IV. RISK MODEL FOR ADJUSTING HOSPITAL MORTALITY RATES, 2000-2002

Patients treated at different hospitals often vary in the severity of their pre-operative clinical condition. To fairly compare outcomes at different hospitals, it is necessary to adjust for differences in the case mix of patients across hospitals. CCMRP "levels the playing field" by accounting for the pre-operative condition of each patient. Hospitals that routinely handle complex cases (e.g., sicker prior to surgery) get a larger risk-adjustment weighting in the risk model, while hospitals that handle less complex cases get a smaller weighting.

CCMRP used a multivariable logistic regression model to determine the relationship between each of the demographic and pre-operative risk variables and the likelihood of in-hospital mortality. Multivariable logistic regression models relate the probability of death to the explanatory factor (e.g., *Patient Age, Creatinine Level, Type of Arrhythmia*), while controlling for all other explanatory factors in the model.

In model development, the three-year dataset was divided into two parts: Data for 2000 and 2001 were used as a “training set” to develop the model, and data for 2002 were used as a “test set” to validate the model. After a final model was chosen and tested, the coefficients were re-estimated using the entire three-year dataset.

Table 3 presents the final model based on the 2000-2002 dataset. Although the risk adjustment model is based on data from 83 hospitals, a risk-adjusted mortality rate is reported for only 77 hospitals: Six hospitals provided data but did not want their results published.

The final risk model included all variables used in the 1999 CCMRP risk model with the exception of angina. The 1999 audit and subsequent analyses revealed uneven coding of that risk factor across hospitals, so it was dropped. In addition, *Cardiogenic Shock* (Yes/No), *NYHA* (Class IV), and *BMI* were added to the model.

Table 3: Logistic Regression Risk Model for Inpatient Mortality, 2000-2002

Explanatory Factor		Coefficient	Standard Error	p-value	Significance	OR
Intercept		-9.74	0.32	0.00	***	
Age (Years)		0.06	0.00	0.00	***	1.06
Gender	Male ^			Reference Group		
	Female	0.37	0.06	0.00	***	1.44
Race	White ^			Reference Group		
	Non-White	0.16	0.06	0.01	**	1.18
Body Mass Index	18.5-39.9 ^			Reference Group		
	< 18.5	1.07	0.16	0.00	***	2.91
	≥ 40.0	0.42	0.15	0.01	**	1.52
Acuity	Elective ^			Reference Group		
	Urgent	0.26	0.07	0.00	***	1.29
	Emergent	1.07	0.11	0.00	***	2.91
	Salvage	2.59	0.22	0.00	***	13.32
Creatinine (mg/dl)		1.15	0.12	0.00	***	3.16
Hypertension		0.06	0.07	0.37		1.06
Dialysis		0.47	0.14	0.00	***	1.59
Peripheral Vascular Disease		0.22	0.07	0.00	***	1.25
Cerebrovascular Disease		0.04	0.07	0.58		1.04
Cardiogenic Shock		0.91	0.11	0.00	***	2.49
Congestive Heart Failure		0.26	0.07	0.00	***	1.30
Diabetes		0.09	0.06	0.12		1.10
Arrhythmia Type	None ^			Reference Group		
	Afib/Flutter	0.39	0.09	0.00	***	1.48
	Heart Block	0.38	0.12	0.00	**	1.47
	Sustained VT/VF	0.50	0.12	0.00	***	1.65
Chronic Lung Disease	None ^			Reference Group		
	Mild	0.28	0.09	0.00	**	1.33
	Moderate	0.34	0.10	0.00	**	1.41
	Severe	0.92	0.12	0.00	***	2.52
Myocardial Infarction	None ^			Reference Group		
	21 or more days ago	0.14	0.08	0.07		1.15
	8 to 20 days ago	-0.14	0.14	0.31		0.87
	1-7 days ago	0.40	0.07	0.00	***	1.49
	Within 24 Hours	0.44	0.12	0.00	***	1.55
NYHA Class IV		0.31	0.07	0.00	***	1.37
Left Main Disease > 50%		0.08	0.06	0.15		1.09
Prior Operations on Pump	None ^			Reference Group		
	1	0.99	0.08	0.00	***	2.69
	2 or more	1.20	0.22	0.00	***	3.32
PTCA	None ^			Reference Group		
	≤ 6 Hours	-0.07	0.07	0.36		0.94
	> 6 Hours	0.01	0.21	0.94		1.02
Ejection Fraction (%)		-0.01	0.00	0.00	***	0.99
Number of Diseased Vessels	One ^			Reference Group		
	Two	-0.02	0.16	0.92		0.99
	Three or More	0.09	0.15	0.54		1.09
Mitral Insufficiency	None ^			Reference Group		
	Trivial	0.03	0.10	0.74		1.04
	Mild	0.06	0.09	0.49		1.06
	Moderate	0.08	0.12	0.51		1.08
	Severe	0.29	0.26	0.28		1.33

