Predictive Values of Initial Severity Scores at Intensive Care Unit Admission for Patients With Overnight or Prolonged Stay

Koji Hosokawa (khosok@u-fukui.ac.jp)
University of Fukui: Fukui Daigaku https://orcid.org/0000-0002-8449-4460
Nobuaki Shime
Hiroshima University: Hiroshima Daigaku, Hiroshima University: Hiroshima Daigaku

---

Original research

Keywords: Acute Physiology and Chronic Health Evaluation score, Sequential Organ Failure Assessment score, Simplified Acute Physiology score, intensive care, mortality

Posted Date: September 30th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-915561/v1

Abstract

Background:

The predictive value of disease severity scores for intensive care unit (ICU) patients is occasionally inaccurate because ICU patients with mild symptoms are also considered. We, thus, aimed to evaluate the accuracy of severity scores in predicting mortality of patients with complicated conditions admitted for > 24 hours.

Methods:

Overall, 35,353 adult patients using nationwide ICU data were divided into two groups: (1) overnight ICU stay after elective surgery and alive on discharge within 24 hours and (2) death within 24 hours or prolonged stay. The performance and accuracy of Sequential Organ Failure Assessment (SOFA), Acute Physiology and Chronic Health Evaluation (APACHE) II and III, and Simplified Acute Physiology Score (SAPS) II scores in predicting in-hospital mortality were evaluated.

Results:

In the overnight stay group, the correlation between SOFA and APACHE III scores or SAPS II was low because many had a SOFA score of 0. In the prolonged stay group, the predictive value of SAPS II and APACHE II and III showed high accuracy but that of SOFA was moderate.

Conclusions:

When overnight ICU stay patients were not included, the high predictive value for in-hospital mortality of SAPS II and APACHE II and III was evident.

Background

Disease severity scores are widely used to evaluate patients in the intensive care unit (ICU) [1]; the common tools include Sequential Organ Failure Assessment (SOFA), Acute Physiology and Chronic Health Evaluation (APACHE), and Simplified Acute Physiology Score (SAPS). They help determine the severity of organ failure and inform on the required life-support treatments. All scores are also useful for predicting mortality [2, 3]; however, their accuracy differs.

Severity scores were developed for assessing critically ill ICU patients whose physiological parameters are considerably deranged [1]. However, patients admitted to the ICU also include those who have undergone elective surgeries. Furthermore, for those who undergo complicated surgeries, their physiological status need not be necessarily impacted. Hence, severity scores might not be suitable for evaluating disease conditions of mild severity. We hypothesized that the accuracy of severity scores for predicting mortality would improve if patients were categorized as having moderate or high severity and if those with postoperative overnight ICU stay who were alive on discharge within 24 hours were not included.

The aim of this study was to evaluate the accuracy of disease severity scores in predicting mortality in both the overnight ICU stay group and prolonged ICU stay group.

Methods

Study design and ethical concerns

The was a retrospective observational study conducted using a nationwide database. Ethical approval was obtained from Hiroshima University (E-1997, 2020.5.7), the principal researcher's affiliation. The authors were part of the Medical Insurance Strategy Committee of the Japanese Society of Intensive Care Medicine.

Selection of patients and database

We obtained data from the Japanese Intensive Care PAtient Database (JIPAD) for January–December 2018. Adult patients (aged ≥ 16 years per the JIPAD) were included. Patients were divided into two groups: (1) overnight ICU stay and alive at discharge within 24 hours and (2) prolonged ICU stay or died within 24 hours. The former group comprised planned admissions after elective surgery or involved performance of ICU procedures and planned discharge within 24 hours. The latter group comprised patients who died within 24 hours or with ICU stay for more than 24 hours.

We collected data on patient demographics, including age, sex, disease category, rate of sepsis, treatment in the ICU, duration of ICU stay, and in-hospital mortality. Sepsis cases included those whose major and secondary disease categories related to the sepsis category (numbers 501–504 in the JIPAD). Life-support treatments in the ICU included catecholamine infusion, mechanical ventilation, renal replacement therapy, and extracorporeal membrane oxygenation.

Outcomes

The primary outcome was the accuracy of predictive values for in-hospital mortality of different disease severity scores. This was separately analyzed for both the overnight and prolonged ICU stay group and then comparatively evaluated. The secondary outcomes were the difference in score values between groups and predictive values for prolonged ICU stay among cases with different scores. Additionally, the correlation between disease severity scores was evaluated to understand the performance of different scoring systems.

Statistical analysis

A histogram of the different disease severity scores was constructed, the median, interquartile range, and skewness were calculated, and the normality of distribution was examined using the Anderson–Darling test. The Wilcoxon test was used to determine the difference in the median values of disease severity scores between groups. For the correlation analysis, a non-linear regression analysis using the least-squares method was performed. Pearson’s correlation coefficient was calculated. $R^2 \geq 0.75$ was considered high; 0.50–0.75, moderate; and 0.25–0.50, low. For the analysis of prediction for prolonged ICU stay and in-hospital mortality, receiver operating characteristic (ROC) curves were constructed, and the area under the curve (AUC) was compared. An AUC ≥ 0.85 was considered highly accurate; 0.75–0.85, moderately accurate; and < 0.75, poorly accurate. Graphs were constructed and statistical analyses were conducted using JMP® Pro 15 (SAS Institute Inc., Cary, NC, USA). Sample size calculation was not performed because the research had an exploratory aim using a large database.

Results

Population

We recruited 35,353 adult ICU patients from the JIPAD (Table 1). The number of postsurgical patients was 24,935 (70.5%), including 20,268 (57.3%) who underwent elective surgeries. The patients were divided into the postoperative overnight ICU stay (n = 12,697 [35.9%]) and prolonged ICU stay (n = 22,656 [64.1%]) groups. The number of life-support treatments administered in the ICU was lower in the overnight than in the prolonged ICU stay group. In-hospital mortality was 0.8% for overnight ICU stay patients and 13.2% for prolonged ICU stay patients.

Table 1
Patient characteristics

	All patients (n = 35,353)	Overnight (n = 12,697)	Prolonged (n = 22,656)
Number of institutes that entered data	46	43	46
Age (years)	70 [60–78]	71 [60–79]	70 [60–77]
Male sex	61.4%	59.2%	62.6%
Postsurgical admission	70.5%	98.3%	55.0%
Elective surgery	57.3%	98.3%	34.4%
Emergent ICU admission	42.7%	0.0%	66.6%
Disease category			
Cardiovascular	34.5%	20.3%	42.5%
Pulmonary	15.1%	20.9%	11.8%
Digestive	19.6%	23.6%	17.3%
Cerebral	13.9%	17.0%	12.1%
Trauma	2.3%	0.2%	3.4%
Others	14.7%	18.0%	12.9%
Sepsis	4.4%	0.0%	6.8%
Treatment in the ICU			
Catecholamine use	45.5%	26.7%	56.0%
Mechanical ventilation	38.4%	11.4%	53.6%
Renal replacement therapy	8.4%	1.1%	12.4%
ECMO	1.2%	0.0%	1.8%
Duration of ICU stay (days)	1 [1–4]	1 [1–1]	3 [1–5]
Duration of hospital stay (days)	20 [12–38]	15 [10–27]	24 [13–44]
ICU mortality	4.0%	0.0%	6.3%
Hospital mortality	8.8%	0.8%	13.2%

Values are presented as percentages or medians (interquartile range). ECMO, extracorporeal membrane oxygenation; ICU, intensive care unit.

Patient severity scores

The values of four severity scores (SOFA, APACHE II, SAPS II, and APACHE III) at ICU admission were extracted from the JIPAD data. The median values of the four severity scores were lower for overnight ICU stay patients than for prolonged ICU stay patients (2 [1–4] vs. 5 [3–8], $p < 0.001$; 12 [9–14] vs. 16 [12–22], $p < 0.001$; 20 [15–26] vs. 34 [26–47], $p < 0.001$; 43 [34–53] vs. 62 [47–81], $p < 0.001$, respectively; Supplemental Fig. 1). In the histogram, the skewness of SOFA was similar between the overnight and prolonged ICU stay groups (0.951 and 0.836, respectively; Supplemental Fig. 1), whereas that of APACHE II, SAPS II, and APACHE III was lower for the overnight ICU stay patients than for prolonged ICU stay patients (0.344 and 1.233; 0.529 and 1.194; and 0.538 and 1.320, respectively; Supplemental Fig. 1). The normality of distribution was not proven for all types of analysis.

