Using patient admission characteristics alone to predict mortality of critically ill patients: a comparison of three prognostic scores

Kwok.M. Ho, Teresa A. Williams, Yusra Harahsheh, Thomas L. Higgins

PII: S0883-9441(15)00549-3
DOI: doi: 10.1016/j.jcrc.2015.10.019
Reference: YJCRC 51991

To appear in: *Journal of Critical Care*

Please cite this article as: Ho Kwok.M., Williams Teresa A., Harahsheh Yusra, Higgins Thomas L., Using patient admission characteristics alone to predict mortality of critically ill patients: a comparison of three prognostic scores, *Journal of Critical Care* (2015), doi: 10.1016/j.jcrc.2015.10.019

Using patient admission characteristics alone to predict mortality of critically ill patients: a comparison of three prognostic scores

Short running title: ICU admission prognostic scores

Kwok. M. Ho [1,2,3] ***, Teresa A. Williams*** [4] ***, Yusra Harahsheh*** [5] ***, Thomas L. Higgins*** [6]

[1] Intensivist, Department of Intensive Care, Royal Perth Hospital, Perth, Australia. Email: kwok.ho@health.wa.gov.au
[2] Clinical Associate Professor, School of Population Health, University of Western Australia, Perth, Australia
[3] Adjunct Associate Professor, School of Veterinary & Life Sciences, Murdoch University, Perth, Australia
[4] Senior Research Fellow, School of Nursing, Midwifery and Paramedicine, Curtin University, Perth, Australia. Email: teresa.williams@curtin.edu.au
[5] PhD Candidate, School of Population Health, University of Western Australia, Perth, Australia. Email: Yusra.Harahsheh@health.wa.gov.au
[6] Professor of Medicine, Surgery & Anesthesiology, Baystate Franklin Medical Center, Springfield, MA; Tufts University School of Medicine, Boston, MA, USA. Email: Thomas.Higgins@baystatehealth.org

Correspondence to: *Dr. K.M. Ho, ICU, Royal Perth Hospital, Wellington Street, Perth, WA 6000, Australia. Email: kwok.ho@health.wa.gov.au*

Conflict of interest statements

The author has no involvement with organisation(s) with financial interest in the subject matter of this study.

Abstract

Purpose: This study compared the performance of three admission prognostic scores in predicting hospital mortality.

Materials and Methods: Patient admission characteristics and hospital outcome of 9549 patients were recorded prospectively. The discrimination and calibration of the predicted risks of death derived from the Simplified Acute Physiology Score (SAPS III), Admission Mortality Prediction Model (MPM_0 III), and Admission Acute Physiology and Chronic Health Evaluation (APACHE II) were assessed by the area under the receiver-operating-characteristic curve (AUROC) and a calibration plot, respectively.

Measurements and Main Results: Of the 9549 patients included in the study, 1276 patients (13.3%) died after ICU admission. Patient admission characteristics were significantly different between the survivors and non-survivors. All three prognostic scores had a reasonable ability to discriminate between the survivors and non-survivors (AUROC for SAPS III: 0.836, MPM_0 III: 0.807, Admission APACHE: 0.845), with best discrimination in emergency admissions. The SAPS III model had a slightly better calibration and overall performance (slope of calibration curve: 1.03, Brier score: 0.09, Nagelkerke R^2: 0.297) compared to the MPM_0 III and Admission APACHE II model.

Conclusions: All three ICU admission prognostic scores had a good ability to predict hospital mortality of critically ill patients, with best discrimination in emergency admissions.

KEY WORDS: outcome; prognosis; prediction; risk adjustment; severity of illness

Abbreviations

APACHE, Acute Physiology and Chronic Health Evaluation

ICU, intensive care unit

MPM, Mortality Prediction Model

AUROC, area under the receiver-operating-characteristic curve

SAPS, Simplified Acute Physiology Score

SMR, standardized mortality ratio

Introduction

The ability to accurately adjust for severity of illness of critically ill patients is essential for research and quality assurance purposes [1]. Many prognostic models have been developed in the past three decades and each has its own strengths and weaknesses [1]. Some prognostic models, including the Acute Physiology and Chronic Health Evaluation (APACHE II-IV models), use the worst physiological parameters of the patients within the first 24 hours of intensive care unit (ICU) admission to estimate the risk of death [2], while others – including the Admission Mortality Prediction Model (MPM_0 III) and Simplified Acute Physiology Score (SAPS III) model - rely solely on patient characteristics on admission to the ICU to estimate the patient's risk of death [3,4]. Because the worst first 24-hour physiological data can be treatment dependent, it has been suggested that a higher risk of death may reflect poor clinical management rather than sicker patients [5]. In a randomized controlled trial setting, an active intervention that affects the biochemical and physiological parameters of a patient may also confound the interpretation of the difference in predicted risks of death between the intervention and control group, if the active intervention is initiated within the first 24 hours of ICU admission. Using patient admission characteristics alone before any ICU interventions to estimate their risk of death can avoid these criticisms and, at a practical level, collecting only admission data is also much easier than screening for the worst physiological and biochemical data over a 24-hour period.

Our previous work showed that using the ICU admission physiology and biochemical variables to calculate the Admission APACHE II score and predicted mortality was not substantially inferior to using the worst first 24-hour APACHE II score in predicting hospital mortality (area under the receiver-operating-characteristic

curve (AUROC) 0.838 *vs.* 0.846, respectively)[6]. Whether this modified use of the APACHE II model (Admission APACHE II) is comparable to other ICU admission prognostic scores in predicting hospital mortality remains uncertain. Although the APACHE II model is quite old now, it is still widely used in many ICUs for research and quality assurance purposes due to its ease of use. In this prospective cohort study, we compared the performance of three prognostic scores (SAPS III, Admission MPM_0 III, and Admission APACHE II models) that use patient admission characteristics alone in predicting the mortality of critically ill patients.

Materials and Methods

All patients who were admitted to the Royal Perth Hospital ICU between 1st January 2008 and 31 December 2013 were included in this study, except those who were readmitted to the ICU during the same hospitalization [2]. Royal Perth Hospital was an 800-bed university teaching hospital and the 22-bed ICU was a tertiary ICU that admitted critically ill adult patients of all specialties and was staffed by fully trained intensivists.

During the study period, all the components of the SAPS III and APACHE II and III scores including both the admission and worst first 24-hour physiology and biochemical data were recorded for all patients admitted to the ICU. After the patient was discharged from ICU, the data were checked for transcription errors and completeness by a designated trained clerical staff member using data from the computerized laboratory database and going through the ICU vital signs flow chart again before the data were transferred to the computer. A single data-custodian has been responsible for ensuring data quality. The data were reviewed for internal consistency before annual lock-down, and there were no patients lost to follow-up or

with missing data. The SAPS III, MPM_0III and Admission APACHE scores and predicted mortality were calculated as described by Moreno *et al*., Higgins *et al*., and Knaus *et al*., respectively (**Supplementary material Tables A and B and MPM_0 III calculator**)[2-4]. Admission Sequential Organ Failure Assessment (SOFA) score was not assessed in this study because our previous work has showed that it did not give better or additional predictive power when compared or added to the APACHE II model [7,8]. This study utilized only clinical data that were de-identified and was registered as a clinical audit with the Clinical Safety and Quality Unit (150521-02) and was exempt from review by the Royal Perth Hospital Ethics Committee.

Statistical analysis

We used the AUROC to assess the discrimination ability of the prognostic scores. The difference in AUROC derived from the same cases was calculated by the *z* statistic as described by Hanley and McNeil [9]. To assess the calibration of the model, we used a calibration plot to compare the predicted and observed risks of death. The slope and intercept of the calibration curve were calculated [10,11]. A calibration curve with a slope of 1 and an intercept of 0 indicates perfect calibration. If the slope of the curve is < 1, it indicates that the predicted risks of death are too extreme, with the model tending to underestimate the number of deaths in the low-risk strata and to overestimate number of deaths in the high-risk strata. Conversely, if the slope is > 1, this indicates that the predicted risks of death are not sufficiently different across the risk strata. An intercept < 0 indicates that the predicted risks of death are systematically too high and an intercept > 0 indicates that the predicted risks of death are systematically too low. Because the intercept can be affected by the slope of the calibration curve, as a standard procedure the intercept reported in this study

was estimated with the slope set to 1 [10,11]. The calibration of the model was also assessed by the Hosmer-Lemeshow chi-square statistics [12], with a p value < 0.05 suggestive of imperfect calibration.

We used the Brier score and Nagelkerke's R^2 to assess the overall performance of the three admission prognostic scores [13,14]. These two overall performance indices will reflect both the discrimination and calibration of a prediction model [11]. Brier score is calculated as $\sum (y_i-p_i)^2 / n$, where *y* denotes the observed outcome while *p* denotes the predicted probability of death for subject *i* in the data set of *n* subjects. Brier scores range from 0 to 0.25, with a Brier score of zero indicates a perfect prediction model and a Brier score of 0.25 signifies a useless prediction model [13]. The Nagelkerke's R^2 is a measure of variations explained by the prognostic score calculated on the log-likelihood scale [14]. A p-value <0.05 was taken as significant and no adjustment was made for multiple comparisons in the subgroup analyses. All statistical analyses were performed by SPSS for Windows (version 22.0, IBM, USA) and MedCalc for Windows (version 12.5, Ostend, Belgium).

Results

Of the 9549 patients included in the study, 1276 patients (13.3%) died after their first ICU admission. Patient admission characteristics including age, admission source, chronic health conditions, admission diagnosis and almost all admission physiological and biochemical parameters were significantly different between the survivors and non-survivors (**Table 1 and Supplementary material Table C**).

All three prognostic scores had a reasonable ability to discriminate between the survivors and non-survivors (AUROC SAPS III: 0.836, MPM_0 III: 0.807, Admission APACHE II: 0.845), with best discrimination noted in emergency

admissions (**Figure 1 and Table 2**). Statistically, the discrimination ability between the SAPS III and Admission APACHE II models were not significantly different (p=0.062), but both models were better than the MPM_0 III model (both p<0.001).

In terms of calibration and overall predictive performance of the models, the SAPS III model (slope of the calibration curve: 1.03, Brier score: 0.09, Nagelkerke R^2: 0.297) was the best among the three (**Figure 2 and Table 3**), with the MPM_0 III (slope of the calibration curve: 0.924) and Admission APACHE II models (slope of the calibration curve: 0.916) both tended to over-predict the risk of death for seriously ill patients especially when the predicted risks of death were >50%. The overall mean predicted risks of mortality (and standardized mortality ratio [SMR]) for the whole cohort of patients according to the SAPS III, MPM_0 III, and Admission APACHE II models were 14.6% (SMR=1.09), 23.0% (SMR=1.73), and 16.2% (SMR=1.22), respectively, confirming that the SAPS III model had the best overall calibration among the three admission scores.

Discussion

This study showed that using patient characteristics on ICU admission alone had a reasonably good ability to differentiate between survivors and non-survivors, with best discrimination in emergency ICU admissions. Among all three prognostic scores that are based on patient admission characteristics alone, the SAPS III score was marginally better than the MPM_0 III and Admission APACHE II models.

Firstly, our results showed that all prognostic models had a reasonably good ability to predict mortality of ICU patients, with best discrimination noted in emergency admissions consistent with previous reports. We had thus validated the predictive accuracy of these admission prognostic models (AUROC SAPS III 0.836

vs. 0.848; MPM_0 III 0.807 *vs.* 0.81; Admission APACHE II models 0.845 *vs.* 0.838)[4,6,15]. Judging from the discrimination ability alone, all three admission prognostic scores can be considered adequate in demonstrating baseline imbalance between groups in a clinical trial.

We noted that the calibration of these three admission prognostic models were quite different. Calibration of a prognostic model and the associated SMR can be affected by the performance of an ICU, and thus, a reliable calibration is essential before SMRs are used to benchmark performance of different ICUs. Our results showed that the calibration of SAPS III model was better than the MPM_0 III and Admission APACHE II models. The SAPS III model requires more patient data than the other two admission scores (and hence more resources are also needed to capture all the data), and thus it is not surprising for the SAPS III model to have better calibration across the full spectrum of severity of illness. Furthermore, we had used the SAPS III model calibrated for Australian ICUs in this study. The case mix in our ICU is different from those in the US, where the APACHE and MPM studies were derived (mean age of this cohort 53 *vs.* 64 and 63 years old in the MPM_0 III original and validation cohorts, respectively). This may explain why the SAPS III model was better calibrated than the APACHE and MPM model in this study [16]. Despite a good discrimination and calibration, the predicted risks derived from the SAPS III model or other admission prognostic models should not be used to replace clinical judgement as a triage tool in deciding who should receive life-support therapy, because prognostic models designed for evaluation of groups of ICU patients do not have sufficient sensitivity and specificity to apply to individual patients [17].

