SUMMARY OF NQF-ENDORSED INTENSIVE CARE OUTCOMES MODELS FOR RISK ADJUSTED MORTALITY AND LENGTH OF STAY ($ICOM_{mort}$ and $ICOM_{LOS}$)

BACKGROUND

Importance of the ICU

The modern intensive care unit (ICU) is the highest mortality unit in any hospital. There are approximately 4 million ICU admissions per year in the United States with average mortality rate reported ranging from 8-19%, or about 500,000 deaths annually.[1-3] The ICU is also one of the sites in which medical errors are most likely to occur because of the complexity of care.[4,5] Since the patient population is severely ill and undergoes multiple complex interventions at the same time, these patients are extremely vulnerable to experiencing adverse outcomes.[6,7] In addition to its impact on mortality, critical care is a costly component of the national health care budget, with costs estimated to be $81.7 billion by 2005, accounting for 13.7% of hospital costs, 4.1% of national health expenditures, and 0.66% of the gross domestic product.[8] These costs are largely explained by the length of stay (LOS) in the ICU.[9,10] For these reasons, there has been substantial interest in measuring ICU outcomes, both in terms of mortality and resource utilization.

Variation in ICU Mortality and Resource Utilization

Considerable variation in mortality has also been observed among ICU patients, which persists even after adjustment for patient characteristics present at admission.[11-16] Twofold to threefold differences in ICU risk adjusted mortality that were previously reported[12] are still present in modern ICUs, irrespective of the model that is used to adjust for patient severity of illness.[13] Similar variation has also been seen in ICU length of stay (LOS), again even after accounting for patient risk factors.[16-19]

Extant ICU Outcome Risk Adjustment Models up to 2011

Clinicians and researchers have long recognized how important ICU performance is to overall hospital mortality. A significant amount of work has already been done to develop tools to assess ICU performance. The three most widely used general mortality risk adjustment models are the Mortality Probability Model (MPM), the Acute Physiology and Chronic Health Evaluation, and the Simplified Acute Physiology Score (SAPS). In a comparative analysis of these models, it has been shown that a hospital's mortality performance assessed by a standardized mortality ratio is not much impacted by which model is chosen.[17]

Of note, all models have been shown to need frequent recalibration.[20-24] That is, while it is reasonable to continue using the same clinical variables, the coefficients on those variables, and the possibility of interactions among the variables needs to be evaluated frequently (and whenever the models are applied to a new population).[25-30]

In addition, it has been shown that the variables used in each of these models can be incorporated into risk adjustment systems for ICU length of stay (LOS).[17] Again, the choice of model among the MPM, SAPS, and APACHE systems did not have much impact on each hospital's performance assessment.

There is, however, an important and stable difference between the models, and this relates to the data collection burden. Among randomly selected patients, the average time for chart abstraction for the MPM, SAPS, and APACHE models was 11, 20, and 37 minutes, respectively.[13]

In summary, the prior literature suggests that choice of model has little impact on hospital performance assessments, but major impact on data collection costs, with MPM being by far the least burdensome. For these reasons, we have recommended that models incorporating the MPM variables be used as the primary method of risk adjustment of ICU outcomes and we submitted such models to the National Quality Forum for evaluation.

Summary

Based on the clinical and economic significance of the ICU and the evidence that ICU performance varies, the National Quality Forum has endorsed measures of ICU outcomes (risk adjusted mortality and length of stay) for public reporting. In this document, we describe the requests made by the critical care community during evaluation of these models by the National Quality Forum and subsequently by the Hospital Quality Alliance. We then explain how we have adjusted our models—which we now refer to as the Intensive Care Outcomes Models ($ICOM_{mort}$ and $ICOM_{LOS}$ for ICU mortality and ICU LOS, respectively), in response to those requests.

RECENT POLICY DECISIONS RELATED TO ICU OUTCOMES

Evaluation of ICU Outcome Risk-Adjustment Models by the National Quality Forum and the Hospital Quality Alliance

During the National Quality Forum comment periods and deliberations, concern was raised about the potential for code status (whether a patient or his family allowed the hospital to provide all possible resuscitative support) would influence performance. In addition, some in the critical care community were concerned that public reporting of ICU performance would create an incentive for referral institutions to refuse to accept complex cases in transfer. Thus, although both the National Quality Forum and the Hospital Quality Alliance endorsed the models, they requested that transfer patients be excluded and a variable be included indicating "full code" status or not. In addition, both organizations stipulate that any performance measure cannot be proprietary, so adopting versions of these models that were copyrighted was not an option.

