar when covariates with a *P* value less than .1 in the univariable logistic regression were adjusted (model II), as shown in Table 3.

Table 1**Basic characteristics of the included patients.**

Variables	Total (n = 4828)	Survivors (n = 4018)	Non-survivors (n = 810)	P
Age, yrs	63.34 ± 17.21	62.00 ± 17.47	70.00 ± 14.13	<.001
Male	2580 (53.44%)	2154 (53.61%)	426 (52.59%)	.597
Severity score				
APS III	46 (34–62)	43.50 (33–59)	60 (46–79.75)	<.001
SOFA	5 (4–8)	5 (3–8)	7 (5–11)	<.001
qSOFA	2 (2–2)	2 (2–2)	2 (2–2)	<.001
OASIS	38 (33–43)	37 (33–42)	42 (37–47)	<.001
Length of hospital stay, days	13.39 (7.71–22.60)	13.81 (7.98–22.73)	11.49 (5.95–21.21)	<.001
Sepsis	2970 (61.52%)	2394 (59.58%)	576 (71.11%)	<.001
Elixhauser Comorbidity Index (SID30)	12 (4–22)	11 (3–21)	19 (11–28)	<.001
Comorbidities				
Congestive heart failure	1118 (23.16%)	902 (22.45%)	216 (26.67%)	.009
Cardiac arrhythmias	1044 (21.62%)	811 (20.18%)	233 (28.77%)	<.001
Valvular disease	283 (5.86%)	224 (5.57%)	59 (7.28%)	.059
Pulmonary circulation disorder	286 (5.92%)	236 (5.87%)	50 (6.17%)	.742
Peripheral vascular disorder	362 (7.50%)	286 (7.12%)	76 (9.38%)	.026
Hypertension	510 (10.56%)	416 (10.35%)	94 (11.60%)	.290
Paralysis	188 (3.89%)	172 (4.28%)	16 (1.98%)	.002
Other neurological disease	562 (11.64%)	481 (11.97%)	81 (10.00%)	.111
Chronic pulmonary disease	1203 (24.92%)	1007 (25.06%)	196 (24.20%)	.604
Uncomplicated diabetes	924 (19.14%)	801 (19.94%)	123 (15.19%)	.002
Complicated diabetes	267 (5.53%)	222 (5.53%)	45 (5.56%)	.972
Hypothyroidism	405 (8.39%)	339 (8.44%)	66 (8.15%)	.787
Renal failure	585 (12.12%)	464 (11.55%)	121 (14.94%)	.007
Liver disease	337 (6.98%)	247 (6.15%)	90 (11.11%)	<.001
Peptic ulcer	5 (0.10%)	5 (0.12%)	0 (0.00%)	.597
AIDS	33 (0.68%)	25 (0.62%)	8 (0.99%)	.244
Lymphoma	59 (1.22%)	43 (1.07%)	16 (1.98%)	.051
Metastatic cancer	238 (4.93%)	148 (3.68%)	90 (11.11%)	<.001
Solid tumor	134 (2.78%)	102 (2.54%)	32 (3.95%)	.026
Rheumatoid arthritis	133 (2.75%)	105 (2.61%)	28 (3.46%)	.181
Coagulopathy	828 (17.15%)	605 (15.06%)	223 (27.53%)	<.001
Obesity	386 (8.00%)	355 (8.84%)	31 (3.83%)	<.001
Weight loss	339 (7.02%)	275 (6.84%)	64 (7.90%)	.283
Fluid and electrolyte disorders	2072 (42.92%)	1690 (42.06%)	382 (47.16%)	.007
Blood loss anemia	119 (2.46%)	105 (2.61%)	14 (1.73%)	.138
Deficiency anemia	1048 (21.71%)	921 (22.92%)	127 (15.68%)	<.001
Alcohol abuse	512 (10.60%)	456 (11.35%)	56 (6.91%)	<.001
Drug abuse	271 (5.61%)	257 (6.40%)	14 (1.73%)	<.001
Psychoses	268 (5.55%)	246 (6.12%)	22 (2.72%)	<.001
Depression	410 (8.49%)	370 (9.21%)	40 (4.94%)	<.001

Patients were grouped as survivors and non-survivors determined by hospital mortality status. Statistical significance ($P < .05$) is shown in bold.

