CHAPTER EIGHT:
PROCEDURE FOR DEVELOPING RISK ADJUSTMENT MODELS

This chapter describes the analytical and statistical methods used to develop risk-adjustment models for the California Hospital Outcomes Project. The discussion here is technical and requires some knowledge of statistics and research design. Volume One contains a less technical discussion.

The development of risk-adjustment models followed a series of steps beginning with identification of the outcome of interest (30-day in-hospital mortality for AMI) and potential risk factors. Detailed definitions of the outcome and risk factors are presented in Chapters Five and Seven, respectively.

Each of the nine steps in developing risk-adjustment models is described in detail below. These steps may be briefly summarized as follows:

1. The lists of potential clinical risk factors were reviewed to identify two key subgroups: (a) particularly important factors that should be forced in to all risk-adjustment models, and (b) factors that might represent either comorbidities or complications, and therefore should be used only in selected models.

2. Univariate and bivariate analyses were used to identify and eliminate low-frequency risk factors, eliminate other risk factors that do not affect or have counterintuitive associations with mortality, and summarize multi-level clinical risk factors as either ordinal predictors or multiple dummy variables.

3. Descriptive analyses were performed to select the best method for modeling the effects of age and other non-clinical risk factors.

4. Each sample was split into two separate samples for estimating and validating risk-adjustment models.

5. Clinical risk factors were selected for the primary risk-adjusted model (labeled Model A), using a set of ten random subsamples to choose only risk factors with both robust and statistically significant parameter estimates.

6. Two-way interactions were selected for the primary risk-adjustment model, using a variety of variable selection procedures.

7. Risk-adjusted models were internally validated and refined by applying models developed using the estimation sample to the corresponding validation sample.

8. Additional non-clinical and clinical risk factors were selected for Model B to assess whether hospital outcome statistics were sensitive to including these additional variables in the analysis.

9. Each risk model was re-estimated after combining the estimation and validation samples, to generate more reliable parameter estimates.

STEP 1: REVIEW OF POTENTIAL CLINICAL RISK FACTORS

The potential clinical risk factors listed in Chapter Seven were reviewed to identify two important subsets. These subsets were analyzed in somewhat different ways from the remaining risk factors, as described below.

1.1 **Particularly important clinical risk factors were identified through review of prior literature and discussions with clinical advisors.**

These factors were forced into all risk-adjustment models, to maximize their face validity to clinicians and health services researchers. Risk-adjustment models without these variables would have been vulnerable to unidentified interactions. The stepwise methods later used to select variables might otherwise have eliminated crucial predictors. However, it was important to be very selective in choosing which variables to force into risk models, because unnecessary and irrelevant variables can overburden a model. The risk factors forced into the AMI risk-adjustment models were female sex, infarct site (e.g., anterior wall, inferior wall, subendocardial, other or unspecified) and prior coronary bypass surgery.

1.2 **Clinical risk factors that might represent complications of care were identified through review of prior literature and discussions with clinical advisors.**

California patient discharge abstracts do not distinguish between comorbidities which typically are present at admission, and complications that develop during an inpatient stay. In the absence of specific information on the abstract, judgments were made as to

whether various conditions were more likely to have been present at admission or to have developed later.

This distinction was important because two risk-adjustment models were developed to predict AMI mortality. Model A is a conservative model that includes fewer risk factors; Model B is a more comprehensive model that includes important but potentially biased risk factors. Conditions that almost certainly were present at admission were candidates for inclusion in both Model A and Model B. Conditions likely to have developed later were candidates only for Model B. Model B thereby gives hospitals the benefit of the doubt related to associated conditions that have unclear timing.

AMI risk factors considered for Model B but not for Model A were shock, hypotension, pulmonary edema, complete atrioventricular block, pleural effusion, urinary tract infection, syncope, acidosis, alkalosis, sepsis, paroxysmal ventricular tachycardia, hyponatremia or hyposmolality, hypernatremia or hyperosmolality, gastrointestinal hemorrhage, pneumonia, aspiration pneumonitis, and unstable angina.

Diagnoses from prior hospitalizations were available for 8.1% of AMI cases. Several risk factors were considered for Model A only if they appeared on the discharge abstract from a prior hospitalization, but were considered for Model B no matter which discharge abstract listed the diagnosis. These AMI risk factors included epilepsy, bundle branch block, atrial fibrillation, cerebrovascular disease, skin ulcer, coagulopathy, supraventricular tachycardia, premature beats, arterial emboli or thromboses, acute renal failure, acute peptic ulcer, and other atrioventricular block.

STEP 2: PRELIMINARY ANALYSES OF CLINICAL RISK FACTORS

These analyses were designed to describe the frequency distributions of all clinical risk factors, detect covariates and covariate patterns with very few observations, evaluate the unadjusted bivariate association between each covariate and death, and summarize multi-level clinical risk factors in a manner appropriate for regression modelling.

2.1 **The frequency distribution of each clinical risk factor was determined and very low-frequency risk factors were eliminated or aggregated.**

Binary risk factors present in less than 1% of all cases were examined carefully. Whenever possible, these risk factors were combined with

physiologically related risk factors that were similarly associated with death. If aggregation along clinical lines was impractical, risk factors present in fewer than 20 patients who died were eliminated. Twenty was chosen as the cutoff because it corresponds to the minimum number (n = 6) needed to estimate effect sizes in risk models based on 30% bootstrap samples (see Step 5 for a detailed description of these bootstrap samples).

No AMI risk factors were eliminated because of low frequency among cases **without** prior hospitalizations. However, the following risk factors failed to qualify in the sample of cases **with** prior hospitalizations: chronic peptic ulcer, acute peptic ulcer, chronic liver disease, coagulopathy (from prior hospitalization). In **both** samples, high risk primary malignancy and secondary malignancy were aggregated into one risk factor that qualified for retention.

2.2 **Clinical risk factors not associated with mortality were identified and eliminated, to improve the efficiency of subsequent modeling.**

The unadjusted bivariate association between each clinical risk factor and death was summarized using relative risk estimates with 95% confidence limits and p-values derived from a continuity-adjusted chi-square distribution (with *k-1* degrees of freedom, where *k* equals the number of risk categories).

Risk factors that were not associated with death at a $p < 0.10$ level were eliminated from further consideration. This cutoff was selected to screen out risk factors least likely to contribute significantly to a multivariate model.

The following AMI risk factors were eliminated because they were not significantly related to mortality among cases **without** prior hospitalizations: collagen vascular disease, chronic obstructive pulmonary disease, psychosis, specified or unspecified anemia, cardiomegaly, urinary tract infection, low-risk primary malignancy, mitral valve disease, other valve disease, and personal history of malignancy. Among cases **with** prior hospitalizations, all of these risk factors except collagen vascular disease, low-risk primary malignancy, mitral valve disease, and other valve disease also failed to qualify. The following additional risk factors were eliminated because they were not associated significantly with mortality among cases **with** prior hospitalizations: complicated diabetes, coagulopathy, neurologic disorders, hypertensive heart failure, chronic pulmonary heart disease, other atrioventricular block, supraventricular tachycardia, premature beats, chronic glomerulonephritis,

osteoarthritis, prior pacemaker insertion, hyposmolality or hyponatremia, pleural effusion, gastrointestinal hemorrhage, and syncope. In addition, several risk factors based exclusively on prior hospitalizations were not related significantly to mortality (e.g. epilepsy, bundle branch block, premature beats, acute renal failure, other atrioventricular block).

2.3 Clinical risk factors that had counterintuitive associations with mortality were identified and eliminated, if biased coding appeared to be the most likely explanation.

The directions of all statistically significant associations between risk factors and mortality were examined. These findings were reviewed with the appropriate clinical advisory panel, after considering the literature summarized in Chapter Two. Risk factors that appeared to lower the risk of AMI death when previous literature and clinical experience suggested the opposite relationship, were eliminated from the analysis. Studies using reabstraction[1,2,3] or data linkage[4] have demonstrated substantial underreporting of several such conditions. Counterintuitive risk-outcome associations could be explained by selective underreporting among patients with poor outcomes.[5,6]

The following AMI risk factors were eliminated because they were counterintuitively associated with lower mortality among cases without prior hospitalizations: hyperlipidemia, obesity, gout or osteoarthritis, unstable angina, old AMI, other atrioventricular block, premature beats, asthma, chronic peptic ulcer disease, syncope, uncomplicated

---

[1]Fisher ES, Whaley FS, Krushat WM, Malenka DJ, Fleming C, Baron JA, et al. The accuracy of Medicare's hospital claims data: Progress has been made, but problems remain. *American Journal of Public Health* 1992; 82:243-248.

[2]Romano PS, Mark DH. Bias in the coding of hospital discharge data and its implications for quality assessment. *Medical Care* 1994; 32:81-90.

[3]Romano PS, Luft HS. Getting the most out of messy data: Problems and approaches for dealing with large secondary data sets. In Grady ML, Schwartz H, eds. *Medical Effectiveness Research Data Methods*. Rockville, MD: US Department of Health and Human Services; 1992. AHCPR Pub. No. 92-0056.

[4]Jollis JG, Ancukiewicz M, DeLong E, Pryor DB, Muhlbaier LH, Mark DB. Discordance of databases designed for claims payment versus clinical information systems: Implications for outcomes research. *Annals of Internal Medicine* 1993; 119:844-850.

[5]Jencks SF, Williams DK, Kay TL. Assessing hospital-associated deaths from discharge data: the role of length of stay and comorbidities. *JAMA* 1988; 260:2240-2246.

[6]Iezzoni LI, Foley SM, Daley J, Hughes J, Fisher ES, Heeren T. Comorbidities, complications and coding bias: Does the number of diagnosis codes matter in predicting in-hospital mortality? *JAMA* 1992; 267:2197-2203.

diabetes, and alcohol or drug use. Further analyses suggested that unstable angina patients may have had very small infarcts under inpatient observation; ICD-9-CM coding guidelines state that AMI may be coded as the principal diagnosis in this situation.[7] Among cases with prior hospitalizations, all of the same risk factors failed to qualify.

2.4 **Multi-level clinical risk factors were summarized as either ordinal predictors or multiple dummy (dichotomous) variables, as appropriate.**

Several clinical risk factors could be divided readily into two or more severity categories, based on the fourth or fifth digit of the ICD-9-CM code or the presence or absence of certain associated diagnoses. For example, diabetes may be classified as complicated if it is associated with ketoacidosis, coma, or end-organ disease (e.g., neuropathy, retinopathy, nephropathy).

To determine how to model the effect of multi-level clinical risk factors, the unadjusted association between each such factor and death was summarized using relative risk estimates with 95% confidence limits and p-values derived from a Mantel-Haenszel chi-square for trend. The Kruskal-Wallis test was used instead of analysis of variance when the equal variance assumption was not satisfied. If the relationship between a multi-level predictor and the risk of an adverse outcome was monotonic (and approximately linear on a logit scale), then the predictor was treated as an ordinal variable in regression models. Otherwise, multiple dummy (dichotomous) variables were created to capture the independent effect of each level. Two adjacent levels were combined into one dummy variable if they were associated with the same risk.

The AMI risk factors with multiple levels were diabetes and hypertension; neither displayed a monotonic relationship with the risk of death. Separate dummy variables were created, but only complicated diabetes and uncomplicated hypertension otherwise qualified for inclusion in the risk-adjustment models (as described above).

STEP 3: PRELIMINARY ANALYSES OF NON-CLINICAL RISK FACTORS

These analyses were designed to describe the distributions of all non-clinical risk factors, to evaluate the unadjusted association between each covariate

[7]Sequencing of angina and coronary heart disease. *Coding Clinic* 1990; 7(3):6-10.

and death, and to select the appropriate analytic specification of each non-clinical variable.

3.1 **The distribution of age (and other continuous predictors) and the associations between these predictors and mortality were evaluated.**

Smoothed scatter plots of the logit outcome (log[p/(1-p)]) as a function of **age** were used to determine the best-fitting form of the relationship between mortality and age. Age was categorized in increments of one to five years, so that each age group had a sufficient number of observations for analysis. Specific components of the age-mortality relationship, such as linear and quadratic terms, were tested using a likelihood ratio statistic.

This analysis led to a change in the specification of age in the study of AMI mortality. In 1993, five dummy variables were used to specify the relationship between age and mortality. In the current study, age was truncated at 100 years and specified as a linear predictor. Truncation was important to minimize the influence of patients erroneously reported as being over 100 years of age and to preserve linearity in the association with the logit risk of death. By treating age as a continuous variable instead of multiple dummy variables, it was easier to evaluate interactions involving other risk factors.

The same approach was applied to examine the relationship between the **month of discharge** (ordered sequentially from the beginning to the end of the study period) and mortality. The month of discharge did not appear to be related to the death rate after AMI.

3.2 **The distribution of categorical non-clinical variables and the associations between these variables and each outcome of interest were evaluated.**

Contingency tables were used to evaluate the relationship between each categorical demographic (e.g., gender, race) and hospitalization characteristic (e.g., expected principal source of payment, source of admission, type of admission, day of week of admission) variable and mortality. This made it possible to combine low-frequency categories that were conceptually similar or had similar death rates.

Race was aggregated into four categories: white, African-American, Hispanic, and other. The "other" category included Asian-Americans, Native Americans, and other groups.

Four categories of **expected payment source** were used: Medicare, MediCal, uninsured (including self-pay, no charge, and section 17000 indigent services), and insured (including Blue Cross/Blue Shield, insurance company, health maintenance organization, Worker's Compensation, Title V, and other government or non-government insurance). Although there were enough HMO cases to create a separate category, this was not done because HMO cases tend to be concentrated at certain hospitals. Adjusting for an HMO insurance effect would have made it difficult to evaluate the performance of these hospitals.

Source of admission was grouped into two categories: (1) routine or home health service, and (2) emergency room (ER), inpatient facility (skilled nursing, intermediate care, acute care), other facility, or other source. Transfers from inpatient facilities were excluded from the AMI analysis, for the reasons described in Chapter Three. Admissions from other facilities and other sources were combined with ER admissions because OSHPD's reabstracting study showed that 52% of these cases should have been reported as ER admissions, and because their risk of death was closer to that of ER admissions than to that of routine admissions.

Type of admission was grouped into two categories: elective or urgent versus emergent. This classification was chosen because AMI death rates were very similar between elective and urgent admissions.

3.3 One category of each demographic variable was designated as the reference group.

The most frequent category of each non-clinical variable was generally chosen as the reference group for regression modelling. Males were selected as the reference group in all models. In all models that included race, white was the reference group. In all AMI models that included source of payment, insurance other than Medicare and MediCal was the reference group. In all AMI models that used source of admission, routine or home health service was the reference group. Elective or urgent admissions were the reference group in models that used admission type.

STEP 4: DIVISION OF DATA INTO SEPARATE SAMPLES FOR ESTIMATION AND VALIDATION

The data set was split into an estimation sample and a validation sample, by randomly selecting 60% of the original cases (without replacement) for the

estimation sample and setting aside the remaining 40% for the validation sample. This procedure made it possible to develop risk-adjustment models on the estimation samples and then assess these models on separate validation samples. Such a test of model fit is more rigorous than one that uses the same sample for both estimation and validation. A 60%/40% split was chosen because a larger estimation sample is more likely to contain cases from sparse cells (rare risk factor combinations), and therefore may allow better assessment of interactions.

Sampling was stratified by outcome status (death) to ensure that the overall probability of the outcome was the same in both the estimation and validation samples.

STEP 5: SELECTION OF MAIN EFFECTS RISK FACTORS FOR MODEL A

As described in Step 1, two different models (A and B) were used to adjust for patient differences across hospitals. The demographic and clinical risk factors in Model A were almost certainly present when the patient entered the hospital and therefore reflect his or her health on admission. Model B contains all of the risk factors in Model A as well as others that may reflect either health on admission or quality of care.

The goal of Step 5 was to identify a single best set of "main effects" risk factors for Model A, using a procedure that would be robust in a variety of circumstances. To this end, subsamples of the estimation sample were randomly generated, and covariate selection procedures (described below) were completed for each subsample. The results of this process were reviewed to determine the best set of risk factors. This procedure minimized the risk of overfitting a risk-adjustment model to the peculiarities of a particular sample.

5.1 **Ten independent random subsamples were generated, without replacement and a sampling fraction of 50%, from the 60% estimation sample.**

Sampling without replacement means that the same case would not have been selected more than once for a single subsample. Sampling with replacement has the theoretical advantage of allowing a subsample to contain more cases with a rare risk factor than the population from which that sample was drawn.

5.2 **The best risk factor set for each subsample was determined.**

For each subsample, a multivariate regression model was fit using stepwise forward selection with the significance level tolerance set to

0.10, forcing in the important clinical risk factors identified in Step 1. Probability values to enter and remove variables were based on the likelihood ratio statistic in logistic models with dichotomous outcomes.

5.3 **The subsample results were combined to determine the final Model A risk factor set.**

All risk factors that were significant at $p < 0.10$ in five or more of the ten subsamples were retained in the construction of Model A. The following AMI risk factors were eliminated from the **"no prior hospitalization"** model because they were significant in fewer than five subsamples: hereditary, degenerative, or demyelinating disorders of the nervous system; hypertensive heart failure; dementia or Alzheimer's disease; malnutrition; chronic pulmonary heart disease; systemic atherosclerosis; chronic glomerulonephritis; and prior pacemaker insertion. The following AMI risk factors were eliminated from the **"one or more prior hospitalizations"** model for the same reason: low-risk primary malignancy; malnutrition; other valve disease or prior valve replacement; late cerebrovascular disease; dementia; collagen vascular disease; atrial fibrillation; arterial embolism or thrombosis; unspecified anemia; supraventricular tachycardia; cardiomegaly; and other cerebrovascular disease. The last six of these risk factors were coded as present in Model A only if noted in a prior record.