Response to National Quality Forum and Hospital Quality Alliance Evaluation of ICU Outcome Risk Adjustment Models

At the time of National Quality Forum and the Hospital Quality Alliance endorsement, no existing ICU outcomes model had been calibrated and validated to meet these specifications. We therefore started with the MPM_0-II model (the last non-proprietary version of this model) and added the full code variable, then assessed the model on a population of that excluded transfers. Because of concerns about calibration without considering interactions among the clinical variables,[25] we convened a clinical panel to suggest candidate interactions to be evaluated.

HOSPITAL AND PATIENT SAMPLE

Hospital Sample

The participating hospitals were those who voluntarily contribute patient-level data for public reporting of ICU outcomes in the state of California. In 2009, this sample consisted of 196 hospitals, representing a diverse group of institutions. Hospitals were asked to collect data on the first 100 consecutive patients per quarter who were discharged from the hospitals and had a stay in any of the hospitals' ICUs.

Patient Sample

The inclusion and exclusion criteria reflect the parameters already established by the pre-existing Mortality Probability Model (MPM_0-II), from which the $ICOM_{mort}$ and $ICOM_{LOS}$ models evolved. In 2009, there were 68,122 eligible patients with complete data for risk adjusted mortality calculation. This sample was split into a 60% development set (40,395) and 40% validation set (27,187) for analyses.

Inclusion criteria:

1. Age 18 or older

The original model was developed on a population ≥ 18 years old.[26] The clinical spectrum of diseases for children is significantly different from adult illnesses.

2. Stay in the ICU for at least four hours

Patients are sometimes admitted to the ICU for very short stays for a variety of administrative reasons (such as absence of other beds) or for periprocedural sedation. If these are the only reasons for an ICU stay, these patients are quite distinct from the typical ICU population. Another group of patients with short ICU stays—those who die within a few hours of admission—likely have outcomes determined entirely by clinical events occurring—and care provided prior to—the ICU stay. Therefore, all patients with ICU stays less the four hours are excluded.

Exclusion criteria:

1. Burn patients

Burn patients were excluded from the original model's development population. Physiologic and clinical variables to predict mortality in burn patients are considerably different than those used to predict mortality in a general ICU population. Often these patients are treated in separate, specialized units. Furthermore, specific prognostic systems have been previously developed for this subset of patients.[31]

2. Trauma patients

Currently, in most parts of the United States, trauma patients who are critically ill go to designated regional trauma centers. Thus, those centers would have trauma patients while other hospitals in the region would not. Furthermore, specific prognostic systems have been previously developed for trauma patients[32] and would be more useful for assessing the performance of regional trauma centers (if this were desired) than general ICU models.

3. Coronary artery bypass grafting surgery (CABG)

CABG patients represent a specialized group whose physiologic derangements do not predict the same risk of mortality as other patients in the ICU. Like trauma and burn patients, specific prognostic systems have been previously developed for this subset of patients as well.[33]

4. Patients admitted to rule out myocardial infarction (MI) that are found within 24 hours of ICU admission to not have a MI or another critical illness

Individuals who "rule out" for MI essentially are admitted to the ICU for monitoring of chest pain or a similar symptom. When this symptom is not due to ischemia (or another accepted reason for ICU admission, such as rupture of a thoracic aortic aneurysm), their risk of death is close to zero. Thus, variation in hospital policies about what percentage of patients are admitted to rule out for MI could have a large influence on calculated performance (hospitals that admitted many such patients would have lower than predicted mortality). Since such policies are known to vary and could significantly affect performance, rule out MI patients who are found not have an MI or other critical illness are excluded.

5. Readmissions

Readmissions to the ICU during the same stay are excluded since interventions during the first ICU admission may impact the patient's risk of mortality in the second admission.

6. Transfers from another acute care hospital

Pre-ICU treatments have the potential to alter the relationship between physiologic scoring and outcome.[34] The relationship between lead-time bias and patient outcomes is complex, having inconsistent effects on outcome, and often differing by patient type.[25] Previous reports have also demonstrated the potential negative impact that patients transferred into ICUs might have on the accepting center's outcome measures.[35,36] Transferred patients are therefore excluded from the sample.

MODEL DESCRIPTION AND DEVELOPMENT

The $ICOM_{mort}$ and $ICOM_{LOS}$ models evolved from the MPM_0-II, which itself was developed as an updated and revised version of the original MPM. The goal of the MPM developers was to construct a model that would accurately predict the mortality experience of a patient sample using the fewest variables required to discriminate and calibrate well.[26] Only variables that had clear definitions, could be easily obtained, and could be reliably collected were included in the final model. The model did not require the data collectors to obtain a primary reason for admission. All variables were collected in the window from one hour prior to ICU admission to one hour after ICU admission. Link to Data dictionary. Link to Data Collection tool.