AIDS = acquired immune deficiency syndrome, APS = acute physiology score, OASIS = Oxford Acute Severity of Illness Score, qSOFA = quick SOFA, SOFA = Sequential Organ Failure Assessment score.

3.3. ROC curve analysis

ROC curve analysis showed that the AUCs of APS III, SOFA, qSOFA, and OASIS for predicting hospital mortality were 0.703 (95% confidence interval [CI] 0.683–0.722), 0.653 (95% CI 0.631–0.675), 0.553 (95% CI 0.535–0.572), and 0.664 (95% CI 0.644–0.685), respectively. As shown in Figure 1 and Table 4, APS III had the best predictive performance and qSOFA had the worst predictive performance, while SOFA and OASIS shared similar predictive performance.

4. Discussion

In order to investigate which scoring system is better for acute respiratory failure, the study used data from an electronic database to evaluate prognosis predictive values of 4 disease severity scoring systems for ICU patients admitted with acute

respiratory failure. Results of the study showed that APS III had the highest prognosis predictive power while qSOFA had the worst performance. As we know, there is no research reported currently that examined the value of the above 4 scoring systems together for predicting the prognosis of ICU patients admitted with acute respiratory failure. Results of our study provide evidence for the selection of prognosis prediction scoring systems for patients admitted to ICU with acute respiratory failure.

Assessment of disease severity is very important for critically ill patients,^[12] and numerous disease severity scoring systems have been developed and used. All the existing scoring systems can be roughly categorized as either disease-specific scoring or generic scoring.^[13] The main difference between these 2 types of scoring systems is that a disease-specific scoring system is designed to be applied to a specific disease, such as the Ranson criteria for acute pancreatitis,^[14] while a generic scoring system is usually

Table 2
Univariable logistic regression analysis of hospital mortality in patients with acute respiratory failure.

Variables	OR	95% CI	P
Age	1.03	1.03–1.04	<.001
Sex			
Male	1.0		
Female	1.04	0.90–1.21	.597
Severity score			
APS III	1.03	1.03–1.03	<.001
SOFA	1.17	1.15–1.20	<.001
qSOFA	1.36	1.21–1.52	<.001
OASIS	1.08	1.07–1.09	<.001
Sepsis			
No	1.0		
Yes	1.67	1.42–1.97	<.001
Elixhauser Comorbidity Index (SID30)	1.04	1.03–1.05	<.001
Comorbidities			
Congestive heart failure			
No	1.0		
Yes	1.26	1.06–1.49	.010
Cardiac arrhythmias			
No	1.0		
Yes	1.60	1.35–1.89	<.001
Valvular disease			
No	1.0		
Yes	1.33	0.99–1.79	.060
Pulmonary circulation disorder			
No	1.0		
Yes	1.05	0.77–1.44	.742
Peripheral vascular disorder			
No	1.0		
Yes	1.35	1.04–1.76	.026
Hypertension			
No	1.0		
Yes	1.14	0.90–1.44	.291
Paralysis			
No	1.0		
Yes	0.45	0.27–0.76	.003
Other neurological disease			
No	1.0		
Yes	0.82	0.64–1.05	.111
Chronic pulmonary disease			
No	1.0		
Yes	0.95	0.80–1.14	.604
Uncomplicated diabetes			
No	1.0		
Yes	0.72	0.58–0.88	.002
Complicated diabetes			
No	1.0		
Yes	1.01	0.72–1.40	.972
Hypothyroidism			
No	1.0		
Yes	0.96	0.73–1.27	.787
Renal failure			
No	1.0		
Yes	1.35	1.08–1.67	.007
Liver disease			
No	1.0		
Yes	1.91	1.48–2.46	<.001
AIDS			
No	1.0		
Yes	1.59	0.72–3.54	.254
Lymphoma			
No	1.0		

(continued)

Table 2
(continued).