Note: ^ refers to the category used to replace missing data for a variable.

Guide for Interpreting the Risk Model

Coefficient: The coefficient for each explanatory factor represents the effect that factor has on a patient's likelihood of dying (in the hospital) following bypass surgery. If the value is positive, it means that the characteristic is associated with an increased risk of death compared to not having the characteristic, while controlling for the effect of all other factors. If the coefficient is negative, having that characteristic is associated with a lower risk of death compared to not having it. The larger the value (whether positive or negative), the greater the effect or weight this characteristic has on the risk of dying. For example, the coefficient for *Congestive Heart Failure (CHF)* in the model is 0.26 and statistically significant. This value is positive, so it indicates that CABG patients with congestive heart failure are at an increased risk of dying compared to patients who do not have the condition.

Standard Error: The standard error is the standard deviation of the sampling distribution of an estimate. It measures the statistical reliability of that estimate.

p–value: The p-value is a measure of the statistical significance of the coefficient compared to the reference category. Commonly, p-values of less than 0.05 are considered statistically significant. The smaller the p-value, the more likely the effect of a factor is real, rather than due to chance.

Significance: When the p-value of a coefficient is less than 0.05, it is deemed statistically significant at the 0.05 level and is denoted with one star (*) in the significance column. Two stars (**) indicate statistical significance at the 0.01 level and three stars (***) indicate statistical significance at the 0.001 level. All statistical tests are two-tailed tests.

Odds Ratio: An odds ratio is another way of characterizing the impact of each factor on in-hospital mortality. Mathematically, the odds ratio is the antilogarithm of the coefficient value. The larger the odds ratio, the greater the impact that characteristic has on the risk of dying. An odds ratio close to 1.0 means the effect of the factor is close to neutral. For example, the odds ratio for *CHF* in the model is 1.30. This means that for patients with *CHF*, the odds of dying in-hospital are about 30% higher compared to patients without *CHF*, assuming all other risk factors are the same.

Discrimination

Models that distinguish well between patients who die and those who survive are said to have good discrimination. A commonly used measure of discrimination is the c-index (also known as the c-statistic or the area under the Receiver Operating Curve (ROC)). For all possible pairs of patients, where one dies and the other survives surgery, the c-index describes the proportion of pairs where the patient who died had a higher predicted risk of death than the patient who lived. The c-index ranges from 0 to 1, with higher values indicating better discrimination. For the 2000-2002 data model the c-index is 0.828. In comparison, c-indexes reported in other recently published studies of CABG mortality using logistic regression (including those from New Jersey, New York, Pennsylvania, and the Society of Thoracic Surgeons) range from about 0.78 to 0.82. As such, the CCMRP model appears to discriminate as well as, or better than, models from other programs that produce risk-adjusted outcomes data for isolated CABG surgery.

Calibration

Calibration refers to the ability of a model to match predicted and observed death rates across the entire spread of the data. A model in which the number of observed deaths aligns well with the number of deaths predicted by the model demonstrates good calibration. Good calibration is essential for reliable risk adjustment. A common measure of calibration is the Hosmer and Lemeshow χ^2-statistic, which compares observed and predicted outcomes over deciles of risk. The Hosmer-Lemeshow test statistic is 29.1 (df=8; p-value=0.00) for the 2000-2002 model (i.e., reject the null hypothesis that there is no difference between actual and predicted deaths). This result was not a major cause for concern; with such a large sample it is common to fail the Hosmer-Lemeshow test.

The next step was to inspect the difference between the actual number of deaths and the predicted number of deaths (derived from the risk model) in each of 10 risk groups. The 10 groups are created by sorting all observations by the predicted risk of death and then dividing the sorted observations into deciles of approximately equal size. Table 4 shows the calibration of the 2000-2002 risk-adjustment model.

Table 4: Calibration of 2000-2002 Model (n=57,388)

Group	N	Minimum Predicted Risk	Maximum Predicted Risk	Actual Deaths	Expected Deaths	Difference
1	5,740	0.001	0.004	3	15.2	(12.2)
2	5,739	0.004	0.005	24	26.1	(2.1)
3	5,739	0.005	0.007	35	36.7	(1.7)
4	5,739	0.007	0.010	31	48.8	(17.8)
5	5,741	0.010	0.013	46	63.5	(17.5)
6	5,739	0.013	0.016	77	82.6	(5.6)
7	5,740	0.016	0.022	126	110.3	15.7
8	5,740	0.022	0.033	167	155.4	11.6
9	5,739	0.033	0.057	277	245.8	31.2
10	5,732	0.057	0.962	769	770.6	(1.6)

The first row of Table 4 shows the decile of patients at lowest risk of in-hospital death in the CCMRP model (e.g., the 5,740 patients whose predicted risk of dying ranged from 0.001 to 0.004). Among the first decile, three patients died, but the model predicted death for 15 of the patients. Assuming a Poisson distribution for a binary outcome with mean 0.0026 (15.2 ÷ 5,740), the predicted range of deaths for the first decile is eight to 23. The observed number of three deaths falls below the expected range. However, 49% of actual deaths occurred in the 10^{th} decile, the highest risk decile of patients, where 769 patients died compared to 771 deaths predicted by the model. The predicted range for the tenth decile is 716 to 825 deaths. The number of observed deaths is very nearly the exact number predicted by the model. Overall, in seven of the ten groups, the number of actual deaths is within the range of expected deaths. Although for groups 1, 4, and 5, the number of observed deaths is below the number of expected deaths, the model calibration shows that the risk model has accurately predicted the number of expected deaths, especially for patients with the highest risk of dying.

Key Findings Regarding the Risk Model

- Although some of the risk model variables are not statistically significant (as determined by a p-value of <0.05), all significant coefficients appeared with the expected sign from a clinical standpoint.

- *Age, Acuity* (e.g., urgency of the operation), *Cardiogenic Shock*, *Dialysis, Ejection Fraction, Creatinine,* and the number of *Prior Operations on Pump* were the most important risk model variables.

- Patients who were extremely underweight (*BMI* < 18.5) had a higher risk of dying in-hospital (OR 2.91) than those in the reference group (*BMI* 18.5-39.9). Patients who were extremely overweight (*BMI* > 40.0) were also at increased risk of death (OR 1.52) but not to the extent that the very underweight were. A very low *BMI* may be a proxy for frailty or indicate a wasting comorbid condition not captured by other risk model variables.

- Even after controlling for all other variables, *Gender* had a statistically significant effect, with males having about one-third lower mortality. This gender effect has weakened when compared to the 1997-1999 model, perhaps because of the inclusion of *BMI* in the current model. The literature suggests that *gender* may be a proxy for body size and/or coronary artery size (diameter) and smaller coronary arteries in women may be more prone to thrombosis or restenosis.

- Of the acute comorbidities collected, *Cardiogenic Shock* had the largest effect (OR 2.49). Of the chronic comorbid conditions, severe *Chronic Lung Disease* has the strongest association with inpatient mortality (OR 2.52).

- Patients with *Left Main Disease* > 50% did not appear to be at increased risk (OR 1.09, not significant) of inpatient death. However, when *Left Main Stenosis* was collected as a continuous measure (see 1997-1999 model), patients with Stenoses > 70% were about 50% more likely to die.

- When compared to the 1997-1999 risk model, six variables in the prior model were no longer significant, and two variables not significant in the prior model were significant in the current model. Of most concern from a clinical perspective, there was no increased risk of mortality from *PTCA* <= 6 hours (OR .94, not significant) though a variable definition change might be responsible for this result. Severe *Mitral Insufficiency* was no longer a significant risk factor, which may also go against clinical reasoning. On the other hand, pre-operative *Dialysis* behaved as expected, putting patients at additional risk (OR 1.59); previously, it did not.

- *Creatinine* was entered into the current risk model as a piecewise linear function, so its odds ratio (3.16) is not comparable to prior CCMRP reports where it was entered as a continuous measure.