ORIGINAL ARTICLE

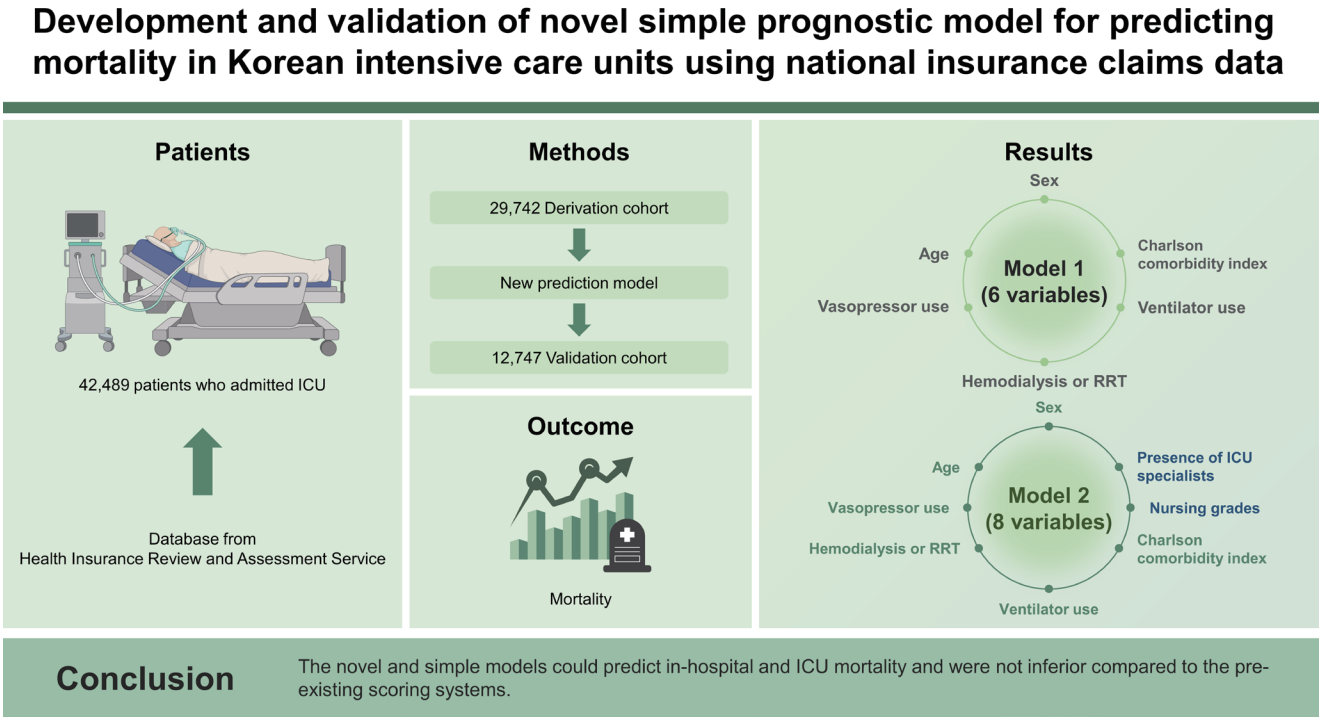
Korean J Intern Med 2024;39:625-639
https://doi.org/10.3904/kjim.2022.311

Development and validation of novel simple prognostic model for predicting mortality in Korean intensive care units using national insurance claims data

Ah Young Leem[1,*], Soyul Han[2,*], Kyung Soo Chung[1], Su Hwan Lee[1], Moo Suk Park[1], Bora Lee[3], and Young Sam Kim[1]

[1]Division of Pulmonology, Department of Internal Medicine, Institute of Chest Disease, Severance Hospital, Yonsei University College of Medicine, Seoul; [2]Department of Statistics, Graduate School of Chung-Ang University, Seoul; [3]Institute of Health & Environment, Seoul National University, Seoul, Korea

*These authors contributed equally to this manuscript.



Background/Aims: Intensive care unit (ICU) quality is largely determined by the mortality rate. Therefore, we aimed to develop and validate a novel prognostic model for predicting mortality in Korean ICUs, using national insurance claims data.

Methods: Data were obtained from the health insurance claims database maintained by the Health Insurance Review and Assessment Service of South Korea. From patients who underwent the third ICU adequacy evaluation, 42,489 cases were enrolled and randomly divided into the derivation and validation cohorts. Using the models derived from the derivation co-

pISSN 1226-3303
eISSN 2005-6648
http://www.kjim.org

hort, we analyzed whether they accurately predicted death in the validation cohort. The models were verified using data from one general and two tertiary hospitals.

Results: Two severity correction models were created from the derivation cohort data, by applying variables selected through statistical analysis, through clinical consensus, and from performing multiple logistic regression analysis. Model 1 included six categorical variables (age, sex, Charlson comorbidity index, ventilator use, hemodialysis or continuous renal replacement therapy, and vasopressor use). Model 2 additionally included presence/absence of ICU specialists and nursing grades. In external validation, the performance of models 1 and 2 for predicting in-hospital and ICU mortality was not inferior to that of pre-existing scoring systems.

Conclusions: The novel and simple models could predict in-hospital and ICU mortality and were not inferior compared to the pre-existing scoring systems.

Keywords: Intensive care unit; Health Insurance Review and Assessment Service; In-hospital mortality; Intensive care unit mortality

INTRODUCTION

The mortality rate is a main key indicator of intensive care unit (ICU) quality. However, patients' disease severity, comorbidities, and demographics significantly impact mortality [1]. Therefore, comparing mortality rates among ICUs without considering severity may provide an incorrect assessment of ICU quality.

A variety of outcome prediction scoring systems have been developed to provide an indication of the risk of death of groups of ICU patients. In 1981, the original Acute Physiology and Chronic Health Evaluation (APACHE) model was published [2], and there have been three subsequent revisions [3-5]. The Simplified Acute Physiology Score (SAPS) was devised as a simplification of APACHE II and has been revised twice since then [6,7]. Furthermore, because organ dysfunction is associated with high rates of ICU morbidity and mortality, organ failure scores such as the Sequential Organ Failure Assessment (SOFA) score have been developed [8]. The Mortality Probability Model (MPM) which is another validated ICU mortality prediction model, have been developed and updated [9,10].

The severity scores predict in-hospital and ICU mortality based on the severity of the patients' conditions. These scoring systems can be used in clinical trials for case-mix comparisons and for the assessment and comparison of ICU quality and performance. However, they were not designed for individual prognostication and the scoring system used by each hospital is different. Furthermore, not all hospitals use a prognostic scoring system. Therefore, there are limitations in predicting prognosis using the existing scoring systems.

Despite the availability of public databases such as Health Insurance Review and Assessment Service (HIRA) claims data, the development of a prognostic model for predicting mortality has not been reported in Asia, including South Korea. Therefore, we aimed to develop and validate a novel prognostic model for predicting mortality in Korean ICUs, using national insurance claims data.

METHODS

Study populations and data

This study used the database of 3rd quality evaluation of HIRA in critical care. Data were obtained from the health insurance claims database maintained by the HIRA of South Korea, the sole nationwide governmental agency that operates a fee-for-service reimbursement system. All Koreans are required to subscribe the National Health Insurance, a single medical insurance system. The insurance qualifications, treatment details, and medical institution information are stored in the HIRA database. Since data on ICU administration and discharge dates are not included in the HIRA claims data, it is not possible to create a severity correction model for ICU mortality. However, HIRA periodically evaluates the adequacy of ICUs, which have a high cost of medical care, to check the report data requested by medical institutions.

ICU mortality can be calculated as the ICU adequacy evaluation has data on ICU admission and discharge dates. The

https://doi.org/10.3904/kjim.2022.311

evaluation includes all ICUs in Korea. The HIRA third ICU adequacy evaluation data were used to develop a severity correction model for in-hospital and ICU mortality. The third ICU adequacy evaluation was conducted from May to July 2019 for institutions (including all general hospitals and tertiary hospitals) providing inpatient care in the ICU. Patients aged 18 years or older admitted to the ICU were included. Patients who were admitted to the ICU for less than 48 hours or were admitted to neonatal or pediatric ICUs, and burn patients, were excluded. Among the 56,926 patients who underwent the third ICU adequacy evaluation, the following cases were excluded: 11,507 not linked to claim data, 12 with claim data recorded after the date of death, and 2,918 with duplicate claims. The remaining 42,489 patients who were accurately identified and whose data were linked to health insurance claim data and date of death were randomly divided into the derivation and validation cohorts in a ratio of 7:3 (Fig. 1). In the derivation cohort, a model for calculating the predicted mortality was developed based on the factors influencing death using multiple logistic regression analysis. Using the model developed in the derivation cohort, we analyzed whether it accurately predicted death in the validation cohort. In addition, the model was verified using data from one general and two tertiary hospitals.

The Institutional Review Board of Yonsei University Health System determined that this study qualified for exempt status (IRB permit number: 4-2021-0212). The requirement for informed consent was waived because the study was a retrospective analysis of claims data from the HIRA. All patients in the HIRA dataset were anonymously recorded.

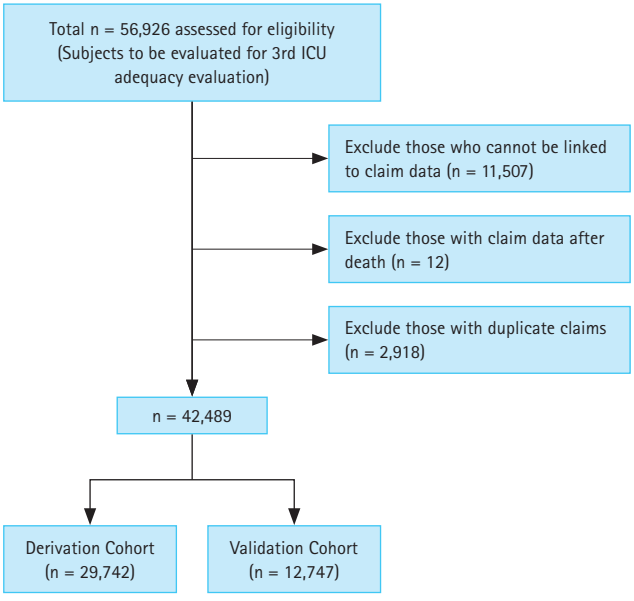


Figure 1. Study flow diagram. ICU, intensive care unit.

Selection of the variables for prognostic model

From prior research, variables that have been recognized as impacting mortality and are integrated into established prognostic scoring systems, were chosen from health insurance claim data [3,6-8,11]. The duration of hospitalization, age, sex, and chronic diseases were investigated to identify patient-related demographic factors. Age was divided into eight categories (18–29, 30–39, 40–49, 50–59, 60–69, 70–79, 80–89, and 90–99 yr). To evaluate chronic diseases, variables were generated by calculating the Charlson comorbidity index (CCI) using the claim code (International Classification of Diseases, 10th Revision) for medical use according to the analysis guide provided by the HIRA. The use of ventilators, hemodialysis or continuous renal replacement therapy (CRRT), and vasopressor drugs (norepinephrine, dopamine, or vasopressin) was identified using the claims code. These variables are factors associated with individual patients.

In addition, given the potential impact of a lack of ICU human resources, the presence of a dedicated ICU specialists [12], and bed-to-nurse grades [13] were also incorporated into a model as factors for predicting ICU and in-hospital mortality.

Statistical analysis

All categorical data are presented as frequency and proportion. Categorical data were analyzed using the chi-square test. The distribution of variables affecting in-hospital and ICU mortality was investigated in the derivation cohort. Bivariate analysis was performed to test the variable significance for in-hospital and ICU mortality. Based on the analysis results, significant variables were selected. In addition, multiple logistic regression analysis was performed to confirm that these factors affect mortality. The choice of variables was informed by the advisory board and clinical expert hearings.

Calibration was evaluated using the Hosmer–Lemeshow goodness-of-fit test (chi-square H). In this test, a large *p* value (> 0.05) indicates that the model is performing well, i.e., that there is not large discrepancy between observed and expected mortality. The area under the receiver operating characteristic (ROC) curve (AUC) was used to evaluate how

Table 1. Baseline characteristics of demographics between ICU survivors and non-survivors in derivation cohort

Variable	Total (n = 29,742)	Survivors (n = 26,263)	Non-survivors (n = 3,479)	*p* value
Sex				0.001
Male	17,029 (57.3)	14,929 (56.8)	2,100 (60.4)	
Female	12,713 (42.7)	11,334 (43.2)	1,379 (39.6)	
Age group (yr)				0.001
18–29	492 (1.7)	451 (1.7)	41 (1.1)	
30–39	796 (2.7)	753 (2.9)	43 (1.2)	
40–49	1,968 (6.6)	1,811 (7.0)	157 (4.5)	
50–59	4,261 (14.3)	3,828 (14.6)	433 (12.5)	
60–69	5,547 (18.7)	4,942 (18.8)	605 (17.4)	
70–79	7,756 (26.1)	6,797 (25.9)	959 (27.6)	
80–89	7,451 (25.1)	6,444 (24.5)	1,007 (29.0)	
90–99	1,471 (5.0)	1,237 (4.7)	234 (6.7)	
CCI				0.001
CCI 0	4,010 (13.5)	3,651 (13.9)	359 (10.3)	
CCI 1	6,693 (22.5)	6,082 (23.2)	611 (17.6)	
CCI 2	6,177 (20.8)	5,443 (20.7)	734 (21.1)	
CCI 3+	12,862 (43.3)	11,087 (42.2)	1,775 (51.0)	
Hospital type				0.884
Tertiary	11,469 (38.6)	10,123 (38.5)	1,346 (38.7)	
General	18,273 (61.4)	16,140 (61.5)	2,133 (61.3)	
Ventilator				0.001
No	19,508 (65.6)	18,489 (70.4)	1,019 (29.3)	
Yes	10,234 (34.4)	7,774 (29.6)	2,460 (70.7)	
CRRT, dialysis				0.001
No	26,223 (88.2)	23,801 (90.6)	2,422 (69.6)	
Yes	3,519 (11.8)	2,462 (9.4)	1,057 (30.4)	
Norepinephrine				0.001
No	18,247 (61.4)	17,443 (66.4)	804 (23.1)	
Yes	11,495 (38.7)	8,820 (33.6)	2,675 (76.9)	
Dopamine				0.001
No	27,002 (90.8)	24,290 (92.5)	2,712 (78.0)	
Yes	2,740 (9.2)	1,973 (7.5)	767 (22.1)	
Vasopressin				0.001
No	27,178 (91.4)	24,755 (94.3)	2,423 (69.7)	
Yes	2,564 (8.6)	1,508 (5.7)	1,056 (30.4)	
Emergency room hospitalization				0.719
No	8,862 (29.8)	7,835 (29.8)	1,027 (29.5)	
Yes	20,880 (70.2)	18,428 (70.2)	2,452 (70.5)	
Nursing grade				0.001
Grade 1	9,185 (30.9)	8,172 (31.1)	1,013 (29.1)	
Grade 2	9,993 (33.6)	8,824 (33.6)	1,169 (33.6)	

https://doi.org/10.3904/kjim.2022.311

Table 1. Continued

Variable	Total (n = 29,742)	Survivors (n = 26,263)	Non-survivors (n = 3,479)	*p* value
Grade 3	3,142 (10.6)	2,776 (10.6)	366 (10.5)	
Grade 4	1,167 (3.9)	1,054 (4.0)	113 (3.3)	
Grade 5	1,632 (5.5)	1,449 (5.5)	183 (5.3)	
Grade 6	1,917 (6.5)	1,649 (6.3)	268 (7.7)	
Grade 7	1,130 (3.8)	963 (3.7)	167 (4.8)	
Grade 8	912 (3.1)	792 (3.0)	120 (3.5)	
Grade 9	664 (2.2)	584 (2.2)	80 (2.3)	
ICU specialists				0.008
No	16,459 (55.3)	14,607 (55.6)	1,852 (53.2)	
Yes	13,283 (44.7)	11,656 (44.4)	1,627 (46.8)	
ECMO				0.001
No	29,414 (98.9)	26,073 (99.3)	3,341 (96.0)	
Yes	328 (1.1)	190 (0.7)	138 (4.0)	
High flow nasal xannula				0.001
No	24,613 (82.8)	22,111 (84.2)	2,502 (71.9)	
Yes	5,129 (17.2)	4,152 (15.8)	977 (28.1)	

Values are presented as number (%) for categorical variables.
ICU, intensive care unit; CCI, Charlson comorbidity index; CRRT, continuous renal replacement therapy; ECMO, extracorporeal membrane oxygenation.
p value was computed by chi-square test or Fisher's exact test for categorical variables as appropriate.

well the model discriminated between patients who lived and patients who died. AUC (which can range from 0 to 1) greater than 0.7 are generally considered evidence of good model identification. The Youden index was used to determine whether death could be determined according to a cut-off point, which is the optimal criterion. The criteria for mortality were compared using the Youden index to determine sensitivity, specificity, and accuracy using a 2 × 2 contingency table based on the cut-off point. Sensitivity predicts how well 1 predicts when the actual value is 1, specificity predicts how well it predicts 0 when the actual value is 0. Accuracy is a measure of how well it predicts 0 if the observed value is 0 and 1 if the observed value is 1.

Using the developed model, the regression coefficients for each variable can be estimated, and the predicted mortality value of a patient can be calculated. Furthermore, using the model, data on newly admitted patients to the ICU could be used to calculate the predicted mortality through the following steps.

1) Logit $g(x)$ could be calculated.:

$$g(x) = \beta_0 + \beta_i x_i + \cdots + \beta_k x_k$$

β_0 is the constant, $\beta_i x_i$ is the estimated coefficient for the i-th variable; i has a value from 1 to k, and k is a function of the model. Since all variables used in the model are categorical variables, dummy variables were created, and 1 was entered; if each variable was not applicable, 0 was entered. We calculated the logit by multiplying the input value by the corresponding coefficient.

2) The logit value was then converted into a probability (death rate) according to the following equation.

$$\text{Probability(mortality)} = [e^{g(x)}] / [1 + e^{g(x)}]$$

Calculation equations of $g(x)$ for predicting mortality are described in Supplementary Material.

Supplementary Table 1 describes the calculation method for predicting mortality using the model (model 1 in Supplementary Table 1). The information recorded in the Table is based on one patient randomly selected from the health insurance claim data. The male patient was aged 70–79 years, with a CCI score of ≥ 3. He did not require ventilation, CRRT, or dialysis during hospitalization. This patient's $g(x)$ was calculated as -2.233 and the intra-ICU mortality was calculated using model 1 as 0.097. Since Youden's index cutoff of ICU

Table 2. Goodness-of-fit of model 1 and model 2 for predicting in-hospital mortality

Variable	Model 1			Model 2		
	β(SE)	Adjusted OR	95% CI	β(SE)	Adjusted OR	95% CI
Constant	-4.258 (0.194)			-4.644 (0.201)		
Sex						
Male		1 (ref)			1 (ref)	
Female	-0.142 (0.037)	0.87	0.81–0.93**	-0.141 (0.037)	0.87	0.81–0.93**
Age group (yr)						
18–29		1 (ref)			1 (ref)	
30–39	0.012 (0.231)	1.01	0.64–1.59	-0.022 (0.233)	0.98	0.62–1.55
40–49	0.259 (0.204)	1.30	0.87–1.93	0.214 (0.206)	1.24	0.83–1.86
50–59	0.438 (0.193)	1.55	1.06–2.26*	0.391 (0.195)	1.48	1.01–2.17*
60–69	0.490 (0.191)	1.63	1.12–2.37*	0.436 (0.193)	1.55	1.06–2.26*
70–79	0.878 (0.189)	2.41	1.66–3.48**	0.799 (0.191)	2.22	1.53–3.23**
80–89	1.327 (0.189)	3.77	2.60–5.46**	1.163 (0.191)	3.20	2.20–4.66**
90–99	1.948 (0.200)	7.01	4.74–10.37**	1.689 (0.202)	5.41	3.64–8.04**
CCI						
CCI 0		1 (ref)			1 (ref)	
CCI 1	0.065 (0.070)	1.07	0.93–1.22	0.087 (0.070)	1.09	0.95–1.25
CCI 2	0.260 (0.068)	1.30	1.14–1.48**	0.287 (0.069)	1.33	1.16–1.52**
CCI 3+	0.464 (0.062)	1.59	1.41–1.80**	0.459 (0.062)	1.58	1.40-1.79**
Respirator						
No		1 (ref)			1 (ref)	
Yes	1.081 (0.039)	2.95	2.73–3.18**	1.282 (0.042)	3.60	3.32–3.91**
CRRT, dialysis						
No		1 (ref)			1.00	
Yes	0.982 (0.044)	2.67	2.45–2.91**	1.088 (0.045)	2.97	2.72–3.25**
Vasopressor						
No		1 (ref)			1 (ref)	
Yes	1.335 (0.042)	3.80	3.50–4.13**	1.453 (0.044)	4.28	3.93–4.66**
Nurse grade						
Grade 1					1 (ref)	
Grade 2				0.154 (0.045)	1.17	1.07–1.27*
Grade 3				0.422 (0.071)	1.52	1.33–1.75**
Grade 4				0.955 (0.102)	2.60	2.13–3.17**
Grade 5				0.506 (0.094)	1.66	1.38–1.99**
Grade 6				1.111 (0.083)	3.04	2.58–3.57**
Grade 7				1.111 (0.098)	3.04	2.51–3.68**
Grade 8				0.970 (0.113)	2.64	2.11–3.29**
Grade 9				0.944 (0.129)	2.57	2.00-3.31**
ICU specialists[a]						
No					1 (ref)	
Yes				-0.098 (0.045)	0.91	0.83–0.99*

https://doi.org/10.3904/kjim.2022.311

Table 2. Continued

Variable	Model 1			Model 2		
	β(SE)	Adjusted OR	95% CI	β(SE)	Adjusted OR	95% CI
Cutoff	0.151			0.144		
Sensitivity	0.761			0.780		
Specificity	0.713			0.713		
Accuracy	0.720			0.723		
AUC (CI)	0.802 (0.802–0.812)			0.811 (0.811–0.821)		
H-L (*p* value)	20.572 (0.008)			11.353 (0.183)		

OR, odds ratio; CI, confidence interval; CCI, Charlson comorbidity index; CRRT, continuous renal replacement therapy; ICU, intensive care unit; Cutoff, Youden index; AUC, area under receiver operating characteristic curve; H-L, Hosmer–Lemeshow.
[a)]Presence of absence of ICU specialists.
The *p* values were obtained by multivariate logistic analysis. **p* < 0.05, ***p* < 0.01.

model 1 was 0.117, the patient was predicted to be survivor; the actual claim data confirmed the prediction.

To compare the discriminatory power of the developed model with APACHE, SAPS, and SOFA, the tools currently in use, the AUCs were compared using the DeLong's test.

All statistical analyses were conducted using the SAS Enterprise Guide 7.1 and R version 3.5.1. Two-sided *p* values < 0.05 were considered statistically significant.

RESULTS

Construction of data for the development of a severity adjustment model

Data from 42,489 cases were randomly divided into the derivation and validation cohorts (7:3). Of the 29,742 patients in the derivation cohort, 3,479 patients had died in the ICU. The baseline demographics of ICU survivors and non-survivors in the derivation cohort are shown in Table 1. The proportion of male, age more than 70 years, and CCI 3 or above was significantly higher in non-survivors than were in survivors. Ventilator use, CRRT or dialysis, and use of vasopressor drugs were significantly more common in the non-survivor group. In the non-survivor group, the proportion of high bed-to-nurse grade was significantly higher and the proportion of the presence of ICU specialists was lower than in the survivor group. Extracorporeal membrane oxygenation and high-flow nasal cannula use were more common in the non-survivor group; however, these parameters were not included in the model because their application frequency was relatively low, and the data was not available for all ICUs.

Derivation and validation of the severity correction model for prediction of in-hospital and ICU mortality

As described in the methods, age, sex, CCI, ventilator use, hemodialysis or CRRT, vasopressor use were selected to create model 1. Furthermore, the presence or absence of ICU specialists and bed-to-nurse grades were added as variables to create model 2 for predicting in-hospital and ICU mortality more accurately using the billing code.

Multivariate logistic regression analysis was performed on the derivation cohort to determine if these variables were related to the mortality rate. The variables selected for the multivariate analysis were audited by a committee consisting of intensivists comprising nine physicians specializing in critical care medicine.

Finally, two severity correction models were developed in the derivation cohort (models 1 and 2) by applying variables selected after multiple logistic regression analysis and clinical consideration.

Model 1 included six patient-related categorical variables (age, sex, CCI, ventilator use, hemodialysis or CRRT, and vasopressor use). Model 2 included the presence or absence of ICU specialists and nursing grades as correction variables, aiming to improve accuracy when predicting ICU and in-hospital mortality.

The criteria for evaluating whether a model composed of selected variables explains the observed outcome well are

Table 3. Goodness-of-fit of model 1 and model 2 for predicting ICU mortality

Variable	Model 1			Model 2		
	β(SE)	Adjusted OR	95% CI	β(SE)	Adjusted OR	95% CI
Constant	-4.336 (0.196)			-4.800 (0.205)		
Sex						
Male		1 (ref)			1 (ref)	
Female	-0.085 (0.041)	0.92	0.85–1.00*	-0.080 (0.041)	0.92	0.85–1.00
Age group (yr)						
18–29		1 (ref)			1(ref)	
30–39	-0.224 (0.239)	0.80	0.50–1.28	-0.265 (0.243)	0.77	0.48–1.24
40–49	0.101 (0.207)	1.11	0.74–1.66	0.043 (0.210)	1.04	0.69–1.57
50–59	0.251 (0.194)	1.29	0.88–1.88	0.194 (0.197)	1.21	0.83–1.79
60–69	0.275 (0.192)	1.32	0.90–1.92	0.208 (0.195)	1.23	0.84–1.80
70–79	0.514 (0.190)	1.67	1.15–2.43*	0.412 (0.193)	1.51	1.03–2.20*
80–89	0.948 (0.190)	2.58	1.78–3.74**	0.746 (0.193)	2.11	1.45–3.08**
90–99	1.449 (0.204)	4.26	2.85–6.35**	1.140 (0.207)	3.13	2.08–4.69**
CCI						
CCI 0		1 (ref)			1 (ref)	
CCI 1	0.029 (0.076)	1.03	0.89–1.19	0.051 (0.077)	1.05	0.90–1.22
CCI 2	0.190 (0.074)	1.21	1.05–1.40*	0.215 (0.075)	1.24	1.07–1.44*
CCI 3+	0.221 (0.068)	1.25	1.09–1.43*	0.201 (0.069)	1.22	1.07–1.44*
Respirator						
No		1 (ref)			1 (ref)	
Yes	1.220 (0.045)	3.39	3.10–3.70**	1.471 (0.048)	4.35	3.96–4.78**
CRRT, dialysis						
No		1 (ref)			1 (ref)	
Yes	1.000 (0.047)	2.72	2.48–2.98**	1.132 (0.049)	3.10	2.82–3.41**
Vasopressor						
No		1 (ref)			1 (ref)	
Yes	1.368 (0.050)	3.93	3.56–4.33**	1.510 (0.051)	4.53	4.09–5.01**
Nurse grade						
Grade 1					1 (ref)	
Grade 2				0.263 (0.050)	1.30	1.18–1.43**
Grade 3				0.554 (0.079)	1.74	1.49–2.03**
Grade 4				0.988 (0.119)	2.69	2.13–3.39**
Grade 5				0.634 (0.105)	1.88	1.53–2.31**
Grade 6				1.265 (0.093)	3.54	2.95–4.25**
Grade 7				1.299 (0.109)	3.67	2.96–4.54**
Grade 8				1.228 (0.125)	3.41	2.67–4.36**
Grade 9				1.080 (0.146)	2.94	2.21–3.92**
ICU doctor[a)]						
No					1 (ref)	
Yes				-0.186 (0.050)	0.83	0.75–0.92**

https://doi.org/10.3904/kjim.2022.311

Table 3. Continued

Variable	Model 1			Model 2		
	β(SE)	Adjusted OR	95% CI	β(SE)	Adjusted OR	95% CI
Cutoff	0.117			0.096		
Sensitivity	0.774			0.818		
Specificity	0.725			0.704		
Accuracy	0.731			0.717		
AUC (CI)	0.812 (0.812–0.823)			0.825 (0.825–0.836)		
H-L (*p* value)	34.423 (< 0.01)			18.234 (0.020)		

ICU, intensive care unit; OR, odds ratio; CI, confidence interval; CCI, Charlson comorbidity index; CRRT, continuous renal replacement therapy; Cutoff, Youden index; AUC, area under receiver operating characteristic curve; H-L, Hosmer–Lemeshow.
[a)]Presence of absence of ICU specialists.
The *p* values were obtained by multivariate logistic analysis. **p* < 0.05, ***p* < 0.01.

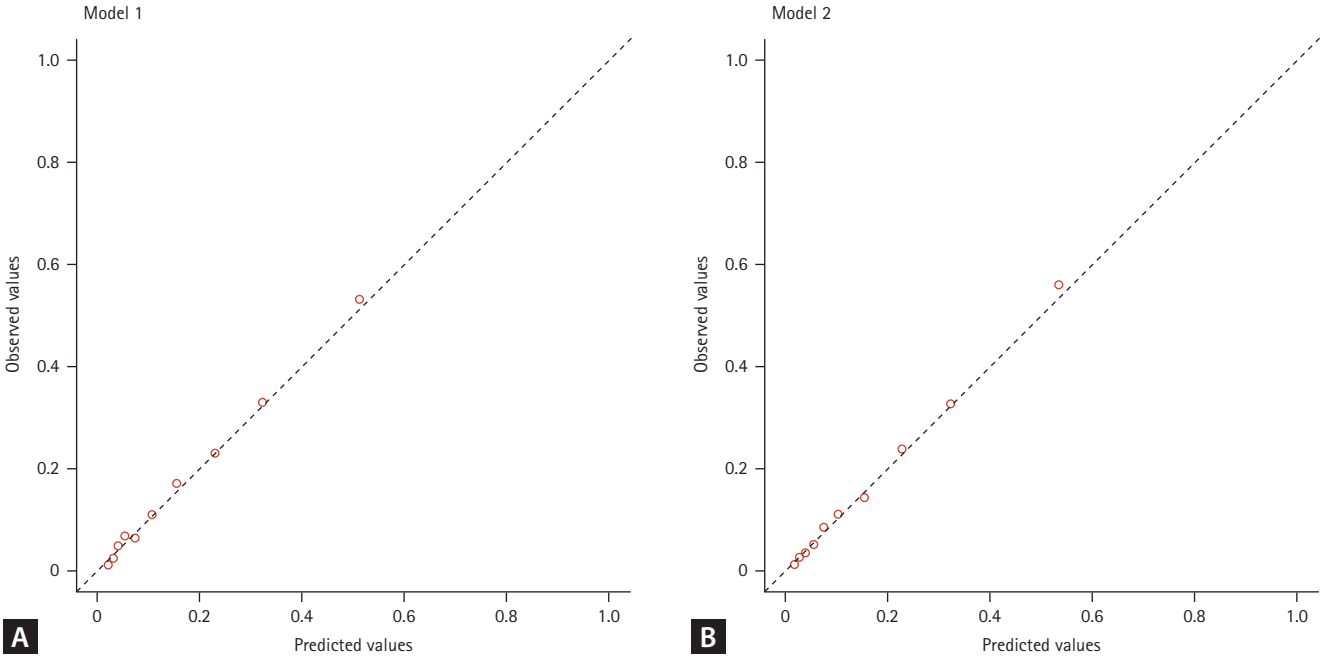


Figure 2. Calibration plots for (A) model 1 and (B) model 2 for predicting in-hospital mortality. The 12,747 patients in the validation cohort were divided into 10 groups according to the value of predicted mortality, to compare the predicted and actual probabilities. The diagonal dotted line represents a perfect agreement between observed and expected mortality estimates.

discrimination, calibration, and overall model performance.

Tables 2 and 3 show the goodness-of-fit evaluations of the models based on 12,747 patients in the validation cohort, demonstrating the degree of agreement between the observed and predicted mortality based on the model probability. For model 1, the goodness-of-fit for in-hospital and ICU mortality was 20.572 and 34.423, respectively, and the Hosmer–Lemeshow statistic *p* value was 0.008 and < 0.01, respectively. The goodness-of-fit for in-hospital and ICU mortality for model 2 was 11.353 and 18.234, respectively, and the Hosmer–Lemeshow statistic *p* value was 0.183 and 0.020, respectively (Table 2, 3). The performance of the models was evaluated using AUCs. For models 1 and 2, the AUCs for the in-hospital mortality rate were 0.802 and 0.811, respectively, confirming that both models had excellent discrimination power of ≥ 0.8. In addition, it was

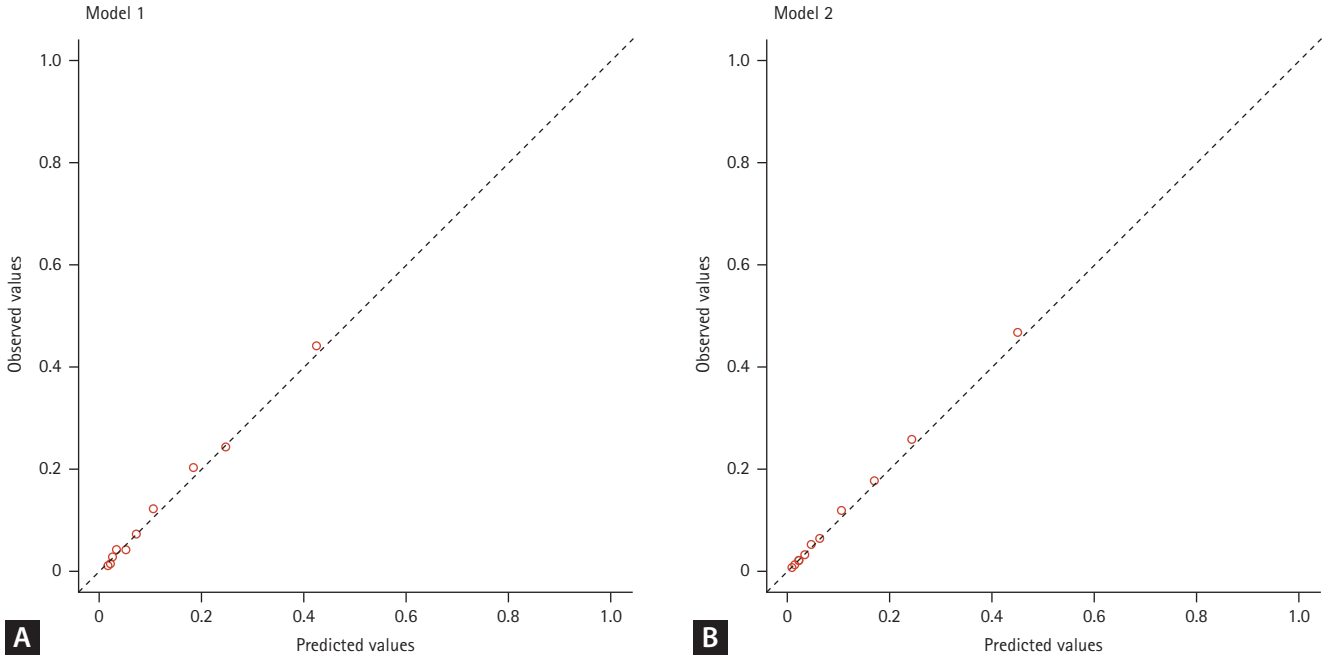


Figure 3. Calibration plots for (A) model 1 and (B) model 2 for predicting ICU mortality. The 12,747 patients in the validation cohort were divided into 10 groups according to the value of predicted mortality, to compare the predicted and actual probabilities. The diagonal dotted line represents a perfect agreement between observed and expected mortality estimates. ICU, intensive care unit.

confirmed that the AUCs of models 1 and 2 for ICU mortality were 0.812 and 0.825, respectively, again demonstrating excellent discrimination. Both severity correction models performed better for ICU mortality than for in-hospital mortality. The cutoff value is the result of calculating the optimal decision reference point using the Youden index, and sensitivity, specificity, and accuracy were calculated using the cutoff value. The cutoff values for models 1 and 2, for both in-hospital and ICU mortality, were ≥ 0.7, which indicates the models are suitable.

Using the developed severity correction models, the regression coefficients for each variable can be estimated, and the predicted mortality value for an individual patient can be calculated. Using model 1 coefficients in Table 4, data on newly admitted patients to the ICU could be used to calculate the predicted mortality.

Calibration plots of the validation cohort for model 1 and model 2, for predicting in-hospital mortality and ICU mortality, are presented in Figures 2 and 3. The 12,747 patients in the validation cohort were divided into 10 equal sized groups by using the deciles of the predicted mortality. The calibration curve represents the relationship between the predicted mortality and the observed mortality. The diagonal dotted lines represents a good agreement between observed and expected mortality estimates.