Variables	OR	95% CI	P
Yes	1.86	1.04–3.32	.035
Metastatic cancer			
No	1.0		
Yes	3.27	2.49–4.30	<.001
Solid tumor			
No	1.0		
Yes	1.58	1.05–2.37	.027
Rheumatoid arthritis			
No	1.0		
Yes	1.33	0.87–2.04	.182
Coagulopathy			
No	1.0		
Yes	2.14	1.80–2.56	<.001
Obesity			
No	1.0		
Yes	0.41	0.28–0.60	<.001
Weight loss			
No	1.0		
Yes	1.17	0.88–1.55	.283
Fluid and electrolyte disorders			
No	1.0		
Yes	1.23	1.06–1.43	.008
Blood loss anemia			
No	1.0		
Yes	0.66	0.37–1.15	.141
Deficiency anemia			
No	1.0		
Yes	0.63	0.51–0.77	<.001
Alcohol abuse			
No	1.0		
Yes	0.58	0.43–0.77	<.001
Drug abuse			
No	1.0		
Yes	0.26	0.15–0.44	<.001
Psychoses			
No	1.0		
Yes	0.43	0.27–0.67	<.001
Depression			
No	1.0		
Yes	0.51	0.37–0.72	<.001

Statistical significance ($P < .05$) is shown in bold.

AIDS=acquired immune deficiency syndrome, APS=acute physiology score, CI=confidence interval, OASIS=Oxford Acute Severity of Illness Score, OR=odds ratio, qSOFA=quick SOFA, SOFA=Sequential Organ Failure Assessment score.

developed to be used in various clinical conditions. Although a disease-specific scoring system might have a better performance for prognosis prediction for its specific disease, but in ICU it might fail to perform as well as it does in general ward, since critically ill patients are usually comorbid with several disorders, thus generic scoring systems are more popular in ICU conditions.^[15] This is why in our study we investigated the 4 scoring systems, 3 of which are already widely used in ICU settings, and 1 of which is a newly developed generic scoring system.^[16,17] APS III is extracted from the APACHE III system, a widely used scoring system in ICU, which was first proposed by Knaus et al^[18] in 1981 and has been updated several times and validated in ICU conditions.^[19] OASIS was actually developed in 2013 from the APACHE IV system by screening and simplifying its parameters.^[3] Unlike the above 2 generic scoring systems,

Table 3
Multivariable logistic regression analysis of hospital mortality in patients with acute respiratory failure.

	Crude			Model I			Model II		
	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P
APS III	1.03	1.03–1.03	<.0001	1.03	1.03–1.03	<.0001	1.03	1.02–1.03	<.0001
SOFA	1.17	1.15–1.20	<.0001	1.16	1.13–1.19	<.0001	1.15	1.12–1.18	<.0001
qSOFA	1.36	1.21–1.52	<.0001	1.22	1.08–1.38	.0011	1.26	1.11–1.42	.0003
OASIS	1.08	1.07–1.09	<.0001	1.06	1.05–1.07	<.0001	1.06	1.05–1.08	<.0001

Model I: age, sepsis, and Elixhauser Comorbidity Index (SID30) were adjusted; Model II: age, sepsis, Elixhauser Comorbidity Index (SID30), congestive heart failure, cardiac arrhythmias, valvular disease, peripheral vascular disorder, paralysis, uncomplicated diabetes, renal failure, liver disease, lymphoma, metastatic cancer, solid tumor, coagulopathy, obesity, fluid and electrolyte disorders, deficiency anemia, alcohol abuse, drug abuse, psychoses, and depression were adjusted.

APS=acute physiology score, CI=confidence interval, OASIS=Oxford Acute Severity of Illness Score, OR=odds ratio, qSOFA=quick SOFA, SOFA=Sequential Organ Failure Assessment score.

SOFA and qSOFA are developed for sepsis (but has been repurposed to predict patient outcomes),^[18,20] and SOFA was created much earlier than qSOFA, which has been reported to be closely related to organ dysfunction and hospital survival.^[20,21] In addition to the popularity, the difference in the way the development of the scores and the complexity of parameters of each score is also one of the reasons why we chose and study these 4 scoring systems. APS III is the acute part of APACHE III system, which consists of 17 physiological variables, acid-base disturbance, age, and the Glasgow Coma Scale.^[10] It was developed from a general ICU patient population using multivariable logistic regression. SOFA, however, was developed based on expert opinion which incorporates organ function scores from 6 organ systems,^[20] and qSOFA was introduced by the Sepsis-3 group in 2016 as a simplified version of SOFA.^[22] Except for qSOFA, laboratory testing is necessary for the calculation of APS III and SOFA. Given that an ideal scoring system should be simple but with good predictive performance, Johnson et al^[3] developed OASIS using machine-learning algorithms, which consists of only 10 parameters. In this way, our study covered scoring systems which are widely used in practice but were developed by different methods and with different complexity of parameters.

There are several studies that specially investigated performance of various disease severity scoring systems in patients with acute respiratory failure or ARDS,^[23] and inconsistent results were found. Meanwhile, efforts are also put into the development of new prediction models,^[24] but to date, no validated scoring system has been available to predict mortality in this patient population. A prospective cohort study including 110 adult ICU patients with ARDS reported the APACHE III scoring systems was superior to that of APACHE II, SOFA, and Simplified APS II in terms of predicting the severity and mortality.^[25] Even for the same scoring system, predictive performance (i.e., evaluated by AUC) reported by these available studies could be inconsistent, which may be related to the different study populations where the scoring system was examined. However, generally speaking, it is consistent that a more complex scoring system usually shows better predictive performance than a simple one. According to results of our study, qSOFA had the lowest predictive power for hospital mortality for acute respiratory failure patients with an AUC of 0.553 only, which was worse than the other 3 scoring systems. Given the very simple parameters of qSOFA, such a result is not surprising. The qSOFA system contains only 3 parameters (respiratory rate, mental status, and systolic blood pressure), and Maitra et al^[26] reported that qSOFA was not sensitive enough for predicting hospital mortality in patients with suspected infection. In contrast, the OASIS system, which contains 10 parameters without any laboratory testing, showed a better predictive power than that of qSOFA. The APS III possesses the most complicated parameters, so it is not surprising to find that it also had the largest numerical AUC. It should be noted that the choice of a scoring system often takes the complexity of the system into account. Our study for the first time compared the predictive performance of the newly developed OASIS with that of other widely used scoring systems. Results of our study indicate a moderate predictive performance of OASIS, which contains only 10 non-laboratory parameters. Although such performance is worse than the APS III but similar to SOFA (both of which have more complex parameters), the simplicity of this new scoring system makes it promising to serve as a prognostic stratification tool. However, for ICU settings where the APACHE system is routinely evaluated, our results suggest that it is reasonable to use this scoring system, which is at least with fair predictive performance.

It should be noted that there were some limitations in the study. First, in order to keep the study population as homogenous as possible, we only included patients admitted to ICU with acute respiratory failure based on ICD codes and who needed oxygen therapy within 24 hours after ICU admission, and thus results of our study could not be directly applied to ICU patients who

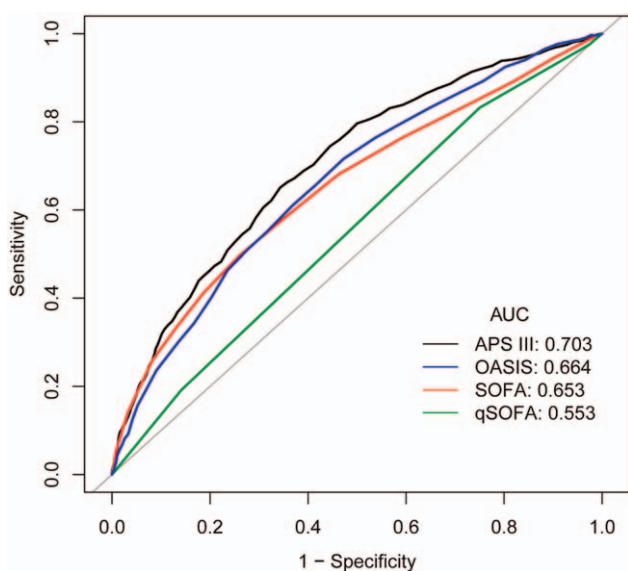


Figure 1. ROC curves of 4 scoring systems for prediction of hospital mortality. APS=acute physiology score, AUC = area under the curve, OASIS=Oxford Acute Severity of Illness Score, qSOFA=quick SOFA, ROC = receiver operating characteristic, SOFA=Sequential Organ Failure Assessment score.