Assessment of Interactions among Clinical Risk Factors

In our dataset, we found that a base model containing only the MPM_0-II variables with the addition of a full code status variable after excluding transferred patients overpredicted mortality, particularly in the higher ranges of risk. Given this overprediction, the most plausible explanation was that there were interactions among clinical variables for those patients with multiple risk factors. For this reason, we convened a clinical panel that suggested evaluation of the following interactions (Table 1).

Table 1. Candidate Interactions between Selected ICOM Clinical Variables

Interaction terms	Rationale for assessing possible interaction
Acute renal failure x chronic renal insufficiency	In the absence of chronic renal insufficiency, a greater insult is required to cause acute renal failure.
Acute renal failure x systolic blood pressure ≤ 90	Low systolic blood pressure may cause acute renal failure.
GI bleed x heart rate ≥ 150 beats/min	High heart rate may indicate a worse GI bleed.
GI bleed x systolic blood pressure ≤ 90	Low systolic blood pressure may indicate a worse GI bleed.
GI bleed x cirrhosis	GI bleed likely to be worse with cirrhosis, but otherwise protective.
CPR before admission x mechanical ventilation	Representative of combined signs of severe insult.
CPR before admission x	Representative of combined signs of severe insult.

coma/deep stupor	
Coma/deep stupor x mechanical ventilation	Representative of combined signs of severe insult
Cerebrovascular incident x coma/deep stupor	Representative of more severe cerebrovascular insult.
Cerebrovascular incident x intracranial mass effect	Representative of more severe cerebrovascular insult.
Intracranial mass effect x coma/deep stupor	Representative of more severe intracranial mass effect.
Cardiac dysrhythmia x heart rate ≥ 150 beats/min	Dysrhythmia may cause higher heart rate.

Treatment of Age in $ICOM_{mort}$

Prior research has shown that the relationship between age and risk is not necessarily simply linear.[25] In our development dataset, univariate plots of mortality risk versus age suggested increasing risk starting in the mid 60s (most particularly at age 65) and again in the mid 80s (most particularly at age 84). Therefore, we modeled age using splines with knots at 65 and 84. These were implemented by including a term which is the maximum of 0 or age minus the knot value.

The complex relationship of age with the other MPM risk factors was further evaluated using age interaction terms similar to the methods used by the developers of the MPM_0-III.[25] Interactions were considered between age and all of the other MPM risk factors.

Length of Stay Model

For the $ICOM_{LOS}$ model, we used methods similar to those in which we previously validated a model using the same set of variables on patients from 2001-2004.[17] For this model, LOS was truncated at 30 days. Variables and candidate interactions were the same as those evaluated in $ICOM_{mort}$.

Estimation of Models

In other studies in which the ICU risk adjustment models have been applied to populations distinct from the ones on which they were developed, each model has maintained adequate discrimination but has shown poor calibration.[20-24] To improve the calibration of our model, we re-estimated the coefficients in the models on our local sample using methods similar to prior studies that also customized the models to new populations. Therefore we divided our data into a randomly selected model development set (60% of the sample, the group on which the model variables were selected) and a model validation set (40% of the sample, the group on which we confirmed adequate calibration).

RESULTS

The coefficients of our customized models for the estimation samples are shown in Table 2 below.

Table 2. $ICOM_{mort}$ and $ICOM_{LOS}$ Model Re-estimated Coefficients

Variable	$ICOM_{mort}$		$ICOM_{LOS}$	
	Coefficient	p-value	Coefficient	p-value
Constant	-5.707	<.0001	0.032	0.9491
Physiology				
Coma/deep stupor (GCS 3 or 4)	1.037	0.0017	1.871	0.0003

Heart rate ≥ 150 beats/min	2.020	<.0001	1.347	0.0105
Systolic blood pressure ≤ 90	0.919	<.0001	1.257	<.0001
Chronic diagnoses				
Chronic renal insufficiency	0.939	0.0002	0.267	0.4096
Cirrhosis	1.693	0.0015	0.827	0.2802
Metastatic neoplasm	2.826	<.0001	0.993	0.0082
Acute diagnoses				
Acute renal failure	1.588	<.0001	2.056	<.0001
Cardiac dysrhythmia	-0.181	0.4104	0.305	0.2424
Cerebrovascular incident	1.655	<.0001	1.963	<.0001
GI bleed	0.536	0.1206	-0.835	0.0305
Intracranial mass effect	-0.171	0.7102	0.549	0.3269
Other				
Age (per year)	0.032	<.0001	0.017	0.0173
Age spline age 65	0.011	0.0177	-0.008	0.1590
Age spline age 84	0.022	0.0268	-0.032	0.0346
CPR before admission	1.766	<.0001	-0.334	0.5464
Mechanical ventilation within 1 hr of admission	1.388	<.0001	2.738	<.0001
Medical or unscheduled surgical admit	2.404	<.0001	1.630	<.0001
Zero factors (no factors other than age from list above)	-0.034	0.9746	1.281	0.0040
Full code	-1.691	<.0001	0.183	0.6154
Interaction terms between clinical variables				
Acute renal failure x chronic renal insufficiency	-0.615	<.0001	-0.382	0.0130
Acute renal failure x systolic blood pressure ≤ 90	-0.203	0.0119	-0.503	0.0001
GI bleed x heart rate ≥ 150 beats/min	-0.345	0.4228	0.479	0.4882
GI bleed x systolic blood pressure ≤ 90	0.126	0.3542	-0.575	0.0047
GI bleed x cirrhosis	-0.389	0.0834	0.064	0.8435
CPR before admission x mechanical ventilation	0.281	0.1791	0.299	0.3617
CPR before admission x coma/deep stupor	0.259	0.1249	-0.440	0.1486
Coma/deep stupor x mechanical ventilation	-0.545	0.0023	-1.030	0.0009
Cerebrovascular incident x coma/deep stupor	0.209	0.3219	-0.201	0.5927
Cerebrovascular incident x intracranial mass effect	0.784	0.0002	0.278	0.3328
Intracranial mass effect x coma/deep stupor	0.950	0.0003	-1.731	0.0001
Cardiac dysrhythmia x heart rate ≥ 150 beats/min	-0.535	0.0113	-0.626	0.0602
Interaction terms between age and other clinical variables				
Age x coma/deep stupor	0.002	0.5944	-0.012	0.0908
Age x heart rate ≥ 150 beats/min	-0.015	0.0048	-0.010	0.2251
Age x systolic blood pressure ≤ 90	0.000	0.8927	-0.007	0.0337
Age x chronic renal insufficiency	-0.005	0.1427	0.002	0.6226
Age x cirrhosis	-0.009	0.2827	-0.013	0.2888
Age x metastatic neoplasm	-0.023	<.0001	-0.008	0.1678
Age x acute renal failure	-0.010	0.0001	-0.014	0.0001
Age x cardiac dysrhythmia	0.003	0.2390	-0.001	0.7962
Age x cerebrovascular incident	-0.017	0.0004	-0.019	0.0021
Age x GI bleed	-0.009	0.0593	0.013	0.0163
Age x intracranial mass effect	0.002	0.7830	-0.001	0.8684

Age x CPR before admission	-0.013	0.0010	-0.004	0.5739
Age x mechanical ventilation	-0.004	0.0761	-0.005	0.0734
Age x medical or unscheduled surgical admit	-0.018	0.0006	-0.012	0.0122
Age x zero factors	-0.005	0.7081	-0.017	0.0107
Age x full code	0.011	0.0005	0.007	0.1300

$ICOM_{mort}$ Model Performance

Discrimination was assessed by using the area under the receiver operating characteristic curve (AUC). The minimum AUC that was considered reasonable discrimination was 0.80.[37] Our model demonstrated adequate discrimination on the validation sample, with an AUC of 0.820 (Table 3).

Calibration was assessed using the Hosmer-Lemeshow goodness-of-fit tests and calibration curves. We performed both the Hosmer-Lemeshow C test and H test. Analyses using the C test divide patients into deciles (i.e. equal number of patients) in ascending order of death. The range of predicted risk of mortality within each decile is determined by the patients in that decile. The H test forms 10 groups based on fixed, equal deciles of risk (i.e. 0.0-0.09%, 0.1%-0.19%, etc.) with variable numbers of patients in each group. The difference between the observed and expected mortality for each strata is summarized by the Pearson chi-square statistic. The statistics are summed over the ten deciles and are compared to the chi-square distribution. The degrees of freedom equal N-2, where N= number of groups, when used on an estimation dataset. However, when used on an application dataset, one in which the coefficients used are not recalculated using the dataset being analyzed, typically the degrees of freedom are the same as the number of groups (10 degrees of freedom.)[37]