The ROC curves helped evaluate the predictive values of the different scores for prolonged ICU stay (Supplemental Fig. 2). The predictive value of APACHE II scores (AUC, 0.748) showed low accuracy. The predictive values of SOFA, SAPS II, and APACHE III scores (AUC = 0.778, 0.827, and 0.761, respectively) showed moderate accuracy. Thus, the differentiating performance of SAPS II between the overnight and prolonged ICU stay group was the best among the four severity scores.

Correlation between the different severity scores

The correlation between the disease severity scores was compared to illustrate the contrast between the overnight and prolonged ICU stay groups. The correlation between SOFA and APACHE III scores was low in the overnight group ($R^2 = 0.263$) and moderate in the prolonged group ($R^2 = 0.548$; Fig. 1A). Similarly, the correlation between SOFA and SAPS II scores was low in the overnight ($R^2 = 0.266$) and moderate in the prolonged group ($R^2 = 0.515$; Fig. 1B). The correlation between SAPS II and APACHE III scores was moderate in the overnight group ($R^2 = 0.622$) and high in the prolonged group ($R^2 = 0.828$; Fig. 1C). Correlations between SAPS II and APACHE II (R^2 = 0.442 and 0.760, respectively), between SOFA and APACHE II (R^2 = 0.182 and 0.491, respectively), and between APACHE II and APACHE III scores (R^2 = 0.558 and 0.822, respectively) were also observed in the overnight and prolonged groups.

Considering the different slopes of the regression curves, the quotient of the SOFA score divided by the APACHE III score was calculated and analyzed. The ratio was the proportional change in the difference in the slope of the linear correlation line between SOFA and APACHE III scores. The median of the ratio was lower in the overnight than in the prolonged ICU stay group (0.058 [0.034–0.087] vs. 0.085 [0.059–0.113], $p < 0.001$). The number of patients with a ratio of zero was higher than the numbers of patients for the other ratios in the overnight group. That property was blurred in the prolonged group (Supplemental Fig. 3A), implying that many patients had a SOFA score of 0. Likewise, the median of the SOFA and SAPS II ratio was lower in the overnight than in the prolonged ICU stay group (0.125 [0.074–0.188] vs. 0.152 [0.105–0.205], $p < 0.001$). The number of patients with a ratio of zero was more remarkable in the overnight group than in the prolonged group (Supplemental Fig. 3B). Additionally, the curved form of the correlation between SOFA and APACHE III scores was partially owing to a SOFA score of 0 or 1 for patients with low APACHE III scores (Fig. 1A, B). Thus, these results indicate the challenges associated with a 0 SOFA score in the overnight ICU stay group.

Accuracy of the scores for predicting in-hospital survival

The predictive values of the four scores for in-hospital survival were analyzed for the overnight and prolonged ICU stay patients. In the overnight ICU stay group, the predictive values of SOFA and APACHE II scores for in-hospital survival had low accuracy (AUC = 0.665 and 0.723, respectively), and those of SAPS II and APACHE III scores had moderate accuracy (AUC = 0.761 and 0.789, respectively; Fig. 2A). In the prolonged ICU stay group, the predictive value of the SOFA score was moderately accurate (AUC = 0.797), and those of SAPS II, APACHE II, and APACHE III scores were highly accurate (AUC = 0.855, 0.855, and 0.869, respectively; Fig. 2B).

Discussion

Using a large database, we evaluated the characteristics of the four severity scores between the patients in the postoperative overnight ICU stay and prolonged ICU stay groups. In the overnight ICU stay group, there was a poor correlation between SOFA and APACHE III scores or SAPS II. In the prolonged ICU stay group, the correlation between SOFA and APACHE III scores or SAPS II was moderately accurate. Moreover, the predictive values of SOFA scores for in-hospital survival showed poor or low accuracy in the overnight ICU stay group and moderate accuracy in the prolonged ICU stay group; the accuracy of the other three scoring systems was high in the prolonged ICU stay group. These results indicate that APACHE II and III scores or SAPS II would be more useful than SOFA scores at ICU admission for the prediction of in-hospital mortality, especially for patients with prolonged ICU stay.