Comparison of the developed model and existing mortality prediction models

The external validity of models 1 and 2 was examined, by evaluating their performance when applied to other patient populations. In addition, we confirmed whether the performance of the models using data collected from three hospitals was comparable to the two measurement tools currently in use (APACHE II and SAPS III) (Table 4, Fig. 4). AUCs were compared using DeLong's test. Using data from 1,000 and 404 ICU patients from tertiary hospitals 1 and 2, the degree of prediction of in-hospital and ICU mortality was determined for models 1 and 2. In the comparison of predictive power for in-hospital and ICU mortality, the performance of models 1 and 2 was not inferior to the existing models (APACHE II and SAPS III). When comparing the predictive power for ICU mortality in tertiary hospital 1, the AUCs of models 1 and 2 were significantly higher than that of APACHE II (model 1, $p = 0.007$; model 2, $p < 0.001$; Table 4, Fig. 4). When models 1 and 2 were used for data from 897 patients in a general hospital, there was no significant

https://doi.org/10.3904/kjim.2022.311

Table 4. Comparison of AUC among scores (APACHE, SAPS, SOFA) and developed modeling

Model	Tertiary hospital 1				Tertiary hospital 2		General hospital 1			
	In-hospital mortality		ICU mortality		ICU mortality		In-hospital mortality		ICU mortality	
	AUC (95% CI)	*p* value	AUC (95% CI)	*p* value	AUC (95% CI)	*p* value	AUC (95% CI)	*p* value	AUC (95% CI)	*p* value
Model 1	0.685 (0.650–0.719)	Reference	0.726 (0.690–0.762)	Reference	0.752 (0.703–0.802)	Reference	0.841 (0.801–0.881)	Reference	0.857 (0.814–0.900)	Reference
APACHE II	0.671 (0.637–0.706)	0.460	0.677 (0.638–0.715)	0.007	0.749 (0.697–0.802)	0.934	0.823 (0.785–0.861)	0.420	0.829 (0.785–0.873)	0.288
SAPS III	0.690 (0.656–0.723)	0.774	0.705 (0.668–0.742)	0.150	0.792 (0.743–0.841)	0.222				
Model 2	0.690 (0.656–0.724)	Reference	0.738 (0.702–0.774)	Reference	0.760 (0.712–0.809)	Reference	0.843 (0.803–0.883)	Reference	0.857 (0.814–0.900)	Reference
APACHE II	0.671 (0.637–0.706)	0.301	0.677 (0.638–0.715)	< 0.01	0.749 (0.697–0.802)	0.747	0.823 (0.785–0.861)	0.356	0.829 (0.785–0.873)	0.286
SAPS III	0.690 (0.656–0.723)	0.984	0.705 (0.668–0.742)	0.102	0.792 (0.743–0.841)	0.326				

AUC, area under ROC curve; APACHE, Acute Physiology and Chronic Health Evaluation; SAPS, Simplified Acute Physiology Score; SOFA, Sequential Organ Failure Assessment; ICU, intensive care unit; CI, confidence interval; ROC, receiver operating characteristic.
DeLong's test for two correlated ROC curves.

https://doi.org/10.3904/kjim.2022.311

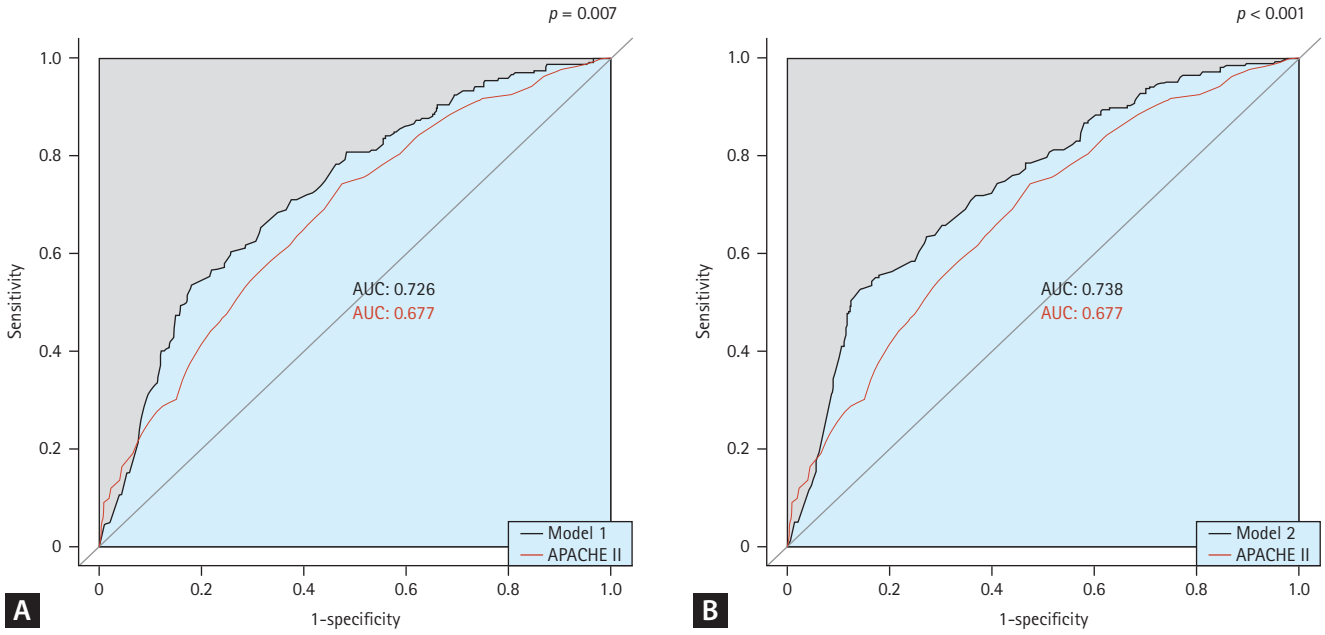


Figure 4. ROC curves of (A) models 1 and (B) 2 and APACHE II for prediction of ICU mortality in tertiary hospital 1. AUC, area under ROC curve; APACHE, Acute Physiology and Chronic Health Evaluation; ICU, intensive care unit; ROC, receiver operating characteristic.

difference in predicting in-hospital and ICU mortality compared to the APACHE II score based on health insurance claims, and the predicted AUC was > 0.8 (Table 4, Fig. 4). Calibration plots of external validation data for model 1 and model 2, for predicting in-hospital mortality and ICU mortality, are presented in Supplementary Figures 1–5. The diagonal dotted lines represents a good agreement between observed and expected mortality estimates.

DISCUSSION

In this study, we developed models that could predict in-hospital and ICU mortality based on the HIRA claims data. Model 1 included six patient-related categorical variables (age, sex, CCI, ventilator use, hemodialysis or CRRT, and vasopressor use). In addition, the presence or absence of ICU specialists and nursing grades were added as correction variables in model 2. Both models 1 and 2 showed results in predicting in-hospital and ICU mortality comparable to the existing scoring systems in the fitness verification of the validation cohort, and in external validity assessment using data from two tertiary hospitals, and one general hospital. This study is significant in that it developed a more convenient model for predicting ICU mortality than existing mortality prediction models through the analysis of data from the entire ICU patients in Korea.

Researchers have previously studied and developed prognostic scoring systems to predict the mortality rate of critically ill patients. APACHE is the most commonly used scoring system. In 1985, Knaus et al. [3] revised 34 items of the original APACHE to 12 items and published the results of validation of 5,815 ICU admission patients in 13 hospitals in the USA. The analysis showed that an increase in score was closely related to an increase in hospital death [3]. According to Zimmerman et al. [11], among 110,558 critically ill patients admitted to 104 ICUs in 45 hospitals in the USA, the new scoring system (APACHE IV) predicted in-hospital mortality with good calibration and discrimination (AUC: 0.880) . In contrast, Le Gall et al. [6] proposed the SAPS II model in a study of 13,152 patients admitted to 137 medical and/or surgical ICUs in 12 countries between September 1991 and February 1992. The SAPS II model predicted in-hospital mortality with AUCs of 0.88 and 0.86 in the developmental and validation samples, respectively [6]. Moreno et al. [7] presented SAPS III, a new model for predicting in-hospital mortality, through a multicenter, multinational cohort study of patients admitted to 303 ICUs from October to December 2002 (AUC: 0.848).

Studies predicting mortality using the SOFA score have

https://doi.org/10.3904/kjim.2022.311

also been reported [14]. Furthermore, Moreno et al. [14] suggested the usefulness of the total maximum SOFA score in predicting ICU mortality based on a multicenter-multinational cohort study among 1,449 patients admitted to 75 ICUs in 16 countries in March 1995 (AUC 0.772).

Several countries have also reported studies on mortality prediction models based on public databases. In the USA, the Mortality Probability Admission Model (MPM0)-III and APACHE-IV prognostic scoring systems are commonly used in ICUs. Efforts have been made at the California Healthcare Foundation and the National Quality Forum (NQF) to develop prosaic scoring systems for quality measurement; therefore, the ICU Outcomes Model (modified and recalibrated version of the MPM0-III model) was developed [15]. Data from 55,304 patients aged ≥ 18 admitted to 55 ICUs at USA hospitals from January 2008 to December 2012 showed that APACHE IV was the most accurate when compared with the ICU Outcomes Model/NQF, and the MPM0-III for predicting in-hospital mortality [15]. In the UK, Harrison et al. [16] developed and announced their own Intensive Care National Audit & Research Centre (ICNARC) model, based on the database of the Case Mix Program. The ICNARC model was developed based on the data of 216,626 patients admitted to 163 critical care units in England, Wales, and Northern Ireland from 1995 to 2003, and showed better discrimination and overall fit than pre-existing risk-prediction models for in-hospital mortality [16].

Several studies conducted in the ICUs in Korea to predict mortality have been reported. In a study which retrospectively reviewed 1,314 patients admitted to the surgical ICU from March 2011 to February 2012 in a university hospital, the overall discrimination and calibration of APACHE IV were similar to those of APACHE II, SAPS 3, and Korean SAPS 3 [17]. Another Korean study investigated the predictive power of APACHE II for in-hospital mortality in ICU patients [18]. In this study, data from the Fever and Antipyretics in Critical Illness Evaluation cohort were collected prospectively between September 1 to November 30, 2019, from adult patients aged ≥ 18 years who were admitted to ICUs in 25 hospitals (15 in Japan and 10 in Korea). The analyses showed that APACHE II predicted in-hospital mortality with poor calibration and modest discrimination. In the study of Lim et al. [19], mortality rates predicted using the general SAPS 3 and its customized equation (Australasia SAPS 3) exhibited good calibration and modest discrimination. However, the Australasia SAPS 3 did not improve the mortality prediction. A prospective multicenter cohort study involving 22 ICUs from 15 centers throughout Korea investigated the validation of the SAPS 3, and customized it, for Korean ICUs [20]. The new equation for Korean ICU patients was tested in a validation cohort, and demonstrated both good discrimination and good calibration. In particular, this study has clinical significance in that in presented a new equation that can be applied to Korean ICUs.

However, all of these previous studies have difficulties in measuring many variables to predict mortality. Furthermore, comparison between hospitals is difficult because not all ICUs use the same scoring system. Therefore, we aimed to develop and validate a novel prognostic model for predicting mortality, using national insurance claims data.

This study is meaningful in that it is the first study in Asia to develop and apply a model to predict in-hospital and ICU mortality using a large-scale national-level database that included data from all ICUs in Korea. This study's data were based on the HIRA of South Korea, the sole nationwide governmental agency that operates a fee-for-service reimbursement system. Therefore, the database is very systematic and contains a lot of information. Moreover, this study distinguishes itself by offering a model for predicting in-hospital and ICU mortality rates more conveniently than the existing scoring systems.

This study has several limitations. First, while the study was based on a large database released by the South Korean government, our findings may not be generalizable to other countries, which may have different patterns of ICU use. Second, because of data limitations, in-hospital mortality could not be analyzed for one of the tertiary hospitals; this limited our testing of external validity. Third, the period of study was spring and summer, and as hospital admissions are affected by the season, our findings may not be applicable to other times of the year. Finally, because of the limited information available, it was difficult to compare and validate the newly developed scoring system with SAPS III and SOFA scores. Therefore, a large-scale systematic study is needed.

In this study, we developed a model that can predict in-hospital and ICU mortality based on the HIRA claims data released by the South Korean government. The novel and simple models were not inferior in predicting in-hospital and ICU mortality compared to the pre-existing scoring systems.

KEY MESSAGE

1. In this study, we developed models that could predict in-hospital and ICU mortality based on the HIRA claims data released by the South Korean government.
2. Model 1 included six patient-related categorical variables (age, sex, CCI, ventilator use, hemodialysis or CRRT, and vasopressor use). In addition, the presence or absence of ICU specialists and nursing grades were added as correction variables in model 2.
3. Both models 1 and 2 showed consistent results in predicting in-hospital and ICU mortality, and were comparable to the existing scoring systems when tested in a validation cohort, and with external data from one general, and two tertiary hospitals.

REFERENCES

1. Garland A. Improving the ICU: part 1. Chest 2005;127:2151-2164.
2. 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;9:591-597.
3. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. Crit Care Med 1985;13:818-829.
4. 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;100:1619-1636.
5. Zimmerman JE, Kramer AA, McNair DS, Malila FM, Shaffer VL. Intensive care unit length of stay: Benchmarking based on Acute Physiology and Chronic Health Evaluation (APACHE) IV. Crit Care Med 2006;34:2517-2529.
6. Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. JAMA 1993;270:2957-2963.
7. Moreno RP, Metnitz PG, Almeida E, et al. SAPS 3--From evaluation of the patient to evaluation of the intensive care unit. Part 2: development of a prognostic model for hospital mortality at ICU admission. Intensive Care Med 2005;31:1345-1355.
8. 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;22:707-710.
9. Lemeshow S, Teres D, Avrunin JS, Gage RW. Refining intensive care unit outcome prediction by using changing probabilities of mortality. Crit Care Med 1988;16:470-477.
10. Higgins TL, Teres D, Copes WS, Nathanson BH, Stark M, Kramer AA. Assessing contemporary intensive care unit outcome: an updated Mortality Probability Admission Model (MPM0-III). Crit Care Med 2007;35:827-835.
11. Zimmerman JE, Kramer AA, McNair DS, Malila FM. Acute Physiology and Chronic Health Evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. Crit Care Med 2006;34:1297-1310.
12. Cho J, Lee HJ, Hong SB, et al. Structure of intensive care unit and clinical outcomes in critically ill patients with influenza A/H1N1 2009. Korean J Crit Care Med 2012;27:65-69.
13. Jung M, Park H, Kang D, et al. The effect of bed-to-nurse ratio on hospital mortality of critically ill children on mechanical ventilation: a nationwide population-based study. Ann Intensive Care 2020;10:159.
14. Moreno R, Vincent JL, Matos R, et al. The use of maximum SOFA score to quantify organ dysfunction/failure in intensive care. Results of a prospective, multicentre study. Working Group on Sepsis related Problems of the ESICM. Intensive Care Med 1999;25:686-696.
15. Kramer AA, Higgins TL, Zimmerman JE. Comparison of the Mortality Probability Admission Model III, National Quality Forum, and Acute Physiology and Chronic Health Evaluation IV hospital mortality models: implications for national benchmarking*. Crit Care Med 2014;42:544-553.
16. Harrison DA, Parry GJ, Carpenter JR, Short A, Rowan K. A new risk prediction model for critical care: the Intensive Care National Audit & Research Centre (ICNARC) model. Crit Care Med 2007;35:1091-1098.
17. Lee H, Shon YJ, Kim H, Paik H, Park HP. Validation of the APACHE IV model and its comparison with the APACHE II, SAPS 3, and Korean SAPS 3 models for the prediction of hospital mortality in a Korean surgical intensive care unit. Korean J Anesthesiol 2014;67:115-122.
18. Kim JY, Lim SY, Jeon K, et al. External validation of the Acute Physiology and Chronic Health Evaluation II in Korean intensive care units. Yonsei Med J 2013;54:425-431.
19. Lim SY, Ham CR, Park SY, et al. Validation of the Simplified

https://doi.org/10.3904/kjim.2022.311

Acute Physiology Score 3 scoring system in a Korean intensive care unit. Yonsei Med J 2011;52:59-64.
20. Lim SY, Koh SO, Jeon K, et al. Validation of SAPS3 admission score and its customization for use in Korean intensive care unit patients: a prospective multicentre study. Respirology 2013;18:989-995.

Received : October 7, 2022
Revised : March 2, 2023
Accepted : November 27, 2023

Correspondence to
Young Sam Kim, M.D. Ph.D.
Division of Pulmonology, Department of Internal Medicine, Institute of Chest Disease, Severance Hospital, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea
Tel: +82-2-2228-1931, Fax: +82-2-393-6884
E-mail: YSAMKIM@yuhs.ac
https://orcid.org/0000-0001-9656-8482

Acknowledgments
This work was supported by the Health Insurance Review & Assessment Service (HIRA). The views expressed are those of the author(s) and not necessarily those of the HIRA.

CRedit authorship contributions
Ah Young Leem: conceptualization, visualization, writing - original draft, writing - review & editing; Soyul Han: data curation, formal analysis; Kyung Soo Chung: conceptualization, methodology, writing - review & editing; Su Hwan Lee: conceptualization, data curation, writing - review & editing; Moo Suk Park: conceptualization, methodology, writing - review & editing; Bora Lee: conceptualization, data curation, formal analysis, methodology; Young Sam Kim: conceptualization, methodology, project administration, visualization

Conflicts of interest
The authors disclose no conflicts.

Funding
None

Supplementary Material. Calculation equations of g(x) for predicting mortality.

g(x) in Model 1 for in-hopital mortality

$$g(x) = -4.258 + (-0.142) \times Sex_2 + 0.012 \times Age\ group_2 + 0.259 \times Age\ group_3 + 0.438 \times Age\ group_4 + 0.490 \times Age\ group_5 + 0.878 \times Age\ group_6 + 1.327 \times Age\ group_7 + 1.948 \times Age\ group_8 + 0.065 \times CCIgroup_2 + 0.26 \times CCIgroup_3 + 0.464 \times CCIgroup_4 + 1.081 \times Respirator_2 + 0.982 \times CRRT,Dialysis_2 + 1.335 \times Drugs_2$$

g(x) in Model 1 for ICU mortality

$$g(x) = -4.336 + (-0.085) \times Sex_2 + (-0.224) \times Age\ group_2 + 0.101 \times Age\ group_3 + 0.251 \times Age\ group_4 + 0.275 \times Age\ group_5 + 0.514 \times Age\ group_6 + 0.948 \times Age\ group_7 + 1.449 \times Age\ group_8 + 0.029 \times CCIgroup_2 + 0.19 \times CCIgroup_3 + 0.221 \times CCIgroup_4 + 1.22 \times Respirator_2 + 1 \times CRRT,Dialysis_2 + 1.368 \times Drugs_2$$

g(x) in Model 2 for in-hopital mortality

$$g(x) = -4.644 + (-0.141) \times Sex_2 + (-0.022) \times Age\ group_2 + 0.214 \times Age\ group_3 + 0.391 \times Age\ group_4 + 0.436 \times Age\ group_5 + 0.799 \times Age\ group_6 + 1.163 \times Age\ group_7 + 1.689 \times Age\ group_8 + 0.087 \times CCIgroup_2 + 0.287 \times CCIgroup_3 + 0.459 \times CCIgroup_4 + 1.282 \times Respirator_2 + 1.088 \times CRRT,Dialysis_2 + 1.453 \times Drugs_2 + 0.154 \times Nrse\ grade_2 + 0.422 \times Nrse\ grade_3 + 0.955 \times Nrse\ grade_4 + 0.506 \times Nrse\ grade_5 + 1.111 \times Nrse\ grade_6 + 1.111 \times Nrse\ grade_7 + 0.97 \times Nrse\ grade_8 + 0.944 \times Nrse\ grade_9 + (-0.098) \times Special\ doctor_2$$

g(x) in Model 2 for ICU mortality

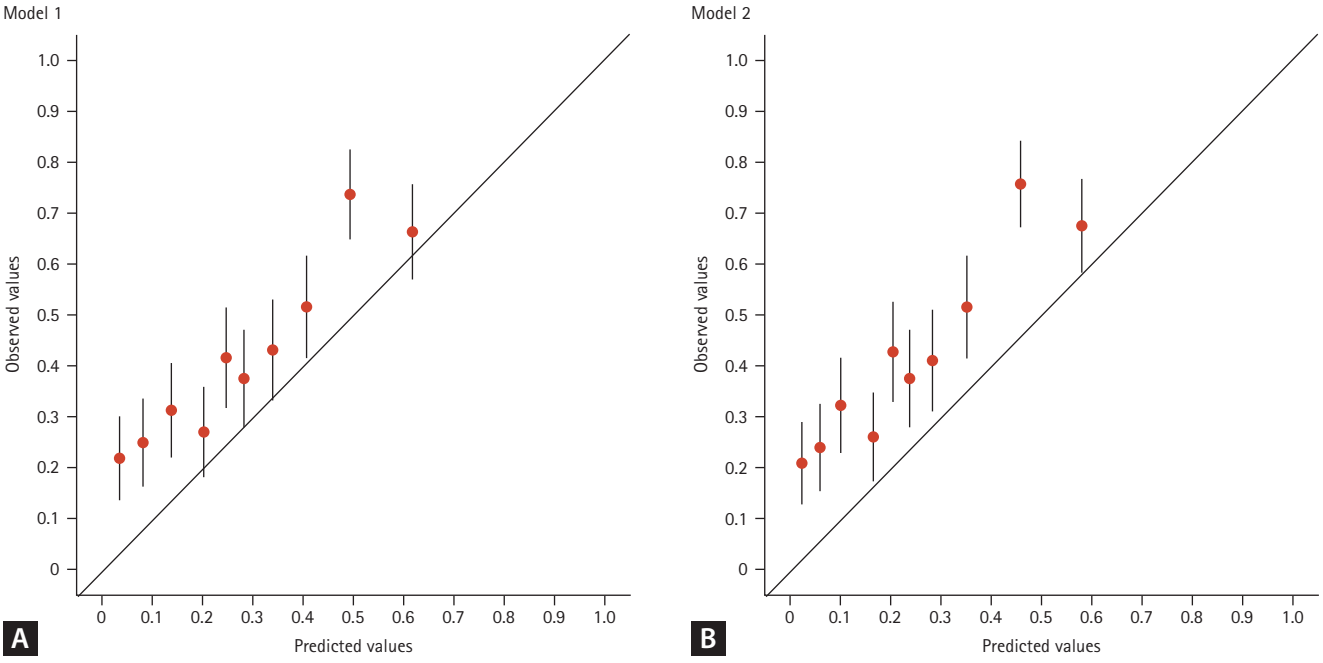
$$g(x) = -4.8 + (-0.08) \times Sex_2 + (-0.265) \times Age\ group_2 + 0.043 \times Age\ group_3 + 0.194 \times Age\ group_4 + 0.208 \times Age\ group_5 + 0.412 \times Age\ group_6 + 0.746 \times Age\ group_7 + 1.14 \times Age\ group_8 + 0.051 \times CCIgroup_2 + 0.215 \times CCIgroup_3 + 0.201 \times CCIgroup_4 + 1.471 \times Respirator_2 + 1.132 \times CRRT,Dialysis_2 + 1.51 \times Drugs_2 + 0.263 \times Nrse\ grade_2 + 0.554 \times Nrse\ grade_3 + 0.988 \times Nrse\ grade_4 + 0.634 \times Nrse\ grade_5 + 1.265 \times Nrse\ grade_6 + 1.299 \times Nrse\ grade_7 + 1.228 \times Nrse\ grade_8 + 1.08 \times Nrse\ grade_9 + (-0.186) \times Special\ doctor_2$$

Supplementary Table 1. Calculating probability of ICU mortality: an example

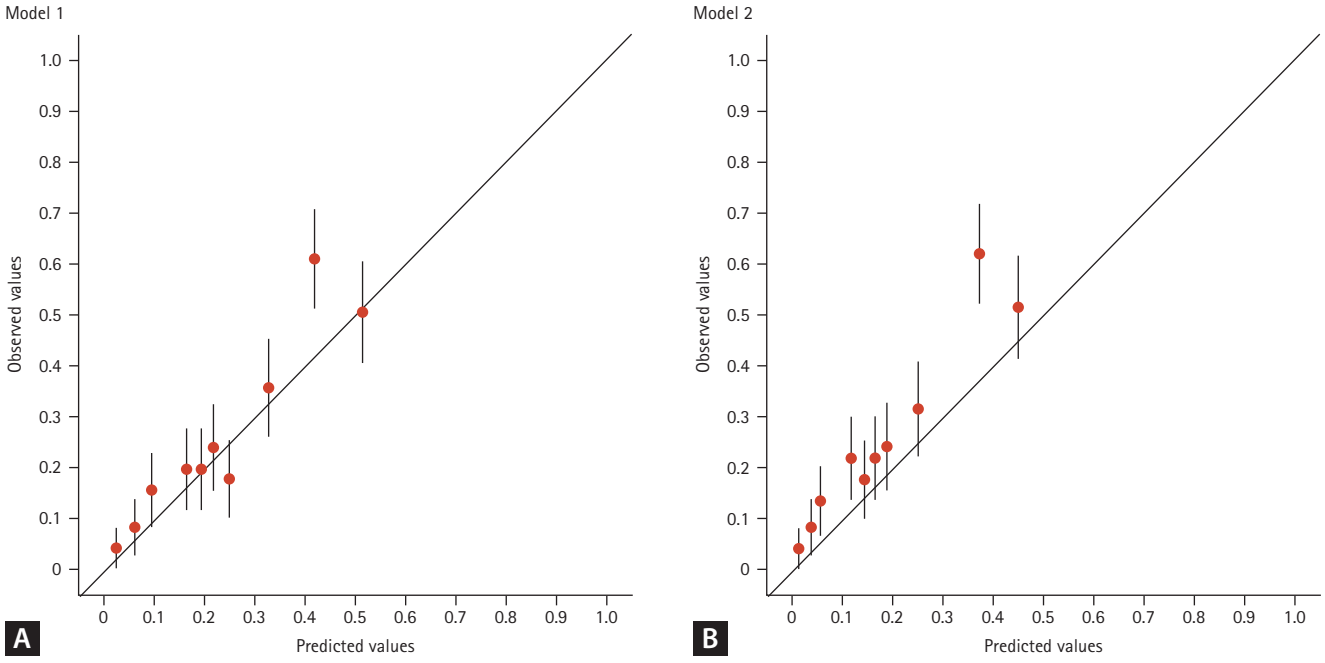
Variable	Model 1		
	β	x	βx
Constant	-4.336		-4.336
Sex			
Female	-0.085	0	0.000
Age group (yr)			
30–39	-0.224	0	0.000
40–49	0.101	0	0.000
50–59	0.251	0	0.000
60–69	0.275	0	0.000
70–79	0.514	1	0.514
80–89	0.948	0	0.000
90–99	1.449	0	0.000
CCI			
CCI 1	0.029	0	0.000
CCI 2	0.190	0	0.000
CCI 3+	0.221	1	0.221
Respirator			
Yes	1.220	0	0.000
CRRT, dialysis			
Yes	1.000	0	0.000
Norepinephrine, dopamine, vasopressin			
Yes	1.368	1	1.368

ICU, intensive care unit; CCI, Charlson comorbidity index; CRRT, continuous renal replacement therapy.
See the "Methods" section in the text for a description of this example patient and the variables of the calculations. g(x) = -2.233. This patient's probability of ICU mortality, as calculated using the Mortality Probability Model system model 1 is $\frac{e^{-2.233}}{1+e^{-2.233}}$ = 0.097.

https://doi.org/10.3904/kjim.2022.311

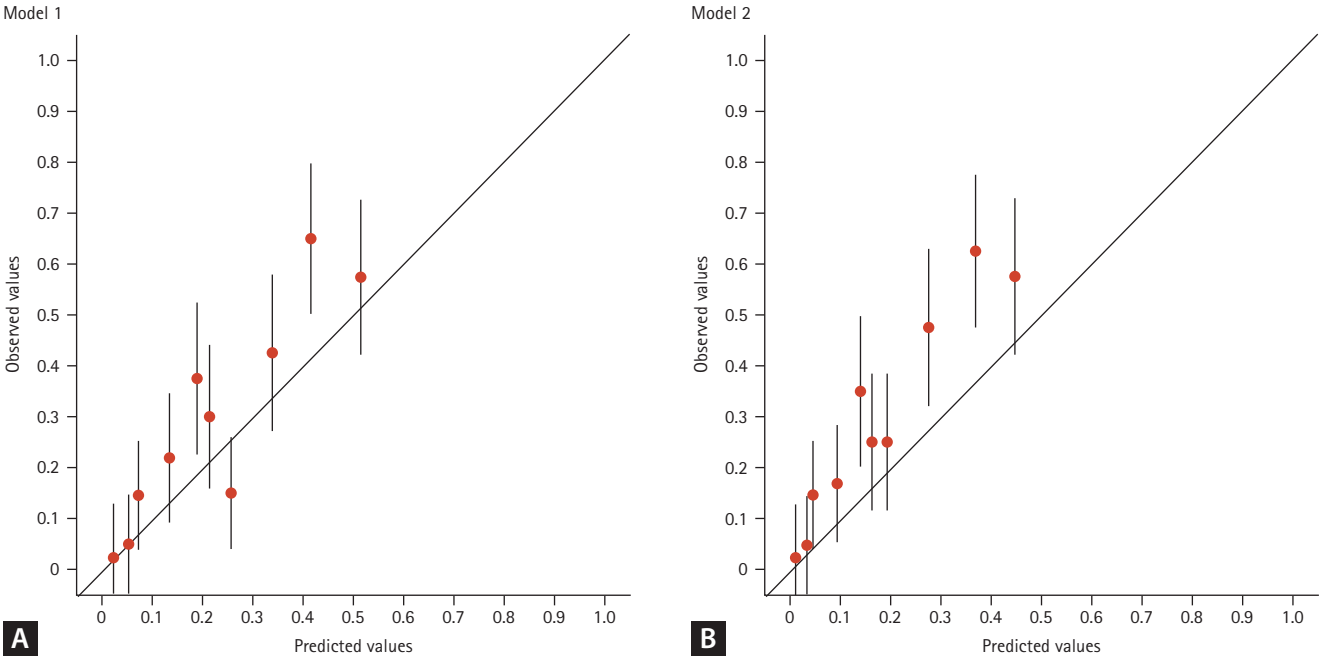


Supplementary Figure 1. Calibration plots for (A) model 1 and (B) model 2 for predicting in-hospital mortality (tertiary hospital 1).

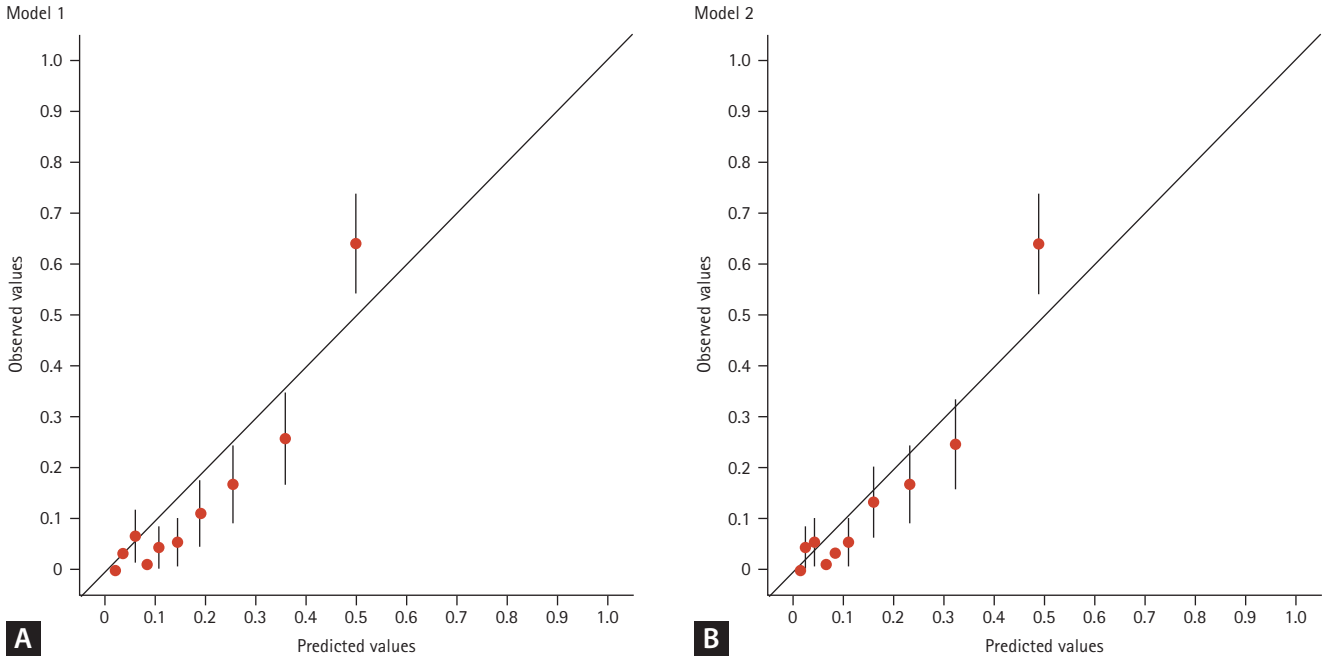


Supplementary Figure 2. Calibration plots for (A) model 1 and (B) model 2 for predicting ICU mortality (tertiary hospital 1). ICU, intensive care unit.

https://doi.org/10.3904/kjim.2022.311

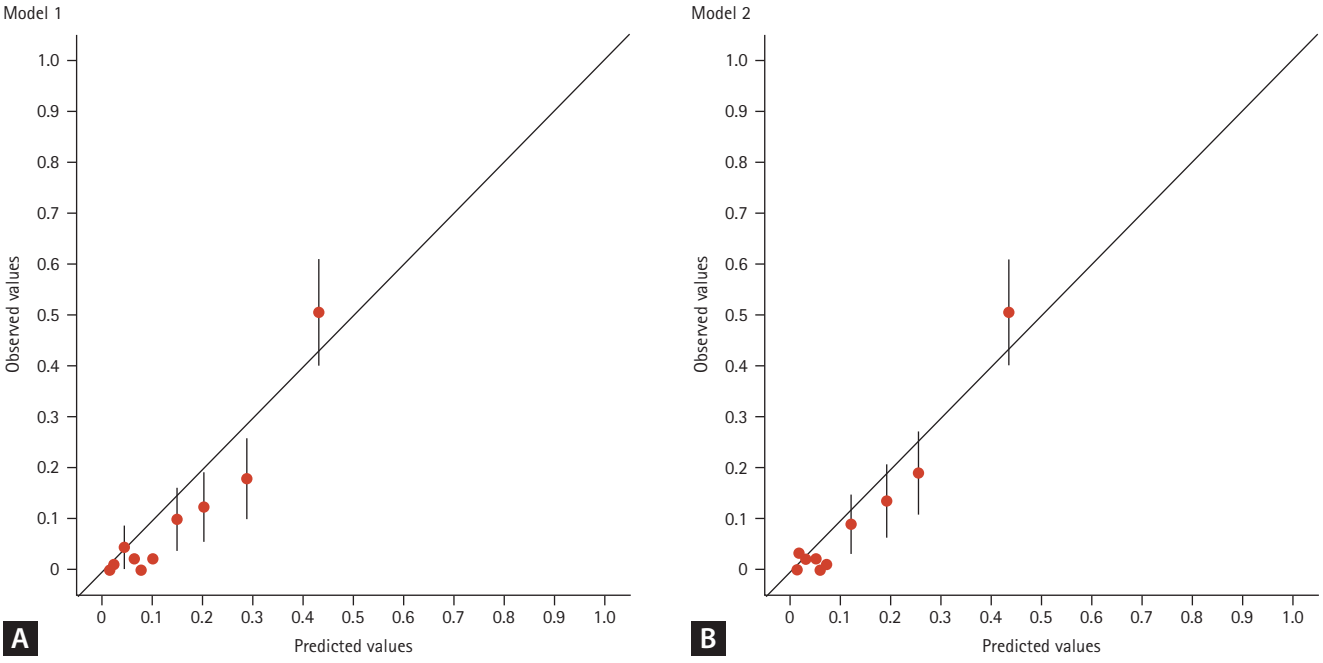


Supplementary Figure 3. Calibration plots for (A) model 1 and (B) model 2 for predicting ICU mortality (tertiary hospital 2). ICU, intensive care unit.



Supplementary Figure 4. Calibration plots for (A) model 1 and (B) model 2 for predicting in-hospital mortality (general hospital 1).

https://doi.org/10.3904/kjim.2022.311



Supplementary Figure 5. Calibration plots for (A) model 1 and (B) model 2 for predicting ICU mortality (general hospital 1). ICU, intensive care unit.