5.4 **The variables confirmed as robust predictors of adverse outcomes were tested in a stepwise regression model on the entire 60% sample.**

One limitation of the multiple subsample method described above is that when several predictors are highly colinear, stepwise models from different subsamples may include different predictors. The contribution of one variable may be fully explained by another variable or combination of variables that did not enter that particular model. Alternatively, competing variables may drop out of a model based on a small (bootstrap) sample, whereas they would stay in a model based on a larger sample. To address these concerns, all risk factors that met the five-sample bootstrap criterion were tested in a stepwise regression using the full 60% estimation sample, with a p-to-enter of $p < 0.01$.

This procedure eliminated no predictors from the analysis of AMI cases **without** prior hospitalizations. However, it resulted in dropping hypothyroidism and atherosclerosis from the analysis of AMI cases **with** prior hospitalizations.

STEP 6: SELECTION OF RISK FACTOR INTERACTIONS FOR MODEL A

The number of Model A risk factors was too large to consider all two-way interactions, let alone three-way and higher order interactions. The choice of approach in the analysis reflects the difficult balance between optimizing model performance and computational efficiency.

6.1 **Clinically plausible interactions involving important main effects were identified and tested.**

This approach was based on the premise that only interactions involving the most important main effects would contribute substantially to risk-adjustment models. In the AMI analyses, only interactions involving age or infarct site (e.g., anterior wall, inferior wall, other or unspecified) were tested.

All of these interactions were tested using the ten randomly generated subsamples described above. For each subsample, a multivariate regression model was fit using stepwise forward selection with the significance level tolerance set to 0.10, forcing in all of the important main effects identified in Steps 1 and 5. Probability values to enter and remove variables were based on the likelihood ratio statistic in logistic models with dichotomous outcomes and on the F statistic in linear models with continuous outcomes. All interactions that were significant at $p < 0.10$ in five or more of the ten subsamples were retained in the construction of Model A.

All risk factors that met the five-sample bootstrap criterion then were tested in a stepwise regression using the full 60% estimation sample, with a p-to-enter of $p < 0.01$. This procedure eliminated several interactions from each analysis.

STEP 7: INTERNAL VALIDATION AND REFINEMENT OF RISK ADJUSTMENT MODELS

To internally validate the final covariate set in each risk-adjustment model, the parameter estimates from the 60% estimation sample were compared to the corresponding parameter estimates derived by fitting the same model to the 40% validation sample. Model specification was considered adequate if a parameter estimate from the 60% estimation sample fell within the corresponding 95% confidence intervals from the 40% validation sample.

Nearly all main effects parameter estimates based on the 60% estimation samples were within the corresponding 95% confidence intervals based on

the 40% validation samples. Lack of overlap in parameter estimates was noted for a larger number of interaction variables. Some of these variables were statistically significant in the estimation sample, but not in the validation sample. A few even had opposite signs in the two samples (e.g., an adverse effect in the estimation sample and a protective effect in the validation sample). All of these variables were examined individually.

The calibration of each risk-adjustment model was assessed with the Hosmer Lemeshow goodness of fit test (further described in Chapter Ten). Specifically, the risk-adjustment model developed on the 60% estimation sample was applied to the 40% validation sample. This was important to ascertain whether the model would fit as well in an independent sample as in the sample used for estimation. This comparison generally demonstrated similar goodness-of-fit across risk strata in the two samples, but some calibration problems were identified and addressed.

As a result of these procedures, two interaction terms were removed from the risk-adjustment model for AMI patients **with** prior hospitalizations. These terms had non-overlapping parameter estimates with opposite signs in the estimation and validation samples. Although the models for AMI patients **without** prior hospitalizations also had several non-overlapping parameter estimates, these variables were not removed because they had strong adverse effects in both samples.

STEP 8: SELECTION OF ADDITIONAL RISK FACTORS FOR MODEL B

To select the additional risk factors for Model B, a procedure was applied similar to that used to select Model A risk factors in Step 5. Ten independent random subsamples were generated, without replacement and a sampling fraction of 50%, from the 60% analytic sample.

Two sets of variables were considered for Model B that were not considered for Model A: clinical characteristics that could represent either comorbidities or complications, and non-clinical characteristics that could be associated with mortality but could also represent confounded or unreliable measures. The clinical characteristics were identified in Step 1.2. The non-clinical characteristics included race, expected principal source of payment, source of admission, and type of admission.

Race and expected payment source were not considered in Model A because they might be associated with differences in the quality of care. They were considered in Model B because they might reflect differences in the severity of illness at admission, perhaps due to delays in seeking care or inadequate outpatient care. Type of admission was not considered in Model A because OSHPD's 1988 reabstracting study noted a 36% error rate for this variable. It

was considered in Model B because physicians may label patients as "emergency" or "urgent" based on clinical features that otherwise would not be captured in risk-adjustment models. Source of admission was not considered in Model A because it may reflect market characteristics, such as proximity to long-term care facilities, rather than patient characteristics. It was considered in Model B because patients transferred from other inpatient facilities may be sicker than average at admission. This difference might not otherwise be captured in risk-adjustment models.

Stepwise forward selection procedures, forcing in all of the main effect and interaction variables from Model A, were used to select covariates. Model A covariates were forced into this model to ascertain the independent effects of additional demographic and clinical factors, controlling for those included in Model A. Candidate risk factors that were significantly associated with mortality at the $p < 0.10$ level in five or more of the ten subsamples were retained in Model B, except that race and at least one category of expected payment source were always included in Model B. This was done to adjust for the effects of socioeconomic variables, even if those effects were statistically insignificant.

The following Model B risk factors for AMI mortality were eliminated from the **"no prior hospitalization"** model because they were significant in fewer than five subsamples: sepsis; hyponatremia or hyposmolality; alkalosis; pneumonia; aspiration pneumonia; gastrointestinal hemorrhage; coagulopathy; bundle branch block; atrial fibrillation; supraventricular tachycardia; arterial thrombosis or embolism; acute peptic ulcer; skin ulcer; nonroutine source of admission; other nonwhite race; and Medicare insurance. The following AMI risk factors were eliminated from the **"one or more prior hospitalizations"** model for the same reason: sepsis; alkalosis; complete atrioventricular block; hypotension; pneumonia; aspiration pneumonia; bundle branch block; atrial fibrillation; arterial thrombosis or embolism; epilepsy; mitral valve disease; skin ulcer; nonroutine source of admission; emergent admission type; other nonwhite race; and Medicare insurance.

As in Model A, all risk factors that met the five-sample bootstrap criterion were tested in a stepwise regression using the full 60% estimation sample, with a p-to-enter of $p < 0.01$. This procedure eliminated no predictors from the AMI analyses; however, pleural effusion was dropped from the **"no prior hospitalization"** model because it had a counterintuitive negative coefficient (perhaps because the diagnosis is not reported in patients with severe pulmonary edema).

STEP 9: RE-ESTIMATION OF MODEL PARAMETERS USING ALL CASES

The 60% estimation sample and the 40% validation sample were re-combined into the full dataset. Model A and Model B were re-estimated by fitting the models developed in Steps 1 through 8 to the complete (100%) data set. The purpose of this step was to generate the most reliable possible estimate of each parameter, using all available data. As described in Step 7, a few interaction variables with questionable clinical significance and inconsistent parameter estimates based on internal validation were dropped at this stage.

AMI Model B demonstrated a serious problem with model fit in both the estimation and validation samples. Although Model B was intended to emphasize discrimination over calibration, it was found to overpredict death among high-risk patients to an unacceptable degree. This problem was attributable primarily to interactions involving the additional clinical risk factors included in Model B. Such interactions were not generally sought, but additional efforts to improve the calibration of AMI Model B were deemed necessary. Interactions were created between shock and related high-risk variables (e.g., anterior wall site, other or unspecified site, CHF, acidosis, hypotension, pulmonary edema, other cerebrovascular disease, acute renal failure). These interactions were tested in a stepwise logistic regression using the complete 100% sample, with a p-to-enter of $p < 0.01$ and all Model A variables and Model B main effects forced in. This effort ameliorated but did not entirely resolve the problem (see Chapter Ten).

The final models re-estimated in this step were used to calculate the predicted probability of an death for each case in the analysis. These predicted probabilities were used in all subsequent analyses of hospital mortality rates.