A systematic review by Minne et al. showed that the AUC of SOFA scores for predicting survival was 0.61–0.88 [3]. They summarized that AUCs of SOFA scores were worse than those of APACHE II/III scores. In a study in the United Kingdom, the prediction values were as follows: APACHE II score (AUC = 0.62) and SOFA score (AUC = 0.61) [4]. Additionally, they described that mortality prediction with SOFA and APACHE II scores was not significant for medical patients; however, for surgical patients, it was significant with APACHE II score (AUC = 0.71) and SOFA score (AUC = 0.62). Our results are in line with the findings of these previous studies, highlighting the different predictive values among postoperative and other patient groups.

The authors have been interested in the admission, triage, and discharge (ADT) criteria of the ICU [5] owing to economic and financial concerns. In Japan, the government mandated a diagnosis-based insurance payment system. Recently, the national provision regarding ICU admission was revised. Hence, hospitals are required to report the SOFA scores at ICU admission and data of patients with sepsis. The ADT guideline documented that any scoring system alone cannot determine the level of care (recommendation 2C) [5]. This was based on a study that showed that a physician's prediction of patient survival was more accurate than that of scoring systems [6]. Moreover, physicians' and nurses' predictions improved the discriminative accuracy of in-hospital survival [7]. Concerning the SOFA score beyond a threshold value not being included in the ICU fee, we aimed to illustrate how the SOFA score alone was unlikely to predict the need for ICU stay.

The demand for ICU stay varies from rigorous monitoring to intensive mechanical life support. The real necessity of ICU stay is difficult to anticipate because ICU admission is subjective, and the severity of a disease does not clearly show a threshold for deciding admissions [8]. Additionally, the admission and triage criteria are occasionally not followed in cases of mass disaster or infection pandemics [5]. In the present study, we considered patients with postoperative overnight ICU stay, the aim of which is mainly concerned with patient monitoring. Irie et al. summarized general patient characteristics in the JIPAD data and analyzed the severity of disease among those requiring monitoring and critically ill adults [9]. They showed the difference in the standardized mortality ratio, APACHE III scores, SAPS II, and in-hospital mortality between the groups. Regarding patient classification criteria, we used the same grouping and the group names were changed to overnight and prolonged ICU stay; however, we also included the SOFA data and focused on the accuracy of the scoring system.

Our study has some limitations. First, the data in the JIPAD were from a limited number of ICUs in Japan. As half of the patients in the database were postsurgical patients, it is difficult to generalize the results to other types of ICUs. Second, the severity scores entered into the database were calculated in various ways. In some institutes, they utilized a semi-computerized method, wherein the physiological values in the electronic medical records were entered onto the data-acquisition platform and expert intensivists authorized by the JIPAD committee proved the total values. At other ICUs, the value was entered manually. Quality control was checked by the JIPAD; however, the quality may need to be validated carefully in each case. Moreover, SOFA scores were first included in the JIPAD in 2018, and our data comprised the SOFA values in the first year after inclusion. It is possible that the quality of SOFA score data would increase over time. Third, there is a possibility of errors in the database. Fourth, the category of postoperative overnight ICU stay patients might not be a well-defined grouping. The decision of a 1-day ICU stay depends not only on the mild disease status but also on various other factors.

Conclusions

Based on the post-hoc analysis of the large ICU database (JIPAD), the comparative analysis of ICU patients in the overnight and prolonged ICU stay groups revealed that the prediction of in-hospital mortality based on distinct disease severity scores (SOFA, APACHE II and III, and SAPS II) showed high accuracy for patients with prolonged ICU stay. For patients with overnight ICU stay, SOFA scores had a lower predictive value than APACHE III scores or SAPS II, probably because many patients had a SOFA score of 0; thus, a modification of predictive scores is necessary for overnight ICU patients.

Abbreviations

ADT: admission, triage, and discharge

APACHE: acute physiology and chronic health evaluation

AUC: area under the curve

ICU: intensive care unit

JIPAD: Japanese Intensive care PAtient Database

ROC: receiver operating characteristic

SAPS: Simplified Acute Physiology Score

SOFA: sequential organ failure assessment

Declarations

Ethics approval and consent to participate

The Ethical Committee for Clinical Epidemiological Research at Hiroshima University approved this study (E-1997, 2020.5.7). Consent to participate was obtained via an opt-out policy uploaded on the university's homepage.

Availability of data and materials

The original data were obtained from the data management committee of the JIPAD in May 2020. The datasets analyzed during the current study are not available from the authors because of their agreement with the committee.

Competing interests

The authors declare that they have no competing interests.

Funding

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

Authors' contributions

KH, Conceptualization, Data curation, Formal analysis, Methodology, Visualization, Roles/Writing - original draft; NS, Conceptualization, Supervision, Validation, Writing - review and editing.

Acknowledgements

We thank Dr. Yusuke Iizuka and Dr. Tetsuhiro Takei for their critical comments on the protocol of the study. We thank the data management committee of JIPAD and the Japanese Society of Intensive Care

Medicine. For statistical analysis, we were supported by Dr. Takahiro Tokunaga (Toyama Hospital, Fukui University).

References

1. Vincent JL, Moreno R. Clinical review: Scoring systems in the critically ill. Crit Care 2010;14(2):207. doi:10.1186/cc8204.
2. Ferreira FL, Bota DP, Bross A, Mélot C, Vincent JL. Serial evaluation of the SOFA score to predict outcome in critically ill patients. JAMA 2001;286(14):1754-8. doi:10.1001/jama.286.14.1754.
3. Minne L, Abu-Hanna A, de Jonge E. Evaluation of SOFA-based models for predicting mortality in the ICU: A systematic review. Crit Care 2008;12(6):R161. doi:10.1186/cc7160.
4. Gosling P, Czyz J, Nightingale P, Manji M. Microalbuminuria in the intensive care unit: Clinical correlates and association with outcomes in 431 patients. Crit Care Med. 2006;34(8):2158-66. doi:10.1097/01.CCM.0000228914.73550.BD.
5. Nates JL, Nunnally M, Kleinpell R, Blosser S, Goldner J, Birriel B, et al. ICU admission, discharge, and triage guidelines: A framework to enhance clinical operations, development of institutional policies, and further research. Crit Care Med. 2016;44(8):1553-602. doi:10.1097/CCM.0000000000001856.
6. Sinuff T, Adhikari NK, Cook DJ, Schünemann HJ, Griffith LE, Rocker G, et al. Mortality predictions in the intensive care unit: Comparing physicians with scoring systems. Crit Care Med. 2006;34(3):878-85. doi:10.1097/01.CCM.0000201881.58644.41.
7. Detsky ME, Harhay MO, Bayard DF, Delman AM, Buehler AE, Kent SA, et al. Discriminative accuracy of physician and nurse predictions for survival and functional outcomes 6 months after an ICU admission. JAMA 2017;317(21):2187-95. doi:10.1001/jama.2017.4078.
8. Marshall JC, Bosco L, Adhikari NK, Connolly B, Diaz JV, Dorman T, et al. What is an intensive care unit? A report of the task force of the World Federation of Societies of Intensive and Critical Care Medicine. J Crit Care 2017;37:270-6. doi:10.1016/j.jcrc.2016.07.015.
9. Irie H, Okamoto H, Uchino S, Endo H, Uchida M, Kawasaki T, et al. The Japanese Intensive care PAtient Database (JIPAD): A national intensive care unit registry in Japan. J Crit Care 2020;55:86-94. doi:10.1016/j.jcrc.2019.09.004.

Figures

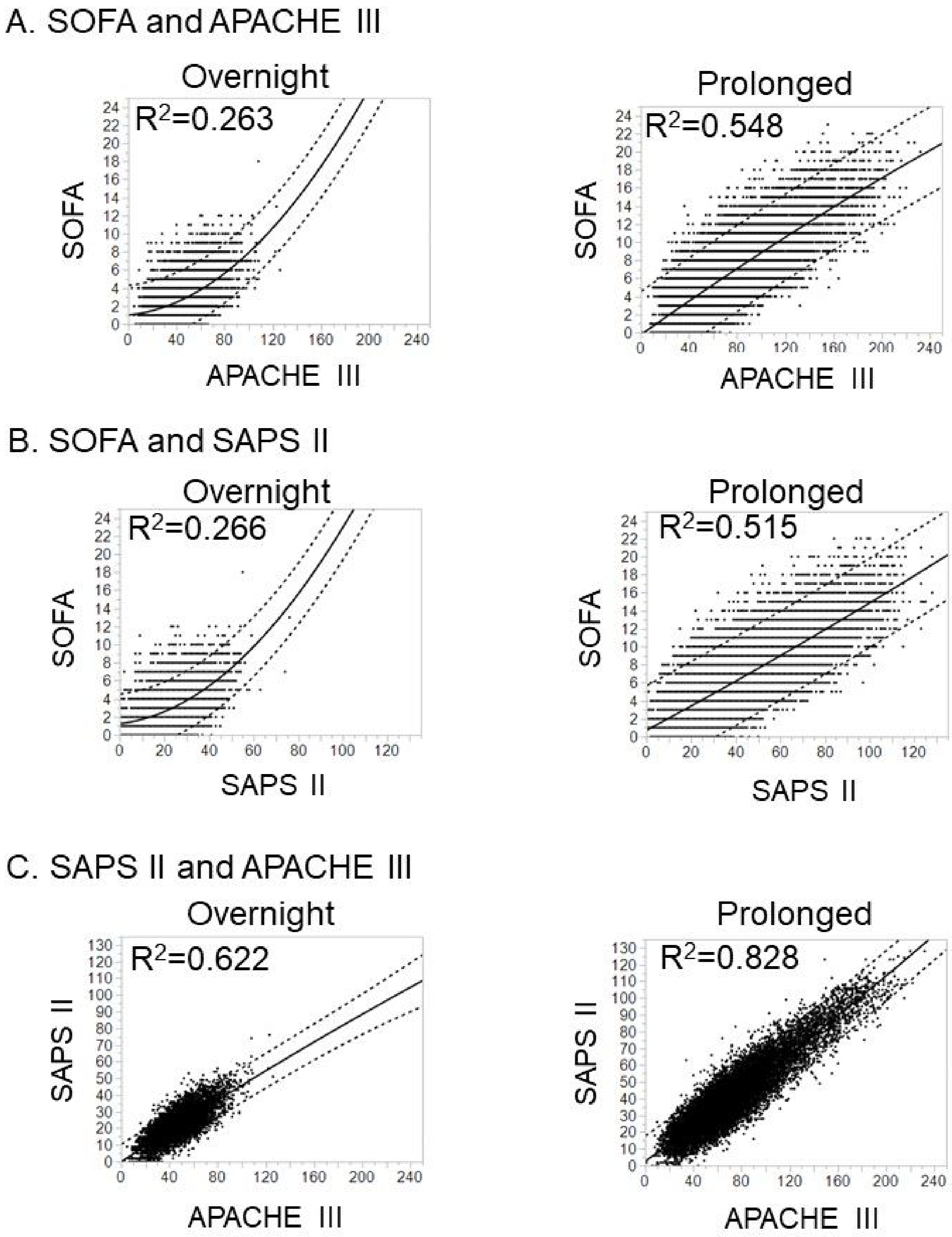

Figure 1

Relation of severity scores between the overnight and prolonged ICU stay groups. Scatter plots showing the relationship between SOFA and APACHE III scores in the overnight and prolonged ICU stay groups. Simple regression with a quadratic curve (black line) drawn with 95% confidence intervals (dotted lines). Pearson’s correlation analysis shows that the correlation between SOFA and APACHE III (A) or SAPS II scores (B) is low in the overnight ICU stay group. In the prolonged ICU stay group, there is a moderate correlation between SOFA and APACHE III scores or SOFA scores and SAPS II. (C) The correlation between APACHE III scores and SAPS II is high in the prolonged group.

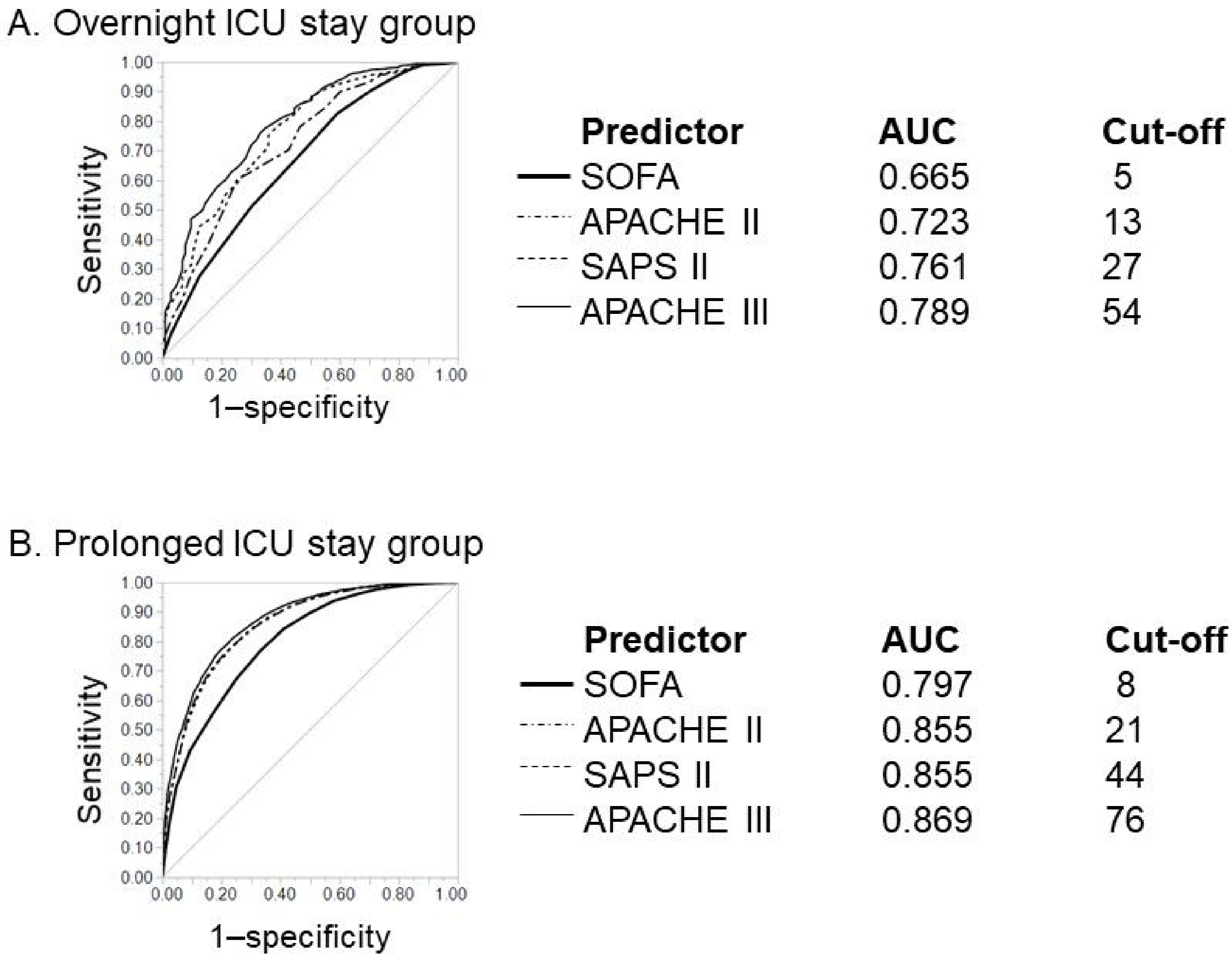

Figure 2

Predictive value of severity scores for in-hospital survival between overnight and prolonged ICU stay groups. In the overnight ICU stay group, the predictive values of SAPS II and APACHE III scores for in-hospital mortality have moderate accuracy (AUC = 0.75–0.85). In the prolonged ICU stay group, the prediction of the four scores for in-hospital mortality shows high accuracy.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- 12SuppFigures.pptx