Table 4
Predictive performance of different scoring systems for predicting hospital mortality.

Scoring system	AUC	95% CI	Best threshold	Specificity	Sensitivity	PPV	NPV
APS III	0.703	0.683–0.722	52.5	0.657	0.651	0.277	0.903
SOFA	0.653	0.631–0.675	7.5	0.740	0.496	0.278	0.879
qSOFA	0.553	0.535–0.572	1.5	0.250	0.832	0.183	0.881
OASIS	0.664	0.644–0.685	37.5	0.527	0.716	0.234	0.902

APS = acute physiology score, AUC = area under the curve, CI = confidence interval, NPV = negative predictive value, OASIS = Oxford Acute Severity of Illness Score, PPV = positive predictive value, qSOFA = quick SOFA, SOFA = Sequential Organ Failure Assessment score.

developed acute respiratory failure during ICU stay. Second, considering that the APACHE III system contains chronic disease parameters, the study extracted the acute physiological score part (i.e., APS) for analysis only, leaving the predicted value of the whole system unknown. Empirically speaking, we speculated that the predictive power of whole system should be better than APS III. Third, we cannot rule out residual confounding in the analysis of the associations between the scoring systems and hospital mortality (such as ethnic background and medication use). However, this should not be a serious concern since the main findings in our study were based on ROC curve analysis. Last but not least, most variables were identified by ICD codes, which may be biased due to risk of misclassification. Future studies should address these limitations to further confirm our results. In addition, it is promising in future researches to improve the predictive performance of OASIS by incorporating new predictor (s), since the only 10 non-laboratory parameters in the original system already showed a moderate predictive performance.

In conclusion, different disease severity scoring systems have different prognostic predictive powers for ICU patients admitted with acute respiratory failure, and APS III had the best predictive prognostic predictive powers. The OASIS scoring system might be a fair choice when taking both predictive power and simplicity into consideration.

Author contributions

WCH and HJX contributed to data analysis, interpretation of the results, and creation of draft of the manuscript. HTF and MHY contributed to interpretation of the results and revision of the manuscript. YCH contributed study design, interpretation of the results, and revision of the manuscript. All authors approved the final version of the manuscript.

Conceptualization: Wen-Cheng Huang, Hong-Jian Xie, Yuan-Cheng Hong.

Data curation: Wen-Cheng Huang, Hong-Jian Xie.

Formal analysis: Wen-Cheng Huang, Hong-Jian Xie, Hong-Tao Fan, Mei-Hao Yan.

Methodology: Hong-Tao Fan.

Supervision: Yuan-Cheng Hong.

Validation: Yuan-Cheng Hong.

Writing – original draft: Wen-Cheng Huang, Hong-Jian Xie.

Writing – review & editing: Hong-Tao Fan, Mei-Hao Yan, Yuan-Cheng Hong.

References

- [1] Breslow MJ, Badawi O. Severity scoring in the critically ill: part 1—interpretation and accuracy of outcome prediction scoring systems. *Chest* 2012;141:245–52.
- [2] Evran T, Serin S, Gürses E, Sungurtekin H. Various scoring systems for predicting mortality in Intensive Care Unit. *Niger J Clin Pract* 2016;19:530–4.
- [3] Johnson AE, Kramer AA, Clifford GD. A new severity of illness scale using a subset of Acute Physiology And Chronic Health Evaluation data elements shows comparable predictive accuracy. *Crit Care Med* 2013;41:1711–8.
- [4] Mendez-Tellez PA, Dorman T. Predicting patient outcomes, futility, and resource utilization in the intensive care unit: the role of severity scoring systems and general outcome prediction models. *Mayo Clin Proc* 2005;80:161–3.
- [5] Przybysz TM, Heffner AC. Early treatment of severe acute respiratory distress syndrome. *Emerg Med Clin North Am* 2016;34:1–14.
- [6] Stefan MS, Shieh MS, Pekow PS, et al. Epidemiology and outcomes of acute respiratory failure in the United States, 2001 to 2009: a national survey. *J Hosp Med* 2013;8:76–82.
- [7] Scala R, Heunks L. Highlights in acute respiratory failure. *Eur Respir Rev* 2018;27:
- [8] Singer M, Deutschman CS, Seymour CW, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* 2016;315:801–10.
- [9] Johnson AE, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. *Sci Data* 2016;3:160035.
- [10] Knaus WA, Wagner DP, Draper EA, et al. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest* 1991;100:1619–36.
- [11] Johnson AE, Stone DJ, Celi LA, Pollard TJ. The MIMIC Code Repository: enabling reproducibility in critical care research. *J Am Med Inform Assoc* 2018;25:32–9.
- [12] Power GS, Harrison DA. Why try to predict ICU outcomes? *Curr Opin Crit Care* 2014;20:544–9.
- [13] Gaudard P, Colson P. Disease-specific scoring or generic scoring in ICU? *J Thorac Dis* 2016;8:1414–6.
- [14] Ranson JH, Rifkind KM, Roses DF, Fink SD, Eng K, Spencer FC. Prognostic signs and the role of operative management in acute pancreatitis. *Surg Gynecol Obstet* 1974;139:69–81.
- [15] Vincent JL, Moreno R. Clinical review: scoring systems in the critically ill. *Crit Care* 2010;14:207.
- [16] Bouch DC, Thompson JP. Severity scoring systems in the critically ill. *Contin Educ Anaesth Crit Care Pain* 2008;8:181–5.
- [17] Sarkar R, Martin C, Mattie H, Gichoya JW, Stone DJ, Celi LA. Performance of intensive care unit severity scoring systems across different ethnicities in the USA: a retrospective observational study. *Lancet Digit Health* 2021;3:e241–9.
- [18] Knaus WA, Zimmerman JE, Wagner DP, Draper EA, Lawrence DE. APACHE-acute physiology and chronic health evaluation: a physiologically based classification system. *Crit Care Med* 1981;9:591–7.
- [19] Cook DA, Joyce CJ, Barnett RJ, et al. Prospective independent validation of APACHE III models in an Australian tertiary adult intensive care unit. *Anaesth Intensive Care* 2002;30:308–15.
- [20] Vincent JL, Moreno R, Takala J, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. *Intensive Care Med* 1996;22:707–10.
- [21] Jones AE, Trzeciak S, Kline JA. The Sequential Organ Failure Assessment score for predicting outcome in patients with severe sepsis and evidence of hypoperfusion at the time of emergency department presentation. *Crit Care Med* 2009;37:1649–54.
- [22] Angus DC, Seymour CW, Coopersmith CM, et al. A framework for the development and interpretation of different sepsis definitions and clinical criteria. *Crit Care Med* 2016;44:e113–21.
- [23] Santos RS, Silva PL, Rocco JR, Pelosi P, Rocco PR. A mortality score for acute respiratory distress syndrome: predicting the future without a crystal ball. *J Thorac Dis* 2016;8:1872–6.

- [24] Ding XF, Li JB, Liang HY, et al. Predictive model for acute respiratory distress syndrome events in ICU patients in China using machine learning algorithms: a secondary analysis of a cohort study. *J Transl Med* 2019; 17:326.
- [25] Saleh A, Ahmed M, Sultan I, Abdel-lateif A. Comparison of the mortality prediction of different ICU scoring systems (APACHE II and III, SAPS II, and SOFA) in a single-center ICU subpopulation with acute respiratory distress syndrome. *Egypt J Chest Dis Tuberc* 2015; 64:843–8.
- [26] Maitra S, Som A, Bhattacharjee S. Accuracy of quick Sequential Organ Failure Assessment (qSOFA) score and systemic inflammatory response syndrome (SIRS) criteria for predicting mortality in hospitalized patients with suspected infection: a meta-analysis of observational studies. *Clin Microbiol Infect* 2018;24:1123–9.