Secondly, our results showed that admission prognostic scores are likely to be less accurate than a prognostic model that utilizes the worst first 24-hour

physiological data [6,15]. This is, indeed, expected because assessing the worst physiological data over the first 24-hour period after ICU admission will have the benefits of assessing the response of the patients to life-support therapies. If a patient's physiological status gets worse despite life-support therapies, a higher risk of death is expected and this strengthens the performance of the prognostic model. The use of worst first 24-hour physiological data for risk adjustment appeared particularly important for elective ICU admissions because complications after major surgery may take time to manifest and their associated impact on outcome may be missed by using patient admission data alone [8]. Because elective admissions are more likely to cluster near the lower end of the spectrum in terms of severity of illness, this may also have affected the prognostic models' discrimination. Our results clearly demonstrated that the discrimination ability of all three ICU admission prognostic scores did not fare as well when applied to elective ICU admissions (**Table 2**). The reason behind why the Hosmer-Lemeshow chi-square statistics were apparently better in elective admissions may be due to the fact that Hosmer-Lemeshow chi-square statistics is sensitive to sample size [18] or most elective admissions are associated with a low predicted risk of mortality.

Finally, we would like to acknowledge the limitations of this study. Although we had included a reasonable number of patients, this was still a single center study. Hence, our results may not be applicable to centres with very different case-mix [16,19,20]. Our study was also underpowered to assess the difference in performance of the three admission prognostic scores in patients with different admission diagnoses [6,10] or whether these models' performance had deteriorated during the 6-year study period. That said, we noted that the Admission APACHE II model's performance had not substantially changed in the past 15 years [6]. We would also

like to acknowledge that new versions of the APACHE models have been widely used worldwide. Whether modifying the APACHE III or IV model using only patient admission (instead of worst first 24 hours) data will perform better than the Admission APACHE II model remains uncertain, but this merits further investigation.

In conclusion, all three ICU admission prognostic scores (SAPS III, MPM_0 III and Admission APACHE II) had a reasonably good ability to differentiate between survivors and non-survivors, with best discrimination in emergency admissions. The calibration and overall performance of the SAPS III model were slightly better than the MPM_0 III and Admission APACHE II models. The choice of using which ICU admission prognostic score for research and quality assurance purposes would depend on resources available and local calibration of the model.

Acknowledgements:

We would like to thank Drs. Geoffrey Clarke and John Weekes for their part in initiating the Royal Perth Hospital ICU database and Dr. Andrew Kramer for providing the MPM_0 III calculator for this study. This study was solely funded by Department of Intensive Care Medicine, Royal Perth Hospital. KMH is funded by Raine Clinical Research Fellowship from Raine Medical Research Foundation and WA Department of Health. No funding was received from the National Institutes of Health, NHMRC, and Wellcome Trust.

References

[1] Gunning K, Rowan K. ABC of intensive care: outcome data and scoring systems. *BMJ* 1999;319:241-4.

[2] Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Crit Care Med* 1985;13:818-29.

[3] Higgins TL, Kramer AA, Nathanson BH, Copes W, Stark M, Teres D. Prospective validation of the intensive care unit admission Mortality Probability Model (MPM_0-III). *Crit Care Med* 2009;37:1619-23.

[4] Moreno RP, Metnitz PG, Almeida E, Jordan B, Bauer P, Campos RA, et al.; SAPS 3 Investigators. 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;31:1345-55.

[5] Shann F. Mortality prediction model is preferable to APACHE. *BMJ* 2000;320:714.

[6] Ho KM, Dobb GJ, Knuiman M, Finn J, Lee KY, Webb SA. A comparison of admission and worst 24-hour Acute Physiology and Chronic Health Evaluation II scores in predicting hospital mortality: a retrospective cohort study. *Crit Care* 2006;10:R4.

[7] Ho KM, Lee KY, Williams T, Finn J, Knuiman M, Webb SA. Comparison of Acute Physiology and Chronic Health Evaluation (APACHE) II score with organ failure scores to predict hospital mortality. *Anaesthesia* 2007;62:466-73.

[8] Ho KM. Combining sequential organ failure assessment (SOFA) score with acute physiology and chronic health evaluation (APACHE) II score to predict hospital mortality of critically ill patients. *Anaesth Intensive Care* 2007;35:515-21.

[9] Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983;148:839-43.

[10] Ho KM. Forest and funnel plots illustrated the calibration of a prognostic model: a descriptive study. *J Clin Epidemiol* 2007;60:746-51.

[11] Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JD. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *J Clin Epidemiol* 2005;58:475-83.

[12] Lemeshow S, Hosmer DW. A review of goodness of fit statistics for use in the development of logistic regression model. *Am J Epidemiol* 1982;115:92-106.

[13] Arkes HR, Dawson NV, Speroff T, Harrell FE Jr, Alzola C, Phillips R, et al. The covariance decomposition of the probability score and its use in evaluating prognostic estimates. *Med Decis Making* 1995;15:120-131.

[14] Nagelkerke NJ. A note on a general definition of the coefficient of determination. *Biometrika* 1991;78:691-2.

[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;42:544-53.

[16] Kramer AA, Higgins TL, Zimmerman JE. Comparing observed and predicted mortality among ICUs using different prognostic systems: why do performance assessments differ? *Crit Care Med* 2015;43:261-9.

[17] Honeybul S, Ho K, O'Hanlon S. Access to reliable information about long-term prognosis influences clinical opinion on use of lifesaving intervention. *PLOS One* 2012;7:e32375.

[18] Peek N, Arts DG, Bosman RJ, van der Voort PH, de Keizer NF. External validation of prognostic models for critically ill patients required substantial sample sizes. *J Clin Epidemiol* 2007;60:491-501.

[19] Rowan KM, Kerr JH, Major E, McPherson K, Short A, Vessey MP. Intensive Care Society's APACHE II study in Britain and Ireland--II: Outcome comparisons of intensive care units after adjustment for case mix by the American APACHE II method. *BMJ* 1993;307:977-81.

[20] Murphy-Filkins R, Teres D, Lemeshow S, Hosmer DW. Effect of changing patient mix on the performance of an intensive care unit severity-of-illness model: how to distinguish a general from a specialty intensive care unit. *Crit Care Med* 1996;24:1968-73.

Table 1. Characteristics of the study cohort.

Variable	Total cohort (N=9549)	Survivors (n=8273)	Non-survivors (n=1276, 13.3%)	P value[#]
Age	53.2 (36-68)	51.6 (35-67)	62.8 (48-75)	0.001
Male, no. (%)	6284 (66)	5478 (66)	806 (63)	0.034
Admission source, no. (%)				0.001
- Operating theatre	4032 (42)	3725 (45)	307 (24)	
- Emergency Department	2840 (30)	2394 (29)	446 (35)	
- Ward	1141(12)	851 (10)	291 (23)	
- Other critical care areas	152 (1)	122 (2)	30 (2)	
- Other hospital	1283 (14)	1104 (13)	179 (14)	
- Other hospital ICU	100 (1)	77 (1)	23 (2)	
Elective surgery, no. (%)	1884 (20)	1824 (22)	60 (5)	0.001
Hospital stay prior to ICU, days	4 (2-10)	4 (2-10)	5 (2-13)	
Mechanical ventilation on adm., (%)	7320 (77)	6303 (76)	1017 (80)	0.003
Acute renal failure in ICU, no. (%)[*]	532 (6)	256 (3)	276 (22)	0.001
Worst 24-hr APACHE II score	17.0 (7.7)	16 (12-21)	27 (22-32)	0.001
Adm. APACHE II score	12 (8-18)	12 (8-16)	21 (16-26)	0.001
SAPS III score	43 (34-53)	41 (33-50)	60 (50-69)	0.001
Adm. APACHE II predicted risk, %	7.7 (3-21)	6.3 (3-16)	36.3 (18-60)	0.001
SAPS III predicted risk, %	7.9 (3-20)	6.3 (2-16)	32.8 (16-50)	0.001
MPM_0 III predicted risk, %	15.6 (8-31)	13.6 (7-26)	42.3 (23-71)	0.001
Length of ICU stay, days	3 (2-6)	3 (2-6)	4 (2-7)	0.001
Length of hospital stay, days	13 (6-25)	14 (7-26)	6 (3-17)	0.001
Chronic health conditions (%):[*]				
- Respiratory	446 (5)	377 (5)	69 (5)	0.176
- Cardiovascular	937 (10)	786 (10)	151 (12)	0.011
- Liver	237 (3)	176 (2)	61 (5)	0.001
- Renal	472 (5)	357 (4)	115 (9)	0.001
- Immune disease	87 (1)	63 (0.8)	24 (1.9)	0.001
- Immune treatment	338 (4)	250 (3)	88 (7)	0.001
- Metastatic cancer	116 (1)	89 (1)	27 (2)	0.003
- Lymphoma	51 (0.5)	34 (0.4)	17 (1.3)	0.001
- Leukaemia / myeloma	117 (1)	76 (0.9)	41 (3)	0.001
- AIDS	9 (0.1)	3 (0.04)	6 (5)	0.001
Major diagnoses, no. (%):				
Cardiac or respiratory arrest	465 (5)	239 (3)	226 (18)	0.001
Pneumonia	393 (4)	331 (4)	62 (5)	0.151
Septic shock	639 (7)	476 (6)	163 (13)	0.001
Multiple trauma	696 (7)	648 (8)	48 (4)	0.001
Head trauma	864 (9)	732 (9)	132 (10)	0.084
Intracranial haemorrhage	314 (3)	197 (2)	117 (9)	0.001
Drug overdose	672 (7)	662 (8)	10 (0.8)	0.001
Congestive heart failure, ischaemic heart disease, cardiogenic shock	240 (3)	169 (2)	71 (6)	0.001
Peripheral vascular disease or aortic aneurysm	275 (3)	248 (3)	27 (2)	0.087
GI obstruction or perforation	189 (2)	159 (2)	30 (2)	0.330
Aspiration	114 (1)	103 (1)	11 (0.9)	0.270
Obstructive airway disease	186 (2)	172 (2)	14 (1)	0.016
Heart valve surgery	633 (7)	615 (7)	18 (1)	0.001
Coronary artery bypass graft	1281 (13)	1246 (15)	35 (3)	0.001
Acute lung injury	40 (0.4)	33 (0.4)	7 (0.5)	0.481
Gastrointestinal bleeding	169 (2)	145 (2)	24 (2)	0.732
Pulmonary embolism	31 (0.3)	23 (0.3)	8 (0.6)	0.058

All values are median and interquartile range in parenthesis unless stated otherwise. Adm., Admission. GI, Gastrointestinal. ICU, intensive care unit. APACHE, Acute Physiology and Chronic Health Evaluation. MPM_0, Mortality Prediction Model on admission. SAPS, Simplified Acute Physiology Score. *According to the definitions by the APACHE model. [#] P values generated by either Mann-Whitney or Chi-square test.

Table 2. The areas under the receiver operating characteristic (AUROC) curve of the Mortality Prediction Model (MPM_0 III), Simplified Acute Physiology Score (SAPS III), and Admission Acute Physiology and Chronic Health Evaluation (Admission APACHE II) predicted risks in predicting hospital mortality.

Predictive model	**Mean AUROC (95% confidence interval)**		
	All patients (N=9549)	**Elective admissions** (n=1884)	**Emergency admissions** (n=6582)
MPM_0 III	0.807 (0.794-0.820)	0.717 (0.647-0.788)	0.801 (0.786-0.815)
SAPS III	0.836 (0.825-0.847)	0.761 0.703-0.819)	0.820 (0.807-0.834)
Admission APACHE II	0.845 (0.834-0.856)	0.805 (0.751-0.859)	0.830 (0.817-0.843)

The AUROC for the SAPS III and Admission APACHE II models were not statistically different ($p=0.062$), but both models were better than the MPM_0 III model (both $p<0.001$).

Table 3. The differences in the calibration (including the intercept and slope of the calibration curves) and overall performance measures of the Mortality Prediction Model (MPM_0 III), Simplified Acute Physiology Score (SAPS III), and Admission Acute Physiology and Chronic Health Evaluation (Admission APACHE II) predicted risks in predicting hospital mortality.

	Brier score	**Nagelkerke R^2**	**H-L χ^2 (p value)**	**Intercept (SE)**	**Slope (SE)**
MPM_0 III:					
All patients	0.11	0.260	62 (0.001)	-0.842 (0.04)	0.924 (0.03)
Elective admissions	0.04	0.111	6 (0.655)	-1.369 (0.28)	1.036 (0.16)
Emergency admissions	0.13	0.258	55 (0.001)	-1.048 (0.4)	0.925 (0.03)
SAPS III:					
All patients	0.09	0.297	119 (0.001)	-0.100 (0.05)	1.026 (0.03)
Elective admissions	0.03	0.100	12 (0.176)	0.950 (0.28)	0.764 (0.11)
Emergency admissions	0.10	0.288	72 (0.001)	-0.109 (0.05)	1.005 (0.04)
Admission APACHE II:					
All patients	0.09	0.296	185 (0.001)	-0.367 (0.04)	0.916 (0.03)
Elective admissions	0.03	0.101	19 (0.016)	0.08 (0.38)	1.216 (0.14)
Emergency admissions	0.10	0.293	109 (0.001)	-0.458 (0.05)	0.896 (0.03)

H-L, Hosmer-Lemeshow C statistic. SE, standard error. Intercepts of the calibration curve were set to a slope of one.

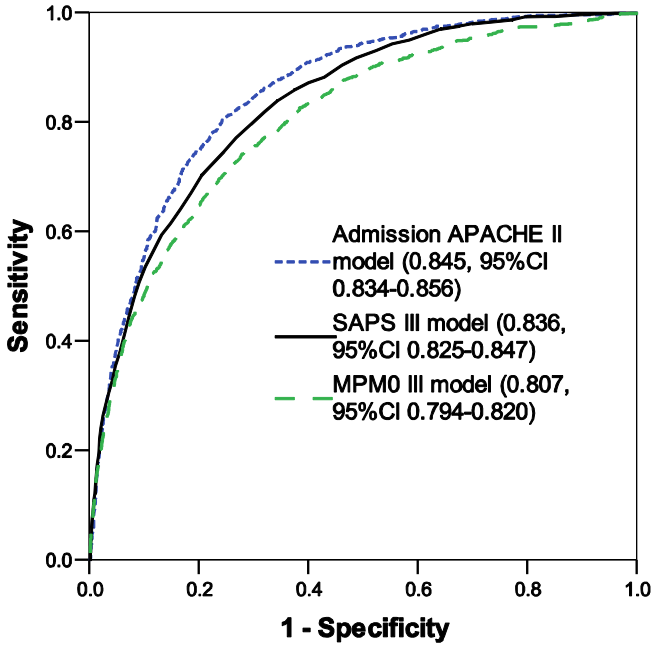

Figure 1

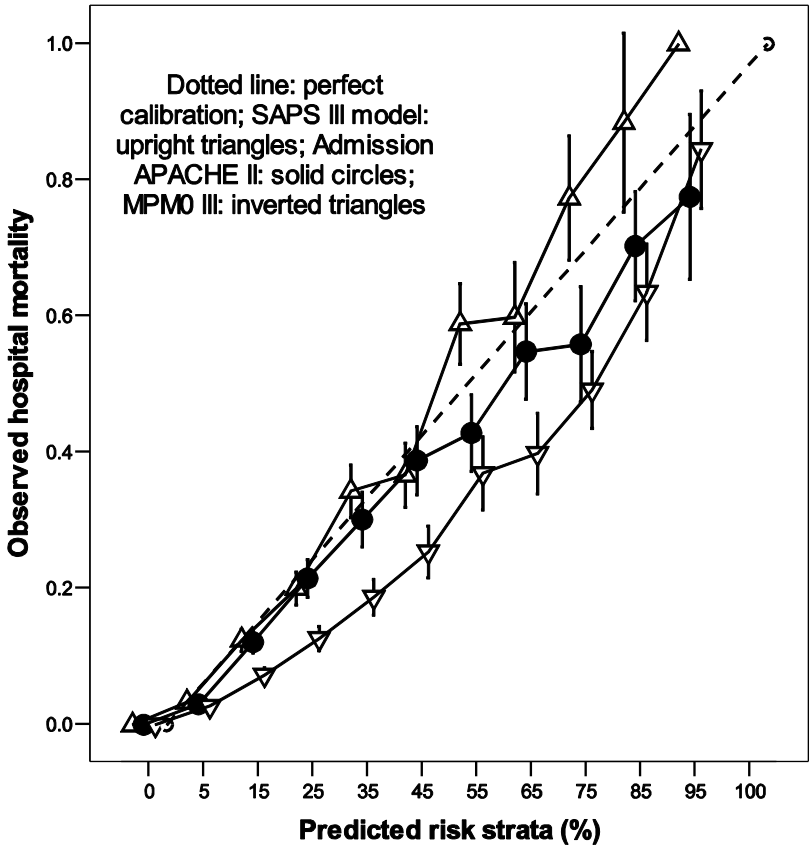

Figure 2