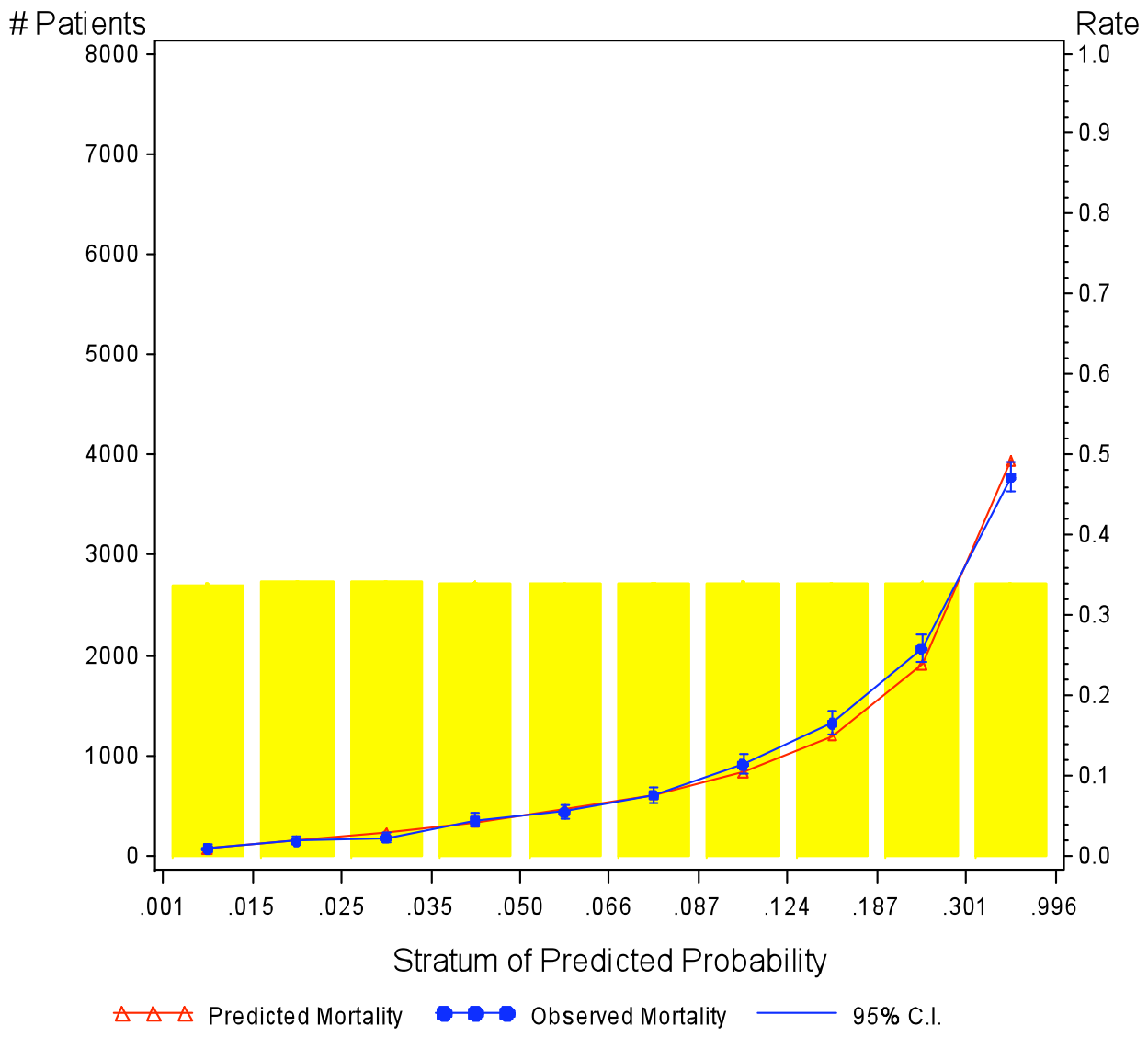
Given the sensitivity of the Hosmer-Lemeshow statistic to sample size,[37,38] after recalibration of the coefficients using logistic regression, calibration was reassessed using 11 random samples of 5,000 patients (Table 3) taken from the validation sample.[38] Nine of the 11 randomly selected samples of 5,000 patients showed non-significant H-L statistics. Calibration was also assessed using the adjunct measure of a graph plotting observed vs. predicted mortality. This plot is depicted in Figure 1.

Table 3. Performance of the Re-estimated $ICOM_{mort}$ Model on the Validation Sample

AUC (95% CI)	**Median H-L statistic***	
	C-test (p-value)	**H-test (p-value)**
0.820 (0.813-0.828)	12.04 (0.28)	16.85 (0.08)

*H-L statistics calculated on 11 random samples of 5,000 patients from the validation sample.

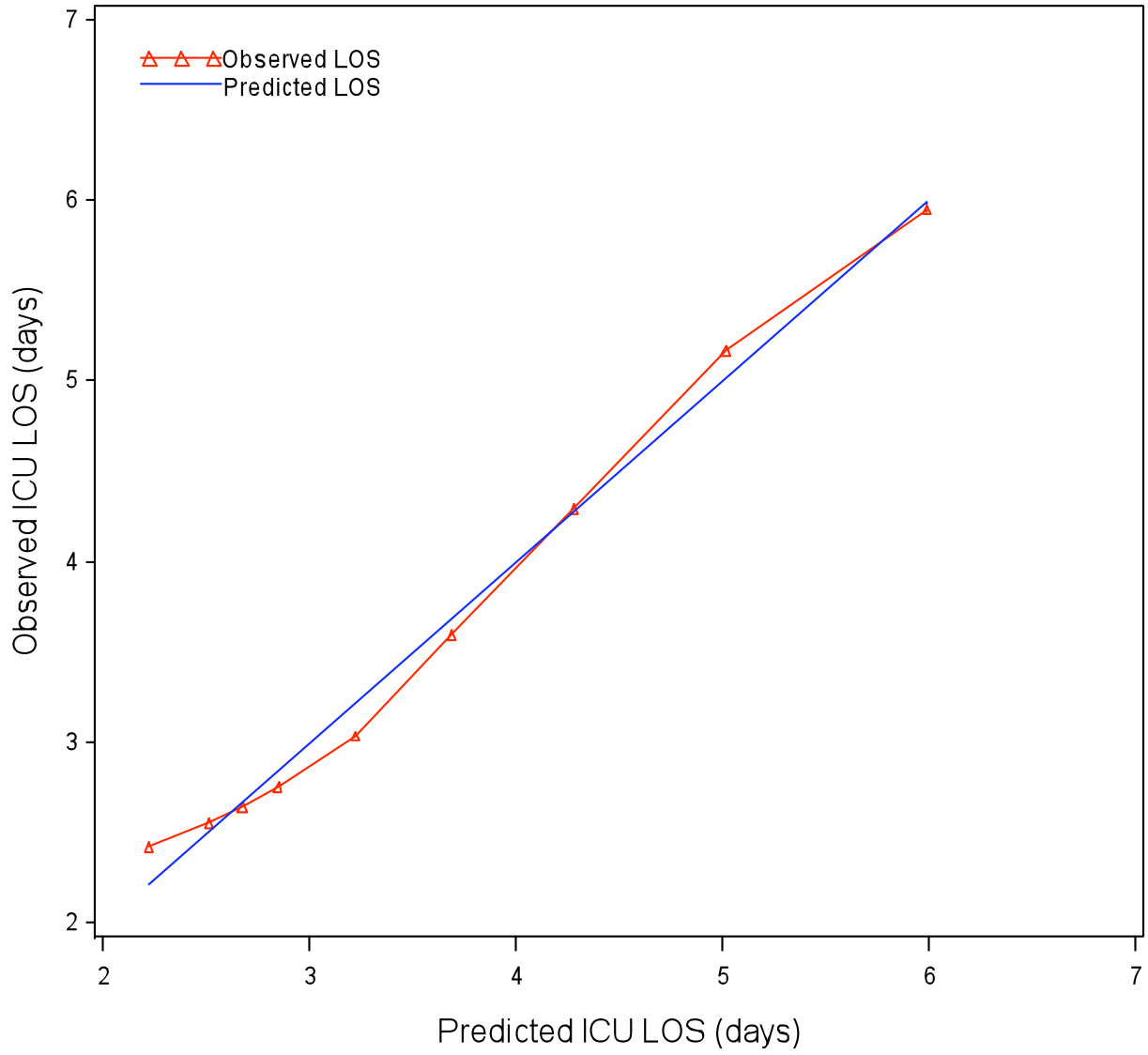
Figure 1. Calibration Curve for ICU Outcomes for Mortality ($ICOM_{mort}$)



$ICOM_{LOS}$ Model Performance

The $ICOM_{LOS}$ model was estimated using linear regression, with the randomly selected 60% estimation sample. The model showed moderate predictive power with adjusted R-square of 0.082. The calibration of the model was assessed within the 40% validation sample. Within deciles of predicted LOS, the following plot (Figure 2) compares the predicted versus observed mean LOS.

Figure 2. Calibration Curve for ICU Outcomes for Length-of-Stay ($ICOM_{LOS}$)



Because the patients used in evaluating LOS included some patients who died (approximately 12%), some elements of the model predicting LOS reflect the fact that death is associated with shorter than expected lengths of stay. This is reflected in the fact that the signs on some of the coefficients of some variables in $ICOM_{LOS}$ are the opposite of those in $ICOM_{mort}$ (that is, some variables with negative coefficients in $ICOM_{LOS}$ have positive coefficients in $ICOM_{mort}$ and vice versa), as seen in Table 2.

RECOMMENDATIONS FOR IMPLEMENTATION

It is now widely understood that any risk-adjusted ICU outcomes model needs to be re-calibrated (or even re-estimated) when applied to any new population,[20-24] and the $ICOM_{mort}$ and $ICOM_{LOS}$ are not exceptions to this rule. Therefore, although our best current models would involve using only the statistically significant variables (including interaction terms) from Table 2 for $ICOM_{mort}$ and $ICOM_{LOS}$, these models are not likely to be optimal for long.

Rather, we recommend that individuals or organizations wishing to assess risk-adjusted ICU outcomes collect all the variables required for these models and then recalibrate (or re-estimate) the models to fit the population whose outcomes are being evaluated. As above, we believe it is best to perform these tasks first on a randomly selected model development subsample of the overall population and test discrimination and calibration in a model validation sample. Of note, since the Hosmer-Lemeshow statistic is very sensitive to sample size, we recommend assessing calibration among multiple sets of 5,000 randomly selected members of the validation subsample.[38]

SUMMARY

The new $ICOM_{mort}$ and $ICOM_{LOS}$ models both demonstrate adequate performance as measured by discrimination and calibration. With the lowest data burden of any existing ICU outcomes models, $ICOM_{mort}$ and $ICOM_{LOS}$ provide a standardized means by which hospital and ICU performance can be compared between institutions at a reasonable cost. Prior to application to any new population, however, we recommend re-estimation of coefficients on the local sample in addition to confirmation of relevant interactions.

1. Angus DC, Linde-Zwirble WT, Sirio CA, et al. The effect of managed care on ICU length of stay: implications for medicare. *JAMA.* Oct 2 1996;276(13):1075-1082.

2. Wu AW, Pronovost P, Morlock L. ICU incident reporting systems. *J Crit Care.* Jun 2002;17(2):86-94.

3. Young MP, Birkmeyer JD. Potential reduction in mortality rates using an intensivist model to manage intensive care units. *Eff Clin Pract.* Nov-Dec 2000;3(6):284-289.

4. Cullen DJ, Sweitzer BJ, Bates DW, Burdick E, Edmondson A, Leape LL. Preventable adverse drug events in hospitalized patients: a comparative study of intensive care and general care units. *Crit Care Med.* Aug 1997;25(8):1289-1297.

5. Andrews LB, Stocking C, Krizek T, et al. An alternative strategy for studying adverse events in medical care. *Lancet.* Feb 1 1997;349(9048):309-313.

6. Giraud T, Dhainaut JF, Vaxelaire JF, et al. Iatrogenic complications in adult intensive care units: a prospective two-center study. *Crit Care Med.* Jan 1993;21(1):40-51.

7. Pronovost P, Wu AW, Dorman T, Morlock L. Building safety into ICU care. *J Crit Care.* Jun 2002;17(2):78-85.

8. Halpern NA, Pastores SM. Critical care medicine in the United States 2000-2005: an analysis of bed numbers, occupancy rates, payer mix, and costs. *Crit Care Med.* Jan 2010;38(1):65-71.

9. Rapoport J, Teres D, Lemeshow S, Avrunin JS, Haber R. Explaining variability of cost using a severity-of-illness measure for ICU patients. *Med Care.* Apr 1990;28(4):338-348.

10. Rapoport J, Teres D, Lemeshow S, Gehlbach S. A method for assessing the clinical performance and cost-effectiveness of intensive care units: a multicenter inception cohort study. *Crit Care Med.* Sep 1994;22(9):1385-1391.

11. Gunning K, Rowan K. ABC of intensive care: outcome data and scoring systems. *BMJ.* Jul 24 1999;319(7204):241-244.

12. Shortell SM, Zimmerman JE, Gillies RR, et al. Continuously improving patient care: practical lessons and an assessment tool from the National ICU Study. *QRB Qual Rev Bull.* May 1992;18(5):150-155.

13. Kuzniewicz MW, Vasilevskis EE, Lane R, et al. Variation in ICU risk-adjusted mortality: impact of methods of assessment and potential confounders. *Chest.* Jun 2008;133(6):1319-1327.

14. Rothen HU, Stricker K, Einfalt J, et al. Variability in outcome and resource use in intensive care units. *Intensive Care Med.* Aug 2007;33(8):1329-1336.
15. Knaus WA, Wagner DP, Zimmerman JE, Draper EA. Variations in mortality and length of stay in intensive care units. *Ann Intern Med.* May 15 1993;118(10):753-761.
16. Render ML, Kim HM, Deddens J, et al. Variation in outcomes in Veterans Affairs intensive care units with a computerized severity measure. *Crit Care Med.* May 2005;33(5):930-939.
17. Vasilevskis EE, Kuzniewicz MW, Cason BA, et al. Mortality probability model III and simplified acute physiology score II: assessing their value in predicting length of stay and comparison to APACHE IV. *Chest.* Jul 2009;136(1):89-101.
18. Rosenthal GE, Harper DL, Quinn LM, Cooper GS. Severity-adjusted mortality and length of stay in teaching and nonteaching hospitals. Results of a regional study. *JAMA.* Aug 13 1997;278(6):485-490.
19. Woods AW, MacKirdy FN, Livingston BM, Norrie J, Howie JC. Evaluation of predicted and actual length of stay in 22 Scottish intensive care units using the APACHE III system. Acute Physiology and Chronic Health Evaluation. *Anaesthesia.* Nov 2000;55(11):1058-1065.
20. Glance LG, Osler TM, Dick AW. Identifying quality outliers in a large, multiple-institution database by using customized versions of the Simplified Acute Physiology Score II and the Mortality Probability Model II0. *Crit Care Med.* Sep 2002;30(9):1995-2002.
21. Markgraf R, Deutschinoff G, Pientka L, Scholten T, Lorenz C. Performance of the score systems Acute Physiology and Chronic Health Evaluation II and III at an interdisciplinary intensive care unit, after customization. *Crit Care.* 2001;5(1):31-36.
22. 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.* Dec 1996;24(12):1968-1973.
23. Rivera-Fernandez R, Vazquez-Mata G, Bravo M, et al. The Apache III prognostic system: customized mortality predictions for Spanish ICU patients. *Intensive Care Med.* Jun 1998;24(6):574-581.
24. Zhu BP, Lemeshow S, Hosmer DW, Klar J, Avrunin J, Teres D. Factors affecting the performance of the models in the Mortality Probability Model II system and strategies of customization: a simulation study. *Crit Care Med.* Jan 1996;24(1):57-63.
25. 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.* Mar 2007;35(3):827-835.
26. Lemeshow S, Teres D, Klar J, Avrunin JS, Gehlbach SH, Rapoport J. Mortality Probability Models (MPM II) based on an international cohort of intensive care unit patients. *JAMA.* Nov 24 1993;270(20):2478-2486.
27. Knaus WA, Wagner DP, Draper EA, et al. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest.* Dec 1991;100(6):1619-1636.
28. 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.* May 2006;34(5):1297-1310.
29. Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *JAMA.* Dec 22-29 1993;270(24):2957-2963.
30. 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.* Oct 2005;31(10):1345-1355.
31. Galeiras R, Lorente JA, Pertega S, et al. A model for predicting mortality among critically ill burn victims. *Burns.* Mar 2009;35(2):201-209.

32. Boyd CR, Tolson MA, Copes WS. Evaluating trauma care: the TRISS method. Trauma Score and the Injury Severity Score. *J Trauma.* Apr 1987;27(4):370-378.
33. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). *Eur J Cardiothorac Surg.* Jul 1999;16(1):9-13.
34. Tunnell RD, Millar BW, Smith GB. The effect of lead time bias on severity of illness scoring, mortality prediction and standardised mortality ratio in intensive care--a pilot study. *Anaesthesia.* Nov 1998;53(11):1045-1053.
35. Combes A, Luyt CE, Trouillet JL, Chastre J, Gibert C. Adverse effect on a referral intensive care unit's performance of accepting patients transferred from another intensive care unit. *Crit Care Med.* Apr 2005;33(4):705-710.
36. Rosenberg AL, Hofer TP, Strachan C, Watts CM, Hayward RA. Accepting critically ill transfer patients: adverse effect on a referral center's outcome and benchmark measures. *Ann Intern Med.* Jun 3 2003;138(11):882-890.
37. Mourouga P, Goldfrad C, Rowan KM. Does it fit? Is it good? Assessment of scoring systems. *Current Opinion in Critical Care* June 2000;6(3):176-180.
38. Kramer AA, Zimmerman JE. Assessing the calibration of mortality benchmarks in critical care: The Hosmer-Lemeshow test revisited. *Crit Care Med.* Sep 2007;35(9):2052-2056.