Technical Guide

OFFICE OF STATEWIDE HEALTH PLANNING AND DEVELOPMENT
OFFICE OF THE DIRECTOR
1600 Ninth Street, Room 433
Sacramento, California 95814
(916) 654-1606 FAX (916) 653-1448

Dear Colleagues:

On behalf of California's Office of Statewide Health Planning and Development (OSHPD), I am pleased to preface our latest report on the outcomes of care in California's hospitals. These outcome studies, mandated by legislation (Assembly Bill 524) signed by Governor Wilson in 1991, are based on data routinely abstracted from hospital medical records and reported to OSHPD for every patient discharged from a California hospital.

This third report on heart attack mortality rates expands and improves upon earlier studies using a larger body of data, refined risk-adjustment methods, and linkage to death certificate information. The study, therefore, represents an important contribution in efforts to evaluate the quality of health care provided throughout the state.

OSHPD had overall responsibility for the project. Andra Zach, R.R.A., M.P.A., served as coordinator. The statistical studies were performed by a distinguished team of researchers from the University of California medical schools at Davis and San Francisco, led by Patrick S. Romano, M.D., M.P.H., and Harold S. Luft, Ph.D. In addition, the Project had the benefit of valuable suggestions from several advisory bodies: the California Health Policy and Data Advisory Commission; its technical advisory committee, made up of representatives of the health services research, hospital, nursing, medical, health information and consumer communities; and from a panel of clinical experts in the field of cardiovascular disease.

OSHPD's primary goal in conducting such studies on outcomes of care, and reporting the results, is to improve the quality of hospital care available to all California citizens. The report provides hospitals with systematic information about their patient care results in comparison to other facilities, and encourages them to examine their processes of care to determine those which result in the best outcomes.

The AB 524 legislation responded to needs expressed by health care purchasers, providers and consumers to have publicly available information that objectively compares hospital performance in patient care. The legislation called for selection of medical, surgical, and obstetrical conditions for study of outcomes of hospital care. The first conditions selected were heart attack (acute myocardial infarction), back surgery (cervical and lumbar disk excisions), and maternal outcomes of obstetrical care (vaginal and cesarean deliveries). Several reports related to these studies have already been published. A study on the outcomes of care of hip fractures is in progress.

The Office of Statewide Health Planning and Development has made a long-term commitment to provide public information describing the quality of care delivered in California hospitals and, eventually, in other settings of care as well. With the assistance of its advisory bodies and colleagues in the health care community, the Office seeks continued improvements in data collection and analytical methods so as to enhance our ability to evaluate the performance of California's health care institutions.

The Office welcomes your comments and suggestions regarding these reports.

Sincerely,

David Werdegar, M.D., M.P.H.
Director

Report of the California Hospital Outcomes Project

Heart Attack Outcomes, 1991-1993

Office of Statewide Health Planning and Development

Lisa Simonson, Ph.D.
Deputy Director

Andra Zach, R.R.A., M.P.A.
Outcomes Project Manager

Study Consultants

University of California, Davis

Patrick S. Romano, M.D., M.P.H.
Principal Investigator

Julie Rainwater, Ph.D.
Project Director

Hong Zhou, Ph.D.
Senior Statistician

Michael Schembri
Statistician

Benjamin Chan, M.S.
Assistant Statistician

University of California, San Francisco

Harold S. Luft, Ph.D.
Principal Investigator

Linda L. Remy, Ph.D.
Project Director

Theodore H. Clay, M.S.
Statistician / Programmer Analyst

Acknowledgments

This report reflects the efforts of the contractors and staff in the Office of Statewide Health Planning and Development. Work was completed under contracts 93-4158 and 94-5321 with the University of California, Davis, and an intercampus agreement between the University of California, Davis (UC Davis), and the University of California, San Francisco (UC San Francisco). Suggestions to the formatting and presentation style of the *User's Guide* were provided under contract with the University of California, Los Angeles (UC Los Angeles).

Andra Zach, R.R.A., M.P.A. had overall responsibility for all aspects of the project, serving as liaison between OSHPD, the research team, and the included hospitals. OSHPD staff conducted the death certificate linkage and had primary responsibility for writing the *User's Guide.*

The contract team at UC Davis was supervised by Patrick S. Romano, M.D. and Julie Rainwater, Ph.D. The contract team at UC San Francisco was led by Harold S. Luft, Ph.D. Dr. Romano had primary responsibility for writing the *Technical Guide.* Dr. Luft and Dr. Romano had primary responsibility for writing materials accompanying the *Detailed Statistical Results.* Drs. Romano, Luft, and Rainwater, had primary responsibility for writing the *Hospital Guide*, with the assistance of Michael Schembri and Ben Chan, M.S.

Ben Chan, M.S. was responsible for most of the data processing and statistical analyses, with the assistance of Dr. Romano. Michael Schembri prepared the table for the *User's Guide* and provided programming for the *Detailed Statistical Results* and the *Hospital Guide*. Hong Zhou, Ph.D. consulted on statistical issues. Linda Remy, Ph.D., and Theodore Clay, M.S., created the study files from the OSHPD discharge data set.

Rebecka Levan, M.P.H., Leah Vriesman, M.H.A., M.B.A. and Gerald Kominsky, Ph.D. of UC Los Angeles provided suggestions for formatting and presentation style of the *User's Guide*.

A number of individuals provided guidance and advice in the design of the study and provided helpful comments on this or previous reports: Edward Hannan, Ph.D., Clifton Bailey, Ph.D., Lisa Iezzoni, M.D., Gregory S. Binns, Ph.D., along with members of the Technical Advisory Committee and the Clinical Panel.

Suggested Citation

Romano PS, Luft HS, Rainwater JA, Zach AP. *Report on Heart Attack 1991-1993, Volume 2: Technical Guide*. Sacramento, CA: California Office of Statewide Health Planning and Development, December 1997.

California Health Policy and Data Advisory Commission

Clark E. Kerr, Chair

Jacquelyn Paige, Executive Director

Technical Advisory Committee

William A. Waite
Committee Chair; 1992 - 1997
Member, California Health Policy and Data Advisory Commission

Jose Alberto Arevelo, M.D.; since 1992
Prudential Health Care of California

Douglas Bagley; since 1992
Los Angeles County Department of Health Services

Robert H. Brook, M.D., Sc.D.; since 1992
RAND/University of California, Los Angeles

Linda Blanchard, R.R.A., M.P.H.; since 1992
UCLA Medical Center

Mark Blumberg, M.D.; 1992 - 1997
Private consultant

David M. Carlisle, M.D., Ph.D.; since 1992
University of California, Los Angeles

David Chernof, M.D.; 1992 - 1997
Blue Cross of California, retired

Jerry Royer, M.D., M.B.A.; since 1992
Committee Chair; since 1997
Member, California Health Policy and Data Advisory Commission

Nancy Donaldson, R.N., D.N.Sc.; since 1997
Donaldson and Associates

Richard S. Frankenstein, M.D.; since 1992
Private practice, Garden Grove

Marcia Grant, R.N., D.N.Sc.; 1992 - 1997
City of Hope National Medical Center

Suzanne B. Henry, R.N., D.N.Sc.; since 1992
University of California, San Francisco

Lucy Johns, M.P.H.; since 1992
Health Care Planning & Policy

Peter I. Juhn, M.D.; since 1994
Kaiser Permanente

Ronald L. Williams, Ph.D.; 1992 - 1994
University of California, Santa Barbara

Heart Attack Clinical Panel

Melvin Cheitlin, M.D.
University of California
San Francisco

Myrvin Ellestad, M.D.
Memorial Heart Institute
Long Beach

Mark Hlatky, M.D.
Stanford University
Stanford

Julie Neil, R.N., A.R.T.
University of California
Davis

James Ritchie, M.D.
University of Washington
Seattle

Mary Woo, R.N., D.N.Sc.
University of California
Los Angeles

TABLE OF CONTENTS

TABLE OF CONTENTS, continued

Overview

Report on Heart Attack

The California Hospital Outcomes Project is an initiative mandated by the State of California, and conducted by the Office of Statewide Health Planning and Development (OSHPD), to develop public reports comparing hospital outcomes for selected conditions treated in hospitals throughout the state.

The Report on Heart Attack is intended to encourage all California hospitals to improve their care and give credit to the hospitals that are the leaders. It can also help insurers, employers, and consumers to select hospitals based on quality of care.

The California Hospital Outcomes Project

Heart attacks (acute myocardial infarctions or AMIs) were chosen as one of the first conditions to be reported upon by the California Hospital Outcomes Project because they are important, common, and deadly. Every year approximately 40,000 heart attack patients are admitted to 400 California hospitals. More than 5,000 of these persons die.

The mortality rates published in previous heart attack reports have been used in many ways. Hospitals have used their results to evaluate and improve their quality of care. Payers have used the reports to contract with the best hospitals. Consumers have used the reports to make more informed decisions.

The results published in this report are useful because:

- **They have been risk-adjusted.** Patient age, sex, type of heart attack, and chronic diseases were used to adjust for differences in patient risk when calculating hospital mortality rates.
- **They have been validated.** A major validation study involving nearly 1,000 heart attacks at 30 hospitals showed that variations in how hospitals report their data to OSHPD do not significantly affect their risk-adjusted death rates. In general, low-mortality hospitals treat heart attacks more aggressively than high-mortality hospitals.

Content of the Report on Heart Attack

This is the third report on heart attack. The first report was published in December of 1993 and the second report was published in May of 1996. This year's report includes heart attack cases from 1991 through 1993. Although 1991 and 1992 cases were included in last year's report, results shown in the current report may be different because the methodology has been improved. These improvements include:

- Linking with Vital Statistics records to ascertain deaths occurring outside the hospital.
- Refining certain patient risk-factor definitions based on the findings of the 1996 validation study.
- Using six months of pre-heart attack hospital records to more completely describe patient risk factors.

This year's report consists of five components:

The *User's Guide* (Volume 1) is intended for all those interested in hospital performance including hospital staff, employers, government agencies, health plans, and insurance companies. This volume provides a brief description of the study background and methods. It also contains two tables that display the results for individual hospitals based on heart attacks that occurred between 1991 and 1993.

The *Technical Guide* (Volume 2) is intended for health services researchers, health care providers, and others interested in the statistical methods used to calculate risk-adjusted death rates.

The *Detailed Statistical Results* (Volume 3) contains the numerical results for individual hospitals upon which the classifications in the *User's Guide* are based. In addition, there are tables that aggregate the results to the county level. It also contains a graphical representation of both individual hospital and county-wide results, which can be used to examine annual trends. An electronic version of the tables is available on diskette.

The *Hospital Comment Letters* (Volume 4) is intended to give readers of the *Report on Heart Attack* an appreciation of its strengths and weaknesses from the hospitals' perspectives.

The *Hospital Guide* accompanied patient specific information that was sent to each hospital several weeks before the Report on Heart Attack was published. Hospitals used this information to prepare their comment letters, which are provided with each volume of the report. More importantly, hospitals and their physicians can use this information to target areas where heart attack care might be improved.

To obtain these documents contact:

Office of Statewide Health Planning and Development
Data User's Support Group
818 K Street
Sacramento, CA 95814
(916) 322-2814

Documents, excluding the *Hospital Guide*, are available on the internet at **http:\\www.oshpd.cahwnet.gov**

Chapter 1

Using this *Guide*

The Technical Guide is intended for health services researchers, health care providers, and others interested in the statistical methods used to calculate risk-adjusted mortality rates.

Technical Guide Overview

This volume of the California Hospital Outcomes Report provides background information about the risk-adjustment models used to derive hospital-specific results for acute myocardial infarction (AMI). These risk-adjustment models were developed through a multi-step process that involved reviewing the scientific literature, convening an expert panel, developing criteria for including and excluding cases, identifying adverse outcomes, selecting risk factors, estimating statistical models, refining and testing these models, and calculating risk-adjusted outcome measures. The details of this process are described so others can replicate the results or apply the methods to other regions. While the research team believes the models developed and used in the California Hospital Outcomes Project are as good as possible given the available time, resources, and data, suggestions for improvement are welcome.

New in 1997

At the beginning of each chapter, a special section highlights any changes in the study methods since the 1996 report. The most important change since 1996 is the use of 30-day mortality, regardless of location, as the outcome of interest for patients with heart attacks. Previous reports only counted deaths that occurred in the hospital, because hospital discharge data were not linked with vital statistics data until last fall. This report also differs from previous reports in that more of California's acute care hospitals are included; data quality has improved so much that fewer hospitals qualify for exclusion. There are also minor changes in the risk-adjustment models, largely prompted by new information from OSHPD's recent AMI validation study about the coding of specific comorbid diagnoses.

Chapter 2

Literature Summary

An extensive review of the clinical literature in the area of AMI was undertaken. The MEDLINE bibliographic data base was searched for English language references since 1970, using relevant keywords. References also were identified through discussions with clinical advisory panel members and review of reference lists in relevant books and meta-analyses.

New in 1997

The literature review in this report is updated through 1995.

Identification of Risk Factors

All studies reporting on risk factors for in-hospital or 30-day mortality after AMI were obtained and reviewed. Studies from developing countries (e.g., Africa, South and Central America) and studies limited to patients who had specific procedures (e.g., angioplasty) or specific risk factors (e.g., multivessel coronary artery disease, anterior wall infarction) were set aside. Studies focusing on the predictive value of specific diagnostic tests (e.g., angiography, radionuclide ventriculography) were also set aside, along with studies that aggregated death with other adverse outcomes. Finally, studies based exclusively on administrative data, such as hospital discharge abstracts or Medicare claims, were not reviewed because of concerns about data quality. Among the remaining studies, those with at least 250 observations were assigned higher priority than those with fewer observations. Special attention was paid to studies that included multivariate analyses of the independent effects of multiple risk factors in large cohorts. The definitions of risk factors and associated odds ratios or risk ratios were abstracted from these studies.

Table 2.1 lists the major risk factors for in-hospital 30-day mortality after AMI, according to the clinical literature cited at the end of this chapter. These studies include large cohort studies and randomized controlled trials of various therapeutic interventions, such as thrombolysis and angioplasty. Because most of these trials excluded large groups of AMIs (e.g., patients presenting in cardiac arrest or without electrocardiographic changes), their results may not generalize to the entire population of AMI patients.

Each risk estimate in Table 2.1 represents the odds ratio or relative risk of death among patients with the characteristic, compared to those without the characteristic. (Studies that reported neither odds ratios, nor relative risks, nor sufficient raw data to calculate these measures, are not shown here.) The risk estimates listed are adjusted for other patient characteristics by

multivariate logistic regression, whenever possible. If multiple studies reported different risk estimates for the same factor, the range of these point estimates is shown. Point estimates are not shown if they are based on logarithmic or quadratic transformations of a risk factor, because interpretation of such coefficients would be difficult. Note that some of these risk estimates were not statistically significant, but they are shown to demonstrate the complete spectrum of values reported in the literature. Meta-analytic statistical techniques were not applied.

Application of the Literature Summaries

These literature summaries were used in two major ways. First, they were used to identify specific diagnoses generally regarded as risk factors for early death after AMI. These diagnoses were reviewed with all members of the clinical advisory panel and then adapted to ICD-9-CM, as described in Chapter Seven.

Most importantly, findings from previous studies, shown in Table 2.1, were compared with preliminary findings from the present study. Comorbidities that were far less common than expected, based on literature review (e.g., hyperlipidemia), were deleted from the list of candidate risk factors because it was believed they might be underreported to OSHPD. When the direction of the observed association between a risk factor and the adverse outcome differed from that reported in previous studies, further discussions or analyses were undertaken. If there was no apparent reason for the "counterintuitive" finding, that risk factor was deleted from the list of candidate covariates. Asthma and obesity are examples of risk factors which were associated with reduced AMI mortality. The most likely causes of such unexpected findings are either unmeasured confounders or selective underreporting of comorbidities among patients who died.

Table 2.1: Literature review of risk factors for in-hospital or 30-day mortality after acute myocardial infarction[1]

Risk Factor	*Risk Estimate[1]*	*Reference*
Age	1.5-2.5 per 10 yrs	2,3,8,17,19,21,27
	2.0-3.2 (>70 vs ≤70)	14,32,25
	2.2 (61-70 vs ≤60)	20
	2.0 (>55 vs ≤55)	22
	1.6 (>80 vs 65-80)	31
	1.9-6.4 (>65 vs ≤65)	9,23,26
Female sex	0.7-1.5	8,9,14,22,23,30,32
Diabetes mellitus	1.2-1.9	2,3,5,14,30
Previous AMI	0.9-1.9	2,9,10,13,14,22,23,25,26,30,32
Previous coronary artery bypass surgery	<1 (OR not specified)	9
Antecedent angina pectoris		
Any	0.8-2.6 (vs none)	2,3,5,9
CCS class III-IV	1.6 (vs none or class I/II)	9
Tobacco use		
Never smoker	1.4 (vs ever-smoker)	2
Current smoker	0.5-0.9 (vs nonsmoker)	5,9,22
History of prior congestive heart failure	0.9	24
History of hypertension	0.7	5
History of COPD	>1 (OR not specified)	13
History of cancer		
Any	0.8	24
Diffuse/metastatic	1.1	8
Ability to walk	1.1-1.6 (unable vs with assist)	8,24
	1.1-1.3 (with assist vs independent)	8,24
Number of body systems with acute or chronic disease	>1 (OR not specified)	17
Killip classification (see also individual components)	3.5-6.1 (>1 vs 1)	2,18
	3.4 (3 or 4 vs 1 or 2)	23
Bradycardia (first 48 hrs)	2.7	5
Tachycardia >100 BPM (at admission or first 48 hrs)	1.2-2.9	5,13,22,24,31
Heart rate (per minute)	1.14 per 1 unit	8,13,17
Respiratory rate (per minute)	1.04 per 1 unit over 12	24
	32.5 (<12 vs 12)	24
	2.0 (>30 vs ≤30)	31
Fever (first 48 hrs)	1.6	5
Rales (one third up)	2.2	10,13,14
Congestive heart failure, with rales or other signs	1.2-10.0	3,18,19,20,24,26
Systolic hypotension (≤90 or ≤95 mm Hg)	1.8-3.7	2,5,14,17,25,31
Mean arterial pressure	0.65 per 1 mm Hg	8,24
Cardiogenic shock at admission	5.0-36.0 (vs normal)	10,11,18,20,24,26,27,32
Cardiorespiratory arrest at admission	2.5	17,24
Body mass index	0.97 per kg/m^2	24
Heart murmur	2.2	31
Coma, stupor, lethargy, disorientation	4.1	31
APACHE II or III Physiology Score	1.46 per 1 unit	1,8,16,17
Abnormal chest radiograph:		
Cardiomegaly	1.0-3.0 (vs normal)	24,25
Congestive heart failure	1.2-1.9 (any)	8,31
	2.3 (interstitial)	25
	3.5-6.2 (pulmonary edema)	10,18,20,22,25

Table 2.1: Literature review of risk factors for in-hospital or 30-day mortality after acute myocardial infarction[1], *continued*

Risk Factor	*Risk Estimate*[1]	*Reference*
Abnormal electrocardiogram	17	6
ECG consistent with AMI	1.2-5.1	5,24,29,31
Site of infarction		
Anterior wall	1.5-2.7	3,10,12,13,14,17,20,23,24
Lateral wall	1.3 (vs inferior/other)	24
Posterior wall	1.9 (vs inferior/other)	24
Q-wave (transmural) infarction	1.2-4.1 (vs nontransmural)	4,5,7,8,15,23,25,28
Right ventricular infarction	7.7	32
Conduction/rhythm disturbance		
Any	1.5	24
Ventricular fibrillation	2.5-14.9	19,22
Asystole	30.0	19
Complete atrioventricular block	3.1	13,22
Atrial fibrillation	1.8-2.2	14,22
Number of leads with ST elevation	1.5 (4-5 vs 0-3) 2.2 (6-7 vs 0-3) 3.9 (>7 vs 0-3)	20
Hyperkalemia (first 48 hrs)	1.8	5
Serum urea nitrogen	1.19 per 1 mmol/L	8,17,24
Azotemia	1.7 (vs normal)	5
Serum creatinine	1.4 per 1 mg/dl over 2	24
Renal dysfunction (BUN > 10.5 mmol/L, Cr > 150 μmol/L)	2.0	31
Serum albumin	0.6 (>3 vs ≤3)	24
AST score	>1 (OR not specified)	17
Creatine kinase		
Peak CK > 1,000 U/L	1.9	32
Peak CK≥3,000 U/L	1.3	23
Peak CK > 8x normal	1.7 (vs <2x normal)	12
CK-MB > 14%	1.4 (vs 4-14%)	5
CK score	<1 (OR not specified)	13,17
Left ventricular ejection fraction	0.68 per 10%	21
Medications at admission		
Digoxin	1.3-1.7	5,22
Diuretic	1.9	22
Beta blocker	0.5-0.9	9,22
Aspirin	0.8	9
"Do not resuscitate" order at admission	1.2	8
Thrombolytic therapy	0.3	32
Emergent angioplasty, if appropriate	<1 (OR not specified)	30

1. Unless otherwise indicated, these figures represent estimates of the relative risk or odds ratio among those with the risk factor compared to those without the risk factor.

Table 2.2: Selected references for short term outcomes

1. Alemi F, Rice J, Hankins R. Predicting in-hospital survival of myocardial infarction, A comparative study of various severity measures. *Medical Care* 1990;28:762-775.

2. Barbash GI, Modan M, Goldbourt U, White HD, Van de Werf F. Comparative case fatality analysis of the international tissue plasminogen activator/streptokinase mortality trial: variation by country beyond predictive profile. *Journal of the American College of Cardiology* 1993;21:281-286.

3. Behar S, Reicher-Reiss H, Abinader E, Agmon J, Friedman Y, Barzilai J, Kaplinsky E, Kauli N, Kishon Y, Palant A, Peled B, Rabinovich B, Reisin L, Schlesinger Z, Zahavi I, Zion M, Goldbourt U. The prognostic significance of angina pectoris preceding the first acute myocardial infarction in 4166 consecutive hospitalized patients. *American Heart Journal* 1992;123:1481-1486.

4. Berger CJ, Murabito JM, Evans JC, Anderson KM, Levy, D. Prognosis after first myocardial infarction. Comparison of Q-wave and non-Q-wave myocardial infarction in the Framingham Heart Study. *JAMA* 1992;266:1545-1551.

5. Blumberg MS. Expected 30-day AMI mortality from models using UB-82 variables with and without added clinical variables. 1992; Draft submitted for publication.

6. Brush JE, Brand DA, Acampora D, Chalmer B, Wackers FJ. Use of the initial electrocardiogram to predict in-hospital complications of acute myocardial infarction. *The New England Journal of Medicine* 1985;312:1137-1141.

7. Connolly DC, Elvebach LR. Coronary heart disease in residents of Rochester Minnesota. VI. Hospital and posthospital course of patients with transmural and subendocardial myocardial infarction. *Mayo Clinic Proceedings* 1985;60:375-381.

8. Daley J, Jencks S, Draper D, Lenhart G, Thomas N, Walker J. Predicting hospital-associated mortality for Medicare patients: A method for patients with stroke, pneumonia, acute myocardial infarction, and congestive heart failure. *JAMA* 1988;260:3617-3623.

9. Davis KB, Alderman EL, Kosinski AS, Passamani E, Kennedy JW. Early mortality of acute myocardial infarction in patients with and without prior coronary revascularization surgery. A Coronary Artery Surgery Registry Study. *Circulation* 1992;85:2100-2109.

10. Dubois C, Pierard LA, Albert A, Smeets JP, Demoulin JC, Boland J, Kulbertus HE. Short-term risk stratification at admission based on simple clinical data in acute myocardial infarction. *American Journal of Cardiology* 1988;61:216-219.

11. Goldberg RJ, Gore JM, Alpert JS, Osganian V, deGroot J, Bade J, Chen Z, Frid D, Dalen JE. Cardiogenic shock after acute myocardial infarction. Incidence and mortality from a community-wide perspective, 1975 to 1988. *New England Journal of Medicine* 1991;325:1117-1122.

12. Hands ME, Lloyd BL, Robinson JS, De Klerk N, Thompson PL. Prognostic significance of electrocardiographic site of infarction after correction for enzymatic size of infarction. *Circulation* 1986;73:885-891.

13. Henning H, Gilpin EA, Coveli JW, Swan EA, O'Rourke RA and Ross J Jr. Prognosis after acute myocardial infarction:A multivariate analysis of mortality and survival. *Circulation* 1979;59:1124-1136.

14. Hillis LD, Forman S, Braunwald E, and the Thrombolysis in Myocardial Infarction (TIMI) phase II co-investigators.Risk stratification before thrombolytic therapy in patients with acute myocardial infarction. *Journal of the American College of Cardiology* 1990;16:3-315.

Table 2.2: Selected references for short term outcomes, *continued*

15. Hutter AM Jr., DeSanctis RW, Flynn T, Yeatman IA. Nontransmural myocardial infarction: a comparison of hospital and late clinical course of patients with that of matched patients with transmural anterior and transmural inferior myocardial infarction. *American Journal of Cardiology* 1981;48:595-602.

16. Iezzoni LI, Ash AS, Shwartz M, Daley J, Hughes JS, Mackiernan YD. Judging hospitals by severity-adjusted mortality rates: The influence of the severity-adjustment method. *American Journal of Public Health* 1996;86:79-87.

17. Keeler EB, Kahn KL, Draper D, Sherwood MJ, Rubenstein LV, Reinisch EJ, Kosecoff J, Brook RH. Changes in sickness at admission following the introduction of the prospective payment system. *JAMA* 1990;264:1962-1967.

18. Killip T, Kimball JI. Treatment of myocardial infarction in a coronary care unit. A two-year experience with 250 patients. *American Journal of Cardiology* 1967;20:457-464.

19. Madsen EB, Hougaard P, Gilpin E. Dynamic evaluation of prognosis from time-dependent variables in acute myocardial infarction. *American Journal of Cardiology* 1983;51:1579-1583.

20. Maggioni AP, Maseri A, Fresco C, Franzosi MG, Mauri F, Santoro E, Tognoni G, on behalf of the investigators of the Gruppo Italiano per lo Studio della Sopravvivenza nell'Infarto Miocardico (GISSI-2). Age-related increase in mortality among patients with first myocardial infarctions treated with thrombolysis. *New England Journal of Medicine* 1993;329:1442-1448.

21. Muller DWM, Topol EJ, Ellis SG, Sigmon KN, Lee K, Califf RM, and the Thrombolysis and Angioplasty in Myocardial Infarction (TAMI) Study Group. *American Heart Journal* 1991;121:1042-1049.

22. Nidorf SM, Parsons RW, Thompson PL, Jamrozik KD, Hobbs MST. Reduced risk of death at 28 days in patients taking a $ blocker before admission to the hospital with myocardial infarction. *British Medical Journal* 1990;300:71-74.

23. Nishi N, Nakanishi N, Kuroda K, Koizumi K, Yoshikawa J, Tatara K. Relationships of site of infarction and history of previous infarction with short- and long-term prognosis after acute myocardial infarction in Japan. *Journal of Clinical Epidemiology* 1992;45:39-45.

24. Normand SLT, Glickman ME, Sharma GVRK, McNeil BJ. Using admission characteristics to predict short-term mortality from myocardial infarction in elderly patients. Results from the Cooperative Cardiovascular Project. *JAMA* 1996;275:22-28.

25. Norris RM, Brandt PWT, Caughey DE, Lee AJ, Scott PJ. A new coronary prognostic index. *Lancet* 1969;1:274-278.

26. Peel AAF, Semple T, Wang T. Lancaster WM, Dall JLG. A coronary prognostic index for grading the severity of infarction. *British Heart Journal* 1962;24:745-760.

27. The Pennsylvania Health Care Cost Containment Council. *Focus on Heart Attack in Pennsylvania. Research methods and Results*. Harrisburg, PA:1996.

28. Rigo R, Murray M, Taylor DR, Weisfeldt ML, Strauss HW, Pitts B. Hemodynamic and prognostic findings in patients with transmural and nontransmural infarction. *Circulation* 1975;51:1064-1070.

29. Rouan GW, Lee TH, Cook EF, Brand DA, Weisberg MC, Goldman L. Clinical characteristics and outcome of acute myocardial infarction in patients with initially normal or nonspecific electrocardiograms (A report from the Multicenter Chest Pain Study). *American Journal of Cardiology* 1989;64:1087-1092.

Table 2.2: Selected references for short term outcomes, *continued*

30. Stone GW, Grines CL, Browne KF, Marco J, Rothbaum D, O'Keefe J, Hartzler GO, Overlie P, Donohue B, Chelliah N, Timmis GC, Vlietstra R, Strzelecki M, Puchrowicz-Ochocki S, O'Neill WW. Predictors of in-hospital and 6-month outcome after acute myocardial infarction in the reperfusion era: The Primary Angioplasty in Myocardial Infarction (PAMI) trial. *Journal of the American College of Cardiology* 1995;25:370-377.

31. Van Ruiswyk J, Hartz A, Kuhn E, Krakauer H, Young M, Rimm A. A measure of mortality risk for elderly patients with acute myocardial infarction. *Medical Decision Making* 1993;13:152-160.

32. Zehender M, Kasper W, Kauder E, Schonthaler M, Geibel A, Olschewski M, Just H. Right ventricular infarction as an independent predictor of prognosis after acute inferior myocardial infarction. *The New England Journal of Medicine* 1993;328:981-1038.

Chapter 3

Selection Criteria

The AMI analysis was designed to focus on fresh AMI admissions to acute care hospitals in California. The goal was to select patients who had just experienced an acute heart attack due to coronary artery disease. Inclusion and exclusion criteria were developed after careful review of the medical literature and extensive discussions with an expert panel that included cardiologists, health services researchers, a cardiac care nurse, and a health information management professional.

New in 1997

- This report includes AMIs admitted to California hospitals between January 1, 1991 and December 1, 1993. Last year's report included AMIs admitted between August 26, 1990 and May 31, 1992. The next report, to be released by the end of 1997, will include AMIs admitted through 1994.

- A very minor change was made in the list of allowable principal diagnoses, based on the results of last year's AMI validation study.

Inclusion Criteria

AMI cases were identified by reviewing the discharge abstracts from all acute care hospitals in California that report data to the Office of Statewide Health Planning and Development (OSHPD). These hospitals do not include facilities operated by the US Department of Veterans Affairs or the Department of Defense. Discharge abstracts that were identified as coming from a non-acute level of care (e.g., skilled nursing, rehabilitation) were not reviewed.[1] Cases selected for the study were required to meet all four of the inclusion criteria listed below.

1. **A principal diagnosis of acute myocardial infarction, initial or unspecified episode of care (410.x0 or 410.x1), *or* a principal diagnosis of a presumed AMI complication with a secondary diagnosis of AMI, initial or unspecified episode of care.**

   The principal diagnosis is "the condition established, after study, to be chiefly responsible for occasioning the admission of the patient to the

1. Before January 1, 1995, hospitals were not required to submit separate reports (or bundles of discharge abstracts) for each type of care they provide. As a result, it was impossible to identify with certainty the skilled nursing discharges from 61 of the 153 California hospitals that had distinct skilled nursing units in 1990. Similarly, it was impossible to identify with certainty the rehabilitation discharges from 47 of the 80 California hospitals that had distinct rehabilitation units in 1990. In 1993, 81 of 210 hospitals did not report their skilled nursing patients separately and 44 of 88 hospitals did not report their rehab patients separately.

hospital for care."[2] Note that cases with a principal diagnosis of 410.x2 (AMI, subsequent episode of care) were not included because the focus was on fresh admissions requiring urgent diagnosis and management. Cases with a principal diagnosis of 410.x0 (AMI, unspecified episode of care) were included because they were clustered at certain facilities and their overall mortality rate and other characteristics closely resembled 410.x1 cases (AMI, initial episode of care). These facilities appear to be improperly coding some initial AMI hospitalizations as 410.x0.

Table 3.1 lists the principal diagnoses that were presumed to represent AMI complications. At some hospitals, patients who presented with one of these cardiovascular complications were assigned a principal diagnosis of AMI and a secondary diagnosis of the observed complication. At other hospitals, the complication was coded as the principal diagnosis because coders failed to appreciate the temporal sequence. To capture similar cases from both sets of hospitals, patients with principal diagnoses of suspected AMI complications were included in the sample. Cases with other principal diagnoses were not included because their AMIs may have resulted from unrelated conditions. Several conditions that appeared on the list of acceptable principal diagnoses in 1993 and/or 1996, such as arterial thrombosis, hypotension, and complete atrioventricular block, were removed this year because OSHPD's validation study suggested that AMIs in these patients are often secondary to other conditions or procedures, such as arterial bypass surgery, sepsis, and conducting system disease, respectively.

Although coding guidelines allow respiratory failure (518.81-518.82) to be coded as the principal diagnosis when it follows an AMI, it was not included on the list of allowable principal diagnoses because most such cases had an indeterminate infarction site and an underlying diagnosis of pneumonia or chronic obstructive pulmonary disease. These findings suggest that AMIs more often were complications rather than causes of respiratory failure.

2. **Age at admission of 18 years or greater.**

   Children were not included because the pathophysiology of AMI in this population usually relates to a congenital anomaly or an acute ischemic event rather than coronary artery disease.

3. **Source of admission equal to routine (11), emergency room (12), other facility (16), home health service (17), or other (19).**

   Patients transferred in from other acute care hospitals (13) were not included in the primary analysis. Instead, these records were linked whenever possible with the corresponding record from the original admitting hospital, so that the patient's ultimate outcome could be attributed back to the hospital that provided the initial care.

2. OSHPD, 1991. Discharge Data Tape Format Documentation.

Patients transferred in from skilled nursing (14) or intermediate care facilities (15) were not included to minimize the number of patients in the sample with "do not resuscitate" (DNR) orders. Patients with DNR orders have a high risk of death, both because of their underlying medical problems (which may not be captured in the risk-adjustment model) and because they are not candidates for life-prolonging interventions, such as mechanical ventilation. Many of these patients are admitted only for comfort care.

Cases admitted from other facilities (16) were included because OSHPD's 1988 reabstraction study showed that most (69 percent) of these cases should have been reported as emergency room admissions. Therefore these cases were grouped with emergency room admissions.

4. **Date of admission between January 1, 1991 and December 1, 1993 (inclusive).**

   As described in Chapter Four, the encrypted social security number and date of birth were used to link prior and subsequent records for each case. Reporting of social security numbers in California began on July 1, 1990. AMI cases admitted between July 1 and December 31, 1990 were available, but were excluded to provide a full 6 month period before admission to ascertain additional information about severity-of-illness for every case. Including three years of AMI cases made it possible to examine mortality trends over time, both statewide and for individual hospitals.

   Cases admitted from December 2 through December 31, 1993 were excluded because discharge records after December 31, 1993 were not available when this study was conducted. Therefore, 30-day outcomes were unknown for some of these cases. Note that the admission date was always used for case selection because it most closely approximates the actual date of the AMI.

Record Linkage

Records for patients transferred from one hospital to another within California were linked (as described in Chapter Four). Linkage was used to combine multiple records on the same patient into a single episode of care. This means that information from a series of discharge abstracts for a patient transferred from one facility to another was combined, and the disposition of the final hospitalization (e.g., death or survival) was ascribed to the "index" hospital. The "index" hospital was the first facility in a series of linked transfers that reported a **qualifying** AMI admission (based on the above inclusion criteria). That admission was labeled the "index" AMI, and need not have been the first admission in the transfer series.

The purpose of this procedure was to eliminate differing transfer rates as a cause of outcome differences across hospitals, and to accumulate as much

information as possible about each AMI case. A strategy was developed to maximize the number of correct matches and to minimize the number of erroneous matches. This strategy is described in Chapter Four.

Exclusion Criteria

Finally, several exclusion criteria were defined to eliminate cases that may not truly represent fresh AMIs, such as unstable angina that was potentially misdiagnosed by physicians or misinterpreted by coders. Because the index record alone was not always sufficient to establish the presence or absence of these exclusion criteria, they were applied after linkage. Cases with any of the following characteristics were excluded:

1. **One or more prior AMI admissions within the 8 weeks preceding the index AMI admission.**

   An AMI was excluded from the study if it was preceded by a **prior AMI admission** within 8 weeks (from admission date to admission date). Prior AMI admissions were defined by a principal or secondary diagnosis of 410.x0 or 410.x1, without regard to the patient's age, source of admission, or type of care, or to other inclusion and exclusion criteria listed in this chapter. For example, an AMI that occurred in a skilled nursing or intermediate care facility would not have been eligible for this study, but would have counted as a prior AMI and thereby disqualified any AMI admission during the next 8 weeks. An AMI in a patient admitted for gallbladder disease would not have been eligible for this study (because it might have been a postoperative complication), but still would have counted as a prior AMI.

   This exclusion is important for two reasons. First, many patients are admitted for acute management of an AMI, then go home and return to the hospital several weeks later for diagnostic evaluation or coronary revascularization. The *International Classification of Diseases, Ninth Revision, Clinical Modification* (ICD-9-CM) directs coders to classify any AMI less than eight weeks old as acute (410.xx), although it offers a fifth digit to distinguish the initial episode of care from subsequent episodes. If prior AMI admissions had not been sought, the same AMI might have been inadvertently double-counted. It is also important to identify prior AMIs because some people suffer a second AMI very shortly after their first, and these reinfarcts confer an increased risk of death. Such reinfarcts had to be excluded to obtain a relatively homogeneous sample of AMIs.

   Note that a prior AMI did **not** disqualify an index AMI if the patient was transferred from the prior facility to the index facility (e.g., the two records were part of a transfer series). By definition, the index AMI in a transfer series was the first record that met the four inclusion criteria listed above. In addition, a prior AMI did **not** disqualify an index AMI if the prior AMI record was actually part of a separate transfer series (or "episode of

care") that started with another index AMI admission outside the 8 week prior interval.

2. **Total length of stay less than 2 days (e.g., 0-1 day) with an ultimate disposition other than the following: acute hospital (2), discharge against medical advice (6), or death (8).**

   Note that the ultimate disposition is the one reported on the last in a series of linked records, if a patient was transferred from one facility to another. The total length of stay in this situation was calculated by adding the lengths of stay across hospitals.

   After excluding deaths, inter-hospital transfers, and discharges against medical advice (all of which had artificially truncated hospital stays), short hospitalizations were thought to represent remote infarctions, trivial infarctions (e.g., cardiac enzyme elevation without electrocardiographic changes), or patients who actually "ruled out" for AMI. The clinical advisors unanimously agreed that a hospital stay of two or more days remains the standard of care for fresh AMIs in California. ICD-9-CM guidelines require coders to assign the AMI code (410.xx) to the diagnosis of "rule out" myocardial infarction, unless an alternative diagnosis has been established.[3] Previous research has confirmed that patients discharged with a diagnosis code of 410.xx after a short stay often ruled out for AMI or were admitted for post-AMI diagnostic evaluation.[4] Other investigators have excluded short-stay patients for the same reason.[5]

3. **An external cause-of-injury (E) code indicating a transport accident of any type (E800.x-E848) from the index record or any subsequent linked record.**

   These cases were excluded because of concern that traumatic myocardial contusions, which usually result from steering column impact, may be misclassified as AMIs. Traumatic injury can lead to elevated cardiac enzymes and electrocardiographic changes that mimic those seen in acute infarction.

3. *Coding Clinic for ICD-9-CM 1985*; 2(2):3.
4. Iezzoni, LI, Burnside S, Sickles L, Moskowitz MA, Sawitz E, Levine PA. Coding of acute myocardial infarction: Clinical and policy implications, *Annals of Internal Medicine* 1988; 109:745-751.
5. Udvarhelyi IS, Gatsonis C, Epstein AM, Pashos CL, Newhouse JP, McNeil BJ. Acute myocardial infarction in the Medicare population. Process of care and clinical outcomes. *JAMA* 1992; 268:2530-2536.

Table 3.1: ICD-9-CM codes for principal diagnoses presumed to represent AMI complications if the case had a secondary diagnosis of 410.x0 or 410.x1

ICD-9-CM Code	*ICD-9-CM Description*
427.1	Paroxysmal ventricular tachycardia
427.41	Ventricular fibrillation
427.42	Ventricular flutter
427.5	Cardiac arrest
429.5	Rupture of chordae tendinae
429.6	Rupture of papillary muscle
429.71	Acquired cardiac septal defect
429.79	Other sequelae of myocardial infarction
429.81	Other disorders of papillary muscle
518.4	Acute edema of lung, unspecified
780.2	Syncope and collapse
785.51	Cardiogenic shock, without mention of trauma

Chapter 4

Linking Hospitalization and Death Records

Record linkages are important for several reasons. First, linkages with subsequent hospital discharge abstracts and death certificates help identify patients' outcomes (e.g., death within 30 days). Otherwise, hospitals that transfer or prematurely discharge their sickest AMI patients might have unduly low death rates. Second, linkage makes it possible to identify fresh AMIs as described in Chapter Three. Third, linkages provide important information about clinical risk factors. Diabetes and other comorbidities are not always coded on discharge abstracts, so more complete information can be obtained when multiple records are available.

New in 1997

! In this year's report, California hospital discharge abstracts were linked with vital statistics records (death certificates) for the first time. The linkage methods are briefly summarized in this chapter, but a complete description is available upon request from OSHPD.

Linking Hospitalizations

The purpose of this chapter is to describe the linkage methods developed for the acute myocardial infarction (AMI) study. The goal of this linkage process is to identify relevant hospital discharge abstracts, order them temporally, and create a linked single-record analysis file summarizing information from all related abstracts.

Step 1. Identify records that meet initial selection criteria

The first step in record linkage was to create a condition file containing all records that (a) met preliminary inclusion criteria and (b) were within the time window used to select cases.[6] These preliminary inclusion criteria are described in Chapter Three. The window period included admissions between August 26, 1990 and December 1, 1993 (inclusive). This search generated 142,091 records, which became *candidates* for study. Note that many of these candidate records were excluded or relabeled as prior or transfer admissions, after the additional steps described below.

6. The master OSHPD database was used to create the condition file. Before starting the search, all records with valid SSNs were extracted from the master OSHPD hospital discharge database and divided into discrete files containing all records with valid SSNs for each month. The monthly files were sorted by SSN to simplify searching and to improve mainframe data management of the extremely large OSHPD master data set.

Step 2. Find all additional records with linkage potential

The goal of this step was to find any additional records within the study period that might link with the AMI records identified in Step One. To begin, invalid social security numbers (SSNs) were identified in three ways and set to missing: (1) SSNs with certain repetitive patterns that hospitals use to designate missing values (i.e., 111-22-3333), (2) SSNs associated with multiple dates of birth on different records, and (3) SSNs outside the valid values provided by the Social Security Administration. For example, analysis of the OSHPD masterfile found one California facility that assigned the same SSN to every emergency room admission over a 4-month period.

To start the search, the AMI condition file was divided into three subfiles. The largest subfile contained 133,314 records with valid social security numbers, the second subfile contained 8,618 records that lacked SSNs but did have other information useful for linkage (e.g., date of birth, gender, zip code), and the smallest subfile contained 159 records with neither valid SSNs nor demographic data. Among the records that had both SSNs and demographic data, 203 pairs were found to have the same date of birth, gender, and zip code, but SSNs with one discrepant digit. These minor SSN discrepancies were presumed to represent data entry errors, so the SSNs involved were recoded.

Two lookup files were constructed from the largest condition subfile. These lookup files were used to search for other records within the study frame (e.g., prior or transfer hospitalizations) which might be related to the AMI records already pulled. Lookup file 1 contained one entry for each unique SSN (N=120,596). Each entry specified all of the admission dates and birth dates associated with that SSN. No SSN was associated with more than two birth dates. Lookup file 2 contained one entry for each unique combination of birth date, sex, and 5-digit zip code (N=120,553), after setting aside 586 combinations that were associated with two or three SSNs and therefore could not be used to identify individuals.

The lookup files were used to locate all potential records for the study. This process involved four steps:

2.1 Using lookup file 1, all records with an exact SSN match were extracted from the monthly OSHPD master files, if at least two of three birth date elements (i.e., month, day, year) also matched (or if one element matched and the other two were transposed). If an SSN in the lookup file was associated with two different birth dates, both birth dates were checked as potential matches for a candidate record. Birth dates were used to confirm linkages because of the danger of improperly linking records that appeared to have the same SSN because of a data entry error. This step found 222,739 records with matching SSNs (in addition to the 133,314 AMI condition records which had been pulled originally).

2.2 Lookup file 2 was matched against the AMI condition subfile that contained records without SSNs. An exact match was required on birth date, sex, and 5-digit zip code. Records in the condition file that lacked

SSNs but matched entries in the lookup file were assigned the SSNs associated with those entries. As a result, SSNs were assigned to 603 of the 8,618 records that lacked SSNs but had other data elements useful for linkage. This left 8,015 records with valid birth dates, genders, and zip codes, but without SSNs.

2.3 Lookup file 2 was then used to search for matching records in the monthly OSHPD master files. An exact match was required on birth date, sex, and 5-digit zip code; any prospective match in the OSHPD masterfile had to have a soft-matching (8 of 9 digits) SSN or no SSN. (All exact SSN matches were already pulled in Step 2.1.) This step found 24,306 new matches that were not in the original AMI condition file, of which 3,319 had 1,958 soft-matching SSNs. These minor SSN discrepancies were presumed to represent data entry errors, so the SSNs involved were recoded.

2.4 In a final effort to determine SSNs for the 8,015 residual records without SSNs, each was matched against the monthly OSHPD master files using birth date, sex, and 5-digit zip code. An additional 9,274 matching records were pulled from the masterfile, of which 6,120 had 3,328 new SSNs. In each of these 6,120 cases, the SSN found on the demographic-matching record was assigned to the condition record that lacked an SSN.

2.5 There were 5,284 newly identified SSNs from Steps 2.3 and 2.4, which were then used to pull all matching records from the OSHPD masterfile. A total of 1,362 new records were pulled in this step; these records matched exactly on SSN and at least two of three birth date elements (or else one element matched while the other two were transposed). Several cases where two different combinations of birth date, sex, and 5-digit zip code linked to the same SSN were manually reconciled.

Step 3. Delete Duplicate Records and Resequence Record Sets

The files created in Step 2 above were joined and sorted by recoded SSN, admission date, discharge date, date of birth, sex, OSHPD facility number, total charges, diagnosis-related group (DRG), number of diagnoses, and number of procedures. The purpose of sorting by these variables was to identify duplicate records from the same hospital with the same SSN, admission and discharge dates, birth date, sex, total charges, DRG, and numbers of diagnoses and procedures. Of 69 such pairs,[7] all but 6 differed by type of care (e.g., "acute care" versus "rehabilitation"). One randomly selected record from each of the 69 pairs was retained, except that if an "acute care" AMI condition record was paired with a non-acute care record, only the acute care record was retained.

The next sort identified 32 pairs of records from the same hospital with the same SSNs and admission and discharge dates. Twenty of these pairs

7. All numbers cited in this chapter come from analyses performed before certain hospitals were excluded for extreme coding practices. These numbers may therefore differ from numbers that would be obtained from analysis of the final data set.

differed by type of care (e.g., "acute care" versus "rehabilitation"); the acute care record from each pair was retained. Among the remaining 12 pairs, one record generally appeared to be a more complete, or corrected, version of the other. The following variables were reviewed sequentially to determine the more complete record: total charges, number of diagnoses, number of procedures. In each case, the record with the higher value was retained.

The next sort identified 2 pairs of records with the same SSNs and admission and discharge dates, but from different hospitals with different types of care. Only one record in each pair was from an acute-care facility, so this record was retained while the other was discarded.

Finally, 37 pairs of acute-care records with the same SSNs and admission and discharge dates, but from different hospitals, were manually reviewed. These patients were apparently admitted to one acute care hospital, transferred to another, and then discharged, all on the same day. Each set was manually sequenced based on the discharge disposition and admission source. Any record with a disposition of "death" was sequenced last. Any record with a disposition of "general acute care hospital" was sequenced first.

After dropping duplicate records and resequencing these 37 pairs, the file was divided into a subfile containing recoded SSNs with only one record (which did not require linkage) and another subfile containing recoded SSNs with multiple records.

Step 4. Order Records in the Period Around the Admission

All records for a given SSN were extracted in Step 2, including some admissions that were irrelevant to the AMI study. For example, a person treated for AMI could have been admitted several months later for appendicitis. The goals of Step 4 were to define the periadmission period, which consists of the index AMI admission and the records around it, and to delete irrelevant records. This was done in four steps: (1) the index admission was identified, (2) transfer records were identified, (3) prior admissions were identified, and (4) the periadmission number was assigned.

The first step in defining a periadmission period was to identify the index AMI record, according to the inclusion criteria described in Chapter Three. At this point, some admissions and their subsequent transfers or readmissions were marked for exclusion, as described in Chapter Three.

The next step was to identify transfer records. Very specific criteria were established to classify subsequent hospitalizations as transfers. These criteria were necessary because most hospitalizations after AMI relate to evaluation or surgical therapy of coronary artery disease and do not belong to the periadmission period. Subsequent SNF/ICF admissions also do not belong to the periadmission period. Some patients experienced several transfers during the periadmission period; the last transfer represented the outcome record (as long as it occurred within 30 days of the AMI). The

specific criteria used to evaluate potential linkages with subsequent hospitalizations varied as follows:

4.1 Candidate records with a "report type" of skilled nursing and intermediate care (3), psychiatric care (4), alcohol/drug care (5), or rehabilitation care (6) were not evaluated.

Step Two pulled many records that were not from general acute care hospitals. These were used to identify prior admissions, but were not used to identify transfers.

4.2 Candidate records with a "report type" of general acute care (1) were categorized according to the discharge disposition of the immediately prior record and included or excluded, as follows:

a. *Intermediate care facility (03) or skilled nursing facility (04)*. No subsequent records were linked.

b. *Other facility (05)*. OSHPD's 1988 reabstraction study showed that some cases reported to have this discharge disposition were actually transfers to acute care hospitals (02). Therefore, subsequent records were linked when: (1) the admission date was the same as the preceding discharge date, and (2) the hospital identification number was different from that on the preceding record (suggesting that the patient may have remained at the same level of care), and (3) the principal diagnosis on the candidate transfer record was neither rehabilitation (V57.xx) nor psychiatric (290.x-319).

c. *Acute care hospital (02).* Some cases with this discharge disposition appear to have been transferred to lower levels of care. Therefore, subsequent records were linked only when: (1) the hospital identification number was different from that on the preceding record (suggesting that the patient may have remained at the same level of care), and (2) the admission date was up to one day later than the preceding discharge date (allowing for late night transfers), and (3) the principal diagnosis on the candidate transfer record was neither rehabilitation (V57.xx) nor psychiatric (290.x-319). When a patient was readmitted to an acute care hospital more than one day after a prior discharge, the second hospitalization was regarded as a separate episode of care and not a transfer.

d. *Routine (01), against medical advice (06), or home health service (07)*. Some patients were discharged to home or left against medical advice and returned to a hospital later the same day. These patients were still in the acute phase of care when they were readmitted, so their hospitalizations needed to be linked. Subsequent records were linked only when: (1) the admission date was the same as the preceding discharge date, and (2) the principal diagnosis on the candidate transfer record was neither rehabilitation (V57.xx) nor psychiatric (290.x-319).

Next, all records that preceded an index AMI record but fell within the study frame were classified as prior admissions. To be considered as a prior admission, a record had to have a recoded SSN that matched, by at least eight of nine digits, a recoded SSN found on either an index AMI or transfer record. A total of 26,548 records were discarded because they were linked to index AMI or transfer records only by demographic variables (e.g., date of birth, gender, zip code). A lookup file was then created to determine if a record had an admission date 0 to 180 days before the index AMI admission. If so, it was flagged as a prior admission. All prior, index, and transfer admissions related to a single AMI were grouped into a periadmission period. A total of 190,502 records not flagged as prior, index, or transfer admissions were discarded.

After the multiple record file was ordered, it was recombined with the single-admission file from Step 3 to create the periadmission file. A new variable was created to group sets of records (prior, index, transfer) into distinct periadmission periods. This grouping variable was needed because some patients had multiple periadmission periods within the study frame. The periadmission file contained one-to-n periadmissions composed of one-to-n records for each SSN.

Step 5. Create the Linked Single-Record Analysis File

The purpose of this step was to transform the periadmission file into a linked analysis file containing one record per periadmission. The transformation began by running programs which used all clinical information from all records in the periadmission file to describe the frequency of all diagnoses and procedures, and their relationship to the study outcomes.

The periadmission file was then used as input for a complex program summarizing the diagnoses and procedures from prior, index, and transfer records into clinical risk factors, as described in Chapter Seven.[8] Ethnicity and date of birth can be recorded differently from one record to another, and source of payment can change from one hospital to another. Therefore, index-record values for these variables were retained. After eliminating hospitals with unusual coding (Chapter Six) and creating random subsets of the file (Chapter Eight), the linked analysis file was ready for statistical modeling.

Reliability of Linked Hospitalizations

At least two variables that could have been used to help link hospital discharge abstracts were deliberately disregarded in the linkage process.

8. Only variables from the index AMI admission can be returned to the index hospital for review and comment. The risk factor program flagged cases which required special handling for this reason. Two variables were created in the linked analysis file to count records labeled as prior admissions and transfers. If either variable was greater than zero, clinical risk factors that could have been obtained from prior or transfer admissions are set to missing in the file returned to hospitals.

Expected principal source of payment was not used because a patient's insurance status often changes from one hospitalization to the next. Race was not used because the definitions may be subjectively applied, and because the overall error rate was reported as 6 percent with 56 percent underreporting of Asian ancestry in OSHPD's 1988 reabstraction study. Of 13,587 AMI cases with a linked prior or transfer record in OSHPD's 1996 report, race differed from that on the index record for 1,561 (11.3 percent). For example, 610 index AMI records indicated White race but another record with the same SSN indicated a different race.

The variables actually used for linkage were the SSN, date of birth, sex, and zip code. None of these variables are perfectly coded, so linkage problems arise. For example, Hispanic patients were far more likely to be missing SSNs than white patients. Patients in southern California and those admitted to large public hospitals were most likely to be missing SSNs. Reporting practices have not changed substantively over time, except at certain Kaiser facilities in northern California that experienced difficulty implementing the SSN reporting requirement. These findings indicate that patients without SSNs differ systematically from patients with reported SSNs. Although an algorithm for linking records without SSNs was developed, it is still likely that transfer rates were underestimated among patients without SSNs.

Several hundred records were found with the same SSNs as index records, but with different values for demographic variables. For example, a 21 year old Black female and a 75 year old Hispanic male, reportedly with the same SSN, were admitted to the same hospital. The former patient had a normal delivery; the latter patient had an AMI. Possible explanations for this problem include: (1) these SSNs correspond to invalid social security numbers that were not identified by OSHPD staff before encryption; (2) hospital employees entered social security numbers incorrectly; (3) multiple people used the same social security number; and (4) patients reported incorrect social security numbers.

A match on two of three birth date elements (i.e., month, day, year) was used to confirm linkage of records based on SSNs. In OSHPD's 1996 report, date of birth discrepancies occurred in 5.5 percent of all AMI cases with multiple linked records.

Linking Death Certificates to Hospitalizations

In 1996, OSHPD began linking its Patient Discharge Data Set (PDDS) with the Department of Health Services' Vital Statistics (VS) or death certificate file. This linkage was designed to provide complete ascertainment of deaths up to 365 days after hospital discharge, with partial ascertainment of deaths up to two years after discharge (although only deaths within 30 days of an index AMI admission were counted in this report). Each death certificate was linked to all applicable records in the PDDS, but each PDDS record was linked to zero or one death certificates. The linkage was performed deterministically, which means that specific criteria were applied and

evaluated. Key elements of this linkage procedure are described below, but a more complete description is available upon request from OSHPD's Health Policy and Planning Division. Subsequent to creation of the linked files used in this report, OSHPD has modified the linkage procedure thus addressing some of the problems identified here. In addition, a probabilistic linkage methodology is in development.

In summary, OSHPD searched sequentially for three types of linkages: (1) death certificates that matched perfectly on SSN, birthdate, sex, race, and 5-digit zip code; (2) death certificates that matched exactly on birthdate, sex, race, and 5-digit zip code, but not on SSN; and (3) death certificates that matched exactly on SSN, but not on demographic variables. OSHPD did not search for "soft" linkages involving both SSN and demographic variables (e.g., a one-digit discrepancy on both SSN and date of birth). Among the second and third types of linkages, the degree of matching or mismatching were prioritized and labeled. For example, birth date mismatches were classified into the following categories, among others: "wrong century, otherwise perfect," "year and date transposed, month correct" "month and year transposed, date correct," "month and date transposed, year correct," and "dates not equal." The poorest demographic matches (given an exact SSN match) were discarded and the poorest SSN matches were discarded. For example, a hospital record that was an exact SSN match but a total birthdate mismatch with a death record was discarded.

Two problems were identified in the matching algorithm, which required correction. Both problems related to improper prioritization when two or more death certificates potentially linked to the same hospital record.

1. Among PDDS records without a perfect match on all linking variables (e.g., SSN, birthdate, sex, race, zip code) in the VS file, the algorithm searched for exact demographic matches before it searched for exact SSN matches.[9] As a result, a death certificate that was an exact demographic match with a hospital record, but a total SSN mismatch, might have been selected over another death certificate that was an exact SSN match but had a minor demographic mismatch (e.g., one digit of birthdate or zip code). The SSN should be given highest priority in matching because it is generally unique to each individual. Therefore, exact SSN matches should be selected over exact demographic matches.

2. Among hospital records with an exact demographic match in the VS file but without an exact SSN match, the algorithm did not search for a death record with a soft matching SSN. Therefore, a relatively good, but not exact, SSN match might have been overlooked in favor of a non-match if both death certificates were exact demographic matches with the hospital record.

9. This was done to address concerns about shared and missing SSNs. The algorithm has been changed to search for matches on SSN before demographics.

These problems were solved by searching the VS file for additional potential linkages with the 1991-1993 AMI data set. There were 2,530 hospital records that did not have a discharge disposition of "death" but linked to a death record using demographic variables alone. Exact SSN matches were found for 106 of these records. By definition, none of these exact SSN matches was also an exact demographic match. However, 62 of the newly identified VS linkages had minor demographic mismatches, such as solitary differences in race, zip code, or one of the three birthdate components. These 62 closely matched death certificates were substituted for the death certificates originally linked. No death certificates with soft matching SSNs were found among the 2,530 PDDS records that had linked to a VS record using demographic variables alone.

In this report, all patients with a discharge disposition of death (08) on either their index AMI admission or a linked transfer record within 30 days of admission were counted as deaths, regardless whether a matching death certificate was found. Similarly, any death certificate matched by the algorithm was automatically accepted as valid if an index or subsequent record had a discharge disposition of death within 30 days of the index AMI admission. These decisions were based on the following observations:

1. Because the focus of this report is on hospital performance, it is reasonable to rely on hospital discharge abstracts to ascertain deaths that occurred before hospital discharge.

2. The overall accuracy of hospital-reported death exceeds 99.5 percent, according to OSHPD's 1996 AMI validation study and its 1992 study of coding quality for medical-surgical DRGs.

3. Special analyses were performed on the subset of 1993 hospital records with a discharge disposition of death. Two variables that were not used in the linkage process helped to validate matches among these patients: (1) date of hospital discharge reported on the PDDS file, which should equal the date of death reported on the death certificate; and (2) location of death reported on the death certificate, which should indicate inpatient hospital. Among perfect matches with the same SSN, birthdate, sex, race, and 5-digit zip code, 98.1 percent had the same reported date of death and 94.9 percent had a VS-reported location of inpatient hospital.

   Compared with this benchmark, nearly all of the soft matches showed comparable validity. For example, exact demographic matches with soft SSN matches had 97.9 percent agreement on the date of death and 94.6 percent agreement on the location of death. Exact SSN matches with zip mismatches had 96.6 percent agreement on the date of death and 93.1 percent agreement on the location of death.

   Only two types of soft matches seem potentially problematic: (1) exact SSN matches with mismatched or missing birthdates had 76.4 percent agreement using either validity criterion; and (2) exact SSN matches with sex mismatches had 87.0 percent and 78.9 percent agreement, respect-

ively, using the two criteria described above. Birthdate mismatches were discarded at the final step of the linkage algorithm, so the former problem is moot. Sex mismatches (and missing birthdate matches) in the 1993 AMI data set were manually reviewed. Without exception, the linked hospital and death records had compatible dates and ICD-9-CM diagnoses, confirming the validity of the matches and suggesting miscoding of sex.

4. According to the original data set used in this report, 93.0 percent of the AMI deaths reported by hospitals had a matching death certificate. If the index AMI record had a valid SSN, 96.5 percent had a matching VS record. If the index AMI record did not have a valid SSN, only 53.2 percent had a matching death certificate. Over 53 percent of reported but unlinked deaths were attributable to missing SSNs, and many of the remainder were concentrated among patients of "other" race, or at certain large hospitals. Several of these hospitals were contacted by OSHPD staff, and indeed confirmed that the patients in question had died.

Based on these analyses, there is little evidence of hospitals reporting AMI deaths that did not actually occur. Nearly all linked death certificates appear to be valid, using other variables that were not involved in the linkage algorithm. Most unlinked deaths were attributable to specific matching problems, such as the absence of an SSN on the hospital record or an especially unreliable racial category ("other" race).[10] Therefore, all vital statistics linkages established by OSHPD were accepted as valid if the patient's discharge disposition was reported as "death," either on the index AMI record or on any linked record within 30 days of the index AMI admission.

Among patients who were discharged alive (according to PDDS data), a higher standard for vital statistics linkage was necessary. The need for this higher standard was established through special analyses of two types of problematic matches:

1. In 28 cases between 1991 and 1993, the same AMI patient had two different death certificates linked to different hospital discharge abstracts. This is a serious problem because each individual can only have one death certificate. Eleven of these cases had a discharge abstract with a disposition of "death"; in each case, the death certificate linked to that PDDS record was a much better match (based on SSN) than the death certificate linked to other PDDS record(s) for the same patient. In the remaining 17 cases, there was no PDDS record with discharge

10. A minor exception to this statement should be noted. Of 936 unlinked AMI deaths in the final mortality analysis, 7 actually had a linked death certificate more than 30 days after the index AMI admission. Two of these linkages were probably incorrect because the SSN did not match; the correct death certificates for these patients could not be found. Manual review suggested that the VS death date was incorrect in two cases, because the PDDS records listed diagnoses and procedures consistent with death and the VS records matched perfectly. However, the discharge disposition of "death" was clearly incorrect in two cases, each of which had a subsequent PDDS record as well as a subsequent VS record (both of which matched perfectly). One case was a perfect match, but the reason for the death date discrepancy was unclear. Because there were only two or three suspected cases in which a hospital reported a death that did not occur, the data were not altered.

disposition of "death." However, 15 of these cases had one linking death certificate that was an exact SSN match and another that was a total mismatch. This problem was resolved by fixing the logical error in OSHPD's matching algorithm, but it suggests that VS matches may not be valid if the SSN is missing or totally mismatched. Each of the last two cases had two linking death certificates that were both mismatched by SSN.

2. In 22 cases between 1991 and 1993, an AMI patient had a hospital admission dated after a linking death certificate, or a hospital discharge dated more than one day after a linking death certificate. Three of these cases had a discharge abstract with a disposition of "death" which linked appropriately to a VS record; in each case, the subsequent hospital record was linked using only demographic variables (gender, birthdate, zip code) but fell outside the time window for hospital record linkage (see Step 4 under "Hospital Discharge Linkage"). In the remaining 19 cases, there was no hospital record with discharge disposition of "death." However, 17 of these cases had a linking death certificate that was a total mismatch by SSN. Each of the last two cases had a linking death certificate that was an exact SSN match but a subsequent PDDS record that was linked using demographic variables (and fell outside the time window for PDDS linkage).

These analyses suggest that in the absence of a hospital record with a discharge disposition of "death," death certificate matches may not be valid if the SSN is missing or totally mismatched. “Totally mismatched" SSNs were defined by exclusion, after partial SSN matches were identified. Partial SSN matches involved any seven of nine numerical positions or consecutive digits, any two of three hyphenated components, or the last four digits (which have the most variability within California, because they are unrelated to which Social Security office assigned the SSN). Note that transpositions and fixed additions or subtractions (e.g., add one to each SSN digit) were not considered in the definition of partial matches.

Additional analyses suggested that in the absence of a hospital record with a discharge disposition of "death," some VS matches may not be valid even when the SSN matches exactly. For example, at least 34 linked death certificates were exact SSN matches but gender mismatches. Of these 34 records, 29 were complete demographic mismatches and 5 were soft matches on birthdate. Most of these records probably represented spouses who shared SSNs with index AMI patients. As a result of these analyses, the following minimum criteria were adopted for vital statistics linkage among patients who were discharged alive (according to the hospital record):

1. gender had to match exactly; and

2. SSN had to match partially or exactly (as defined above); and

3. birthdate had to match partially or exactly (as defined in Step 2.1 under "Hospital Discharge Linkage"). A birthdate match was defined as partial if

at least two of three elements (i.e., month, day, last two digits of year) also matched, or if any one element matched and the other two were exactly transposed. This definition is slightly stricter than the definition applied by OSHPD to discard birthdate mismatches, but it is consistent with the definition that has been used to link PDDS records in the last three reports of the California Hospital Outcomes Project.

With these criteria, one can be nearly certain that if an AMI patient discharged alive from an acute-care hospital in California has a linking death certificate within 30 days of admission, that patient actually died after discharge.

Chapter 5

Definition of Outcome

Only one outcome of acute myocardial infarction was studied: death within 30 days of admission. In selecting this outcome, several statistical and clinical issues were considered. For example, death is an important and rather frequent outcome of AMI hospitalizations. Medical interventions, such as prompt administration of intravenous thrombolytics, can reduce the risk of early death after an AMI. In addition, two recent studies of OSHPD data have shown that death is reported reliably. These characteristics make it a useful outcome for analysis.

New in 1997

- This report uses death within 30 days, regardless of location, as the outcome for AMI patients. Previous reports only counted in-hospital deaths.

- This improvement was achieved by linking California hospital discharge abstracts to vital statistics records (death certificates). As a result, any potential bias in mortality rates due to differences in the length-of-stay pattern across hospitals has been eliminated.

Ascertainment of Deaths

Deaths within 30 days of admission were ascertained using two different data sources: linked hospital discharge abstracts and vital statistics records (death certificates). Hospital discharge abstracts only record deaths that occur in nonfederal acute care hospitals in California. By contrast, a death certificate is generated whenever a California resident dies, regardless where the death occurs. For the reasons described in Chapter Four, a death certificate cannot always be linked to previous hospital discharge abstracts for the same patient. Therefore, neither hospital discharge abstracts nor vital statistics records capture all deaths that occur within 30 days of an AMI. Among 119,863 fresh AMIs in California between January 1991 and November 1993, 13,976 (11.7 percent) were reported as 30-day deaths in both data sets, 936 (0.8 percent) were reported as inpatient, 30-day deaths on hospital discharge abstracts but were not confirmed by vital statistics data, and 2,501 (2.1 percent) had linked death certificates within 30 days but were discharged alive from the hospital. All 17,413 deaths ascertained from either or both data sources were counted in this study.

In-hospital deaths beyond 30 days were not counted because these late deaths may have resulted from social problems or unrelated illnesses. Not counting late deaths made the outcome comparisons across hospitals more

valid. Other cutoffs were considered but the 30-day limit was adopted because it is consistent with previous research in the field.

Attribution of Deaths

Because 19.5 percent of AMI patients were transferred from the hospital where they were initially admitted to another acute care facility, it was important to define an "episode of care" that included all inpatient treatment for a single AMI. The outcome of each "episode of care" was attributed to the hospital that originally admitted the patient. Attribution of outcomes to the initial hospital was an important and desirable feature of this study. Otherwise, hospitals that transferred many of their AMI patients to other facilities would have had relatively low risk-adjusted morality because some of their patients would have died elsewhere. Conversely, hospitals that neither transferred their own patients elsewhere nor accepted transfers would have had relatively high risk-adjusted mortality. These biases were avoided by attributing linked outcomes to the initial hospital. In addition, the risk of death is highest during the first 24 hours after an AMI and most of the key decisions that affect short-term mortality are made during this period. Determining whether, when, and where to transfer the patient is one of the most important of these decisions.

It appears that patients were often transferred for diagnostic evaluation or coronary revascularization. Of the patients admitted between July 31, 1990 and May 31, 1991, and then transferred elsewhere, 25 percent underwent coronary bypass grafting, 31 percent underwent coronary angioplasty, and 66 percent underwent cardiac catheterization at a subsequent hospital. Although some post-transfer deaths may be attributable to complications of surgery rather than complications of the initial AMI, these two types of complications cannot be distinguished using administrative data. Overall, patients who underwent either angioplasty or coronary bypass grafting had **lower** 30-day mortality than patients who underwent neither of these procedures.

Chapter 6

Selection and Inclusion of Hospitals

Certain hospitals may not be directly comparable with the great majority of hospitals caring for AMI patients in California. For example, non-acute care hospitals are not organized and staffed to treat patients with acute conditions. Any AMI records from such hospitals are probably either miscoded or represent atypical patients. In addition, the data received from several acute care hospitals had important limitations that precluded evaluating these facilities. This chapter describes the universe of hospitals eligible for study and the specific criteria used to exclude eligible hospitals.

New in 1997

- This report does not exclude California hospitals that transferred at least 20 percent of their AMI patients to nonreporting hospitals (e.g., federal hospitals or hospitals in other states). Previous reports excluded these hospitals because of the potential for underestimating their true inpatient mortality rates. This improvement was achieved by linking hospital discharge abstracts to vital statistics records (death certificates). As a result, nearly all deaths within 30 days of admission were ascertained, regardless where the death occurred.

- In this report, each hospital's data were reviewed for possible miscoding on an annual basis, as well as aggregated across all three years. Hospitals that improved their coding practices between 1991 and 1993 were included in the study for the period during which they submitted adequate data. Previous reports did not allow for the possibility of improved coding during the study period.

Hospitals Eligible for Study

The original study sample for the California Hospital Outcomes Project included cases from all non-federal acute care hospitals in California, as noted in Chapter Three. Hospitals operated by the US Department of Veterans Affairs or Department of Defense do not report data to OSHPD and therefore could not be included.

Many hospitals provide more than one type of care (e.g., acute care plus skilled nursing care or rehabilitation). Before January 1, 1995, these hospitals were encouraged but not required to submit separate bundles of abstracts, or reports, from each type of care. If a hospital failed to distinguish its acute care abstracts from its other abstracts, OSHPD assigned the same "type of care" to every discharge abstract from that hospital. (Beginning January 1, 1995, hospitals are required to distinguish the type of care on the discharge

abstract.) This assignment was based on the types of licensed units at the hospital and the proportion of records that fell into each Major Diagnostic Category. In 1993, 50 percent of acute care hospitals with rehabilitation units and 39 percent of acute care hospitals with skilled nursing or intermediate care units did not submit separate reports to OSHPD. Some patients in these skilled nursing or rehabilitation units may have experienced AMIs without being transferred to acute care. Such cases were inadvertently included in this study, whereas they would have been excluded if the type of care had been reported correctly. As a result, the AMI mortality statistics may be misleading for hospitals that provide multiple levels of care but fail to submit separate bundles of abstracts from each level.

Criteria for Excluding Hospitals

Although hospitals devote considerable effort to producing accurate discharge abstracts, the guidelines that professional coders follow when they abstract medical records are sometimes ambiguous and subject to multiple interpretations. Hospitals also face financial incentives that affect how diagnoses are coded, particularly for Medicare beneficiaries. As a result, the prevalence of various AMI risk factors is extremely variable across hospitals. Some hospitals reported these associated conditions on far fewer records than would be expected based on statewide prevalence data. If this variability reflects unusual documentation practices by physicians or coding practices by medical records personnel, it could seriously distort comparisons of risk-adjusted mortality across hospitals.

To avoid this problem, hospitals with the most unusual data related to important clinical risk factors were excluded. These exclusion criteria were applied to all linked records in a single "episode of care," because all such records were used to ascertain clinical risk factors. The criteria listed below were derived after reviewing the prevalence of every risk factor across hospitals, and considering possible reasons for excess variability. For example, the proportion of AMI patients with a history of coronary bypass surgery could vary widely because some hospitals specialize in treating complex patients. On the other hand, conditions such as hypertension and congestive heart failure should be distributed more evenly across hospitals.

Without reviewing individual medical records at excluded hospitals, it was impossible to tell whether these data were incorrect or simply reflected an unusual patient population or an unusual practice pattern. Written comments submitted by several hospitals that were excluded from the 1993 analyses, and a subsequent survey of these excluded hospitals, support both explanations. As more is learned about why certain hospitals have unusual patterns of data, it may be possible to include them in future studies.

There are two basic ways to identify hospitals with unusual patterns of data. First, a fixed cutoff could be applied based on clinical considerations or face validity; all hospitals at which the reported prevalence of a risk factor is below (or above) that level would be excluded. Second, a probability cutoff could be

applied, based on the statistical significance of the difference in the reported prevalence of a risk factor between one hospital and the statewide average. All hospitals at which the true prevalence of a risk factor is extremely unlikely to be the same as the statewide average would be excluded. To minimize the number of excluded hospitals, a set of criteria were developed that included both fixed and probability cutoffs.

The probability cutoffs were designed so there would be only a 5 percent chance of excluding one or more hospitals statewide, under the assumption that all hospitals had the same true prevalence of the risk factors of interest. This procedure is known as a correction for multiple comparisons. Because 431 California hospitals admitted AMI patients during the 1991-93 study period, the probability that any specific hospital was excluded based on its reporting of key risk factors was much smaller than the 5 percent chance that one or more hospitals statewide was excluded. Specifically, the exact probability that a specific hospital was excluded by chance, using a one-tailed test, was $p < 0.000029752$ (or approximately 3 in 100,000).

Probability cutoffs identify hospitals where the prevalence of a risk factor was very significantly different from the statewide average, in a statistical sense. However, they do not address the clinical plausibility of such differences. For this reason, fixed prevalence cutoffs were also established. Hospitals were excluded only if they exceeded **both** the probability cutoff and the fixed prevalence cutoff for a risk factor. These prevalence cutoffs represent the limits of clinical plausibility, based on literature review and discussion with specialists in the field. They were confirmed and slightly adjusted based on the empirical distribution of prevalences across hospitals. For example, the lower cutoff for congestive heart failure was set at 17 percent because the five hospitals with the lowest reported prevalences had values of 8.6 percent, 17.9 percent, 19.5 percent, 20.5 percent, and 21.4 percent (aggregated over three years). The upper cutoff for other and unspecified infarct site was set at 28 percent because the distribution of prevalences demonstrated a clear break between 25 percent and 30 percent.

Risk Factor	*Direction*	*Prevalence Cutoff*	*State Prevalence*
Subendocardial site	undercoded	12.0%	32.1%
Hypertension	undercoded	14.0%	37.8%
Other/unspecified site	overcoded	28.0%	8.1%
Congestive heart failure	undercoded	17.0%	34.7%

The combined effect of these criteria was to exclude 27 hospitals that admitted 2,127 AMI patients in 1991, 17 hospitals that admitted 1,068 AMI patients in 1992, and 13 hospitals that admitted 494 AMI patients in 1993. Overall, 3,689 of 119,863 AMI patients (3.1 percent) were excluded because of hospital data problems. These figures represent a significant decline from the 35 hospitals and 4.8 percent of cases that were excluded from last year's study. Table 6.1 lists the specific hospitals excluded from this AMI study and gives the specific reason for the exclusion. Table 6.2 lists the specific hospitals that were included in the study, but did not have eligible AMI cases in at least one of the three study years.

Table 6.1: Hospitals excluded from AMI models in one or more study years

Hospital	*County*	*1991*		*1992*		*1993*	
		Cases	*Cause*	*Cases*	*Cause*	*Cases*	*Cause*
Alameda Hospital	Alameda	93	b				
Feather River Hospital	Butte	98	a,c				
Brookside Hospital	Contra Costa	138	a,c				
Coalinga Regional Medical Center	Fresno	6	c	16	c	18	c
Selma District Hospital	Fresno	51	c	42	c	37	c
Glenn Medical Center	Glenn	18	c	6	c	6	c
St. Joseph Hospital-Eureka	Humboldt	70	b				
El Centro Reg Medical Center	Imperial	74	a				
East L A Doctors Hospital	Los Angeles	21	c				
L A Comm Hospital of Norwalk	Los Angeles	31	a	19	a	23	a
LA Co/Harbor-UCLA Med Cen	Los Angeles	107	a,c	125	a,c	101	a,c
L.A Co/Olive View-UCLA Med Cen	Los Angeles	43	a	62	a	51	a
George L. Mee Memorial Hospital	Monterey	12	c	16	c	18	c
Salinas Valley Memorial Hospital	Monterey	124	d				
Barstow Community Hospital	San Bernardino	67	a	69	a	56	a
Needles-Desert Communities Hosp	San Bernardino	37	c	28	c	31	c
Villa View Community Hospital	San Diego	14	a,c	19	a,c	9	a,c
Sharp Chula Vista Medical Center	San Diego	103	b				
CA Pacific Med Cen-CA Campus	San Francisco	112	a	131	a		
St. Francis Memorial Hospital	San Francisco	96	a	146	a,c		
St. Luke's Hospital	San Francisco	127	c	106	c	110	c
Medical Center at U.C.S.F.	San Francisco	122	c	95	a		
Columbia Good Samaritan Hospital	Santa Clara	216	d				
Redding Medical Center	Shasta	170	b	149	d		
Siskiyou General Hospital	Siskiyou	30	a,c,d	25	a,c,d	26	a,c,d
Trinity Hospital	Trinity	16	c	14	c	8	c
Rideout Memorial Hospital	Yuba	131	a				

Cause for exclusion
a: Subendocardial site of infarction possibly undercoded
b: Hypertension possibly undercoded
c: Other or unspecified site of infarction possibly overcoded
d: Congestive heart failure possibly undercoded

Table 6.2: Number of cases in hospitals reporting no AMI's in one or two study years

Hospital	*County*	*1991*	*1992*	*1993*
Vencor Hospital-San Leandro	Alameda	33	0	0
Kingsburg District Hospital	Fresno	1	0	1
Tehachapi Hospital	Kern	1	4	0
Avenal District Hospital	Kings	6	4	0
Avalon Municipal Hospital and Clinic	Los Angeles	1	0	1
Barlow Respiratory Hospital	Los Angeles	1	0	0
Dominguez Medical Center	Los Angeles	7	0	0
Valley Hospital-Pomona	Los Angeles	1	0	0
Specialty Hospital of Southern Caliifornia	Los Angeles	46	11	0
Lincoln Hospital Medical Center	Los Angeles	4	0	4
Beverly Hills Medical Center	Los Angeles	24	1	0
Mission Community Hospital-Panorama	Los Angeles	36	0	10
Pico Rivera Medical Center	Los Angeles	15	2	0
Mission Community Hospital-San Fernando	Los Angeles	9	9	0
L.A. County/Rancho Los Amigos Hospital	Los Angeles	1	5	0
USC-University Hospital	Los Angeles	0	3	3
Chowchilla District Memorial Hospital	Madera	0	1	1
Surprise Valley Healthcare District	Modoc	0	1	1
College HospitalCosta Mesa	Orange	22	0	0
Vencor Hospital-Orange County	Orange	117	81	0
Lakeside Hospital	Riverside	0	5	5
Sharp Healthcare Murietta	Riverside	0	46	57
San Diego General Hospital	San Diego	1	0	0
Vencor Hospital-San Diego	San Diego	36	23	0
UCSD / Thornton Hospital	San Diego	0	0	12
Seton Medical Center-Coastside	San Mateo	1	0	0
Del Puerto Hospital	Stanislaus	1	0	0
Memorial Hospital at Exeter	Tulare	5	1	0

Chapter 7

Definitions and Prevalence of Risk Factors

In this study, risk factors were defined as characteristics or conditions that probably existed at the time of admission and may have influenced patient outcomes. Three sets of risk factors were examined.

The first set includes demographic characteristics such as sex, race, and age. The second set includes hospitalization characteristics such as the source and type of admission. The third set represents clinical characteristics such as diabetes and cancer. These clinical factors include both chronic illnesses and conditions or procedures associated with the principal diagnosis (e.g., the portion of the heart involved in an AMI). All clinical risk factors were based on the diagnoses and procedures listed on discharge abstracts and coded using ICD-9-CM. Each patient discharge abstract includes a principal diagnosis and principal procedure, plus as many as 24 other diagnoses and as many as 24 other procedures.

New in 1997

- This report includes minor changes in the definitions of several risk factors. These changes reflect the advent of new ICD-9-CM codes or new national guidelines related to existing codes, the availability of extra information from previous hospitalizations (up to 6 months before the index AMI), and further analyses intended to create more homogeneous clusters of comorbid diagnoses. These changes had little overall impact on the study methods and results.

- In this report, numerous risk factors were deemed ineligible for use in risk-adjustment models, even though they had clinically meaningful and statistically significant effects on mortality. These risk factors were either not reliably coded by California hospitals, according to OSHPD's AMI validation study, or showed implausible variation in prevalence from year to year. Previous reports included these poorly coded risk factors and were therefore more susceptible to coding-related biases.

Demographic and Hospitalization Characteristics

The demographic variables available from patient discharge abstracts are sex, race, and age. Table 7.1 summarizes these characteristics of the AMI sample. Each of these three variables was tested in risk-adjustment models, as described in Chapter Eight. For analytic purposes, race was aggregated into four categories: White, African-American, Hispanic, and other.

Several measures describing the hospitalization were available from patient discharge abstracts: expected principal source of payment, source of admission, type of admission, and disposition. The first three of these variables were tested in risk-adjustment models, as described in Chapter Eight. The expected source of payment was used as a crude indicator of socioeconomic status. The source of admission may help distinguish critically ill patients who are admitted through an emergency room from more stable patients who are admitted directly from a physician's office. The type of admission reflects whether a patient was sick enough to require admission to an intensive care unit. The large number of categories for expected source of payment and source of admission were aggregated into a smaller number of categories for analytic purposes. Table 7.2 summarizes the hospitalization characteristics of the AMI sample.

Criteria for Selecting Clinical Risk Factors

After reviewing the recent medical literature and obtaining the assistance of a clinical advisory panel, a list of potential clinical risk factors for death after AMI was developed. These potential clinical risk factors are listed in Chapter Two and defined in this chapter.

Potential clinical risk factors for death after AMI were adapted to the *International Classification of Diseases, Ninth Revision, Clinical Modification* (ICD-9-CM) by reviewing all volumes of ICD-9-CM; the American Hospital Association's *ICD-9-CM Coding Handbook, 1991 Revised Edition*; *Coding Clinic for ICD-9-CM*; OSHPD's *Discharge Data Review*; and other publications for coding professionals. These adaptations were reviewed by two coding experts. Finally, the number of cases and the mortality rate associated with each five-digit ICD-9-CM diagnosis were examined to ensure that no potential clinical risk factors had been omitted. During this process, many potential clinical risk factors were redefined to capture differences in risk more precisely. The following criteria were used to select potential clinical risk factors:

1. **Prevalence.**
   Extremely rare conditions (e.g., less than 0.1 percent prevalence) were not considered as potential clinical risk factors, because it would have been impossible to estimate their contribution to the risk of death. Some moderately rare conditions were considered as potential clinical risk factors but were dropped or aggregated with other risk factors during the model development process (Chapter Eight).

2. **Ability to define using ICD-9-CM.**
   Potential clinical risk factors without corresponding ICD-9-CM codes were not included because they could not be identified from hospital discharge abstracts.

3. **Confidence that the condition was likely to have been present when the patient was admitted to the hospital.**

Conditions likely to have developed after admission, such as iatrogenic infections, were not considered as potential clinical risk factors. However, it was not always clear whether a condition was "likely to have been present when the patient was admitted" or "likely to have developed after admission." Conditions that could have developed either before or after admission were placed in a special category for further examination (as described in Chapter Eight).

4. **Clinical importance.**
   Conditions with minimal clinical significance, such as skin rashes, were removed from the list of potential risk factors. During the model development process (Chapter Eight), potential risk factors that were not statistically associated with 30-day mortality were also identified and removed (even if they had some clinical significance).

Timing of Clinical Risk Factors

The timing of diagnoses is a critical issue in risk-adjusting hospital outcomes. Any acute or chronic condition diagnosed either at or before admission may be used in risk-adjustment because it reflects severity-of-illness at admission. Any chronic condition diagnosed after admission may also be used in risk-adjustment because it was presumably present, albeit undetected, at admission. By contrast, acute conditions diagnosed after admission are problematic because they may reflect quality of care. Some complications of AMI are potentially preventable with prompt and aggressive treatment, including aspirin, thrombolytic agents, and coronary revascularization if necessary. If one treats these conditions as risk factors by including them in risk-adjustment models, one may inappropriately give hospitals credit when they fail to prevent complications. Until 1996, California hospital discharge abstracts did not include any information on the timing of diagnoses. Therefore, any acute condition could be either a comorbidity (e.g., present at admission) or a complication of care (e.g., present only after admission).

This dilemma was resolved by developing two different models to adjust for differences in patient characteristics across hospitals. Model A is a more conservative model that includes fewer risk factors; Model B is a more comprehensive model that includes several additional risk factors. The risk factors in Model A almost certainly represent comorbidities -- clinical or personal characteristics that were present when the patient entered the hospital. Model B includes all of the risk factors in Model A plus certain demographic variables (e.g., race, source of admission, expected principal source of payment) and clinical characteristics with unclear timing (e.g., shock, pulmonary edema). By comparing the results from Models A and B, one can assess the sensitivity of hospital-specific risk-adjusted mortality rates to assumptions about the timing of acute conditions. The development of these models is further described in Chapter Eight.

The presence or absence of each risk factor was determined after linking serial hospitalizations for AMI patients who were transferred from one

hospital to another. The discharge diagnoses from all hospitals involved in the episode of care were combined into a single list. Thus, a patient who was transferred from Hospital X to Hospital Y but only had hypertension coded in Hospital Y was classified as hypertensive in the analysis of Hospital X's AMI outcomes. Hospitals thereby received credit for clinical risk factors that they might not have had the opportunity to discover or document before transfer. Many inter-hospital transfers occur so quickly that the initial hospital cannot complete its diagnostic evaluation.

During the 6 months before the date of an index AMI admission, 17.1 percent of cases had one or more prior hospitalizations. Among these cases, prior discharge abstracts provided additional information about the presence and timing of clinical risk factors. If a diagnosis was noted on a prior discharge abstract, then it clearly preceded the AMI of interest. For this reason, clinical risk factors were defined somewhat differently according to whether there were any prior hospitalizations. The term "index AMI hospitalization," as used below, includes linked transfer hospitalizations within the same episode of care.

1. Risk factors that represent chronic diseases, such as hypertension and diabetes, were identified from either the index AMI hospitalization or prior hospitalizations. If there were no prior hospitalizations, then the index record alone was used to identify these risk factors. These risk factors were tested in both Model A and Model B.

2. Risk factors that represent fundamental characteristics of an AMI, such as the infarct site, were identified exclusively from the index AMI hospitalization. These risk factors were tested in both Model A and Model B.

3. Risk factors that represent acute complications of an AMI, such as shock, were identified exclusively from the index AMI hospitalization. The presence of one of these complications in association with a previous AMI (or other acute illness requiring hospitalization) has little bearing on the risk of death after the index AMI. These risk factors were tested only in Model B.

4. Risk factors that could represent either chronic comorbidities or acute complications, such as mitral regurgitation, were first identified only from prior hospitalizations (to be certain that they were actually preexisting conditions). These risk factors were tested in both Model A and Model B, but only among cases with one or more prior hospitalizations. The same risk factors were then identified only from index AMI hospitalizations (where complications and comorbidities could not be distinguished). These revised risk factors were tested only in Model B.

5. In some cases, the two versions of risk factors created using method #4 were combined into one version that incorporated information from both prior and index hospitalizations. This was done in two situations: (1) when the risk factor was deemed much more likely to be a chronic comorbidity

than an acute complication (e.g., congestive heart failure), based on OSHPD's AMI validation study; or (2) when the statistical effect of the risk factor on mortality was unrelated to its timing. In the former situation, the risk factor was tested in both Model A and Model B. In the latter situation, the risk factor was tested only in Model B.

New Clinical Risk Factors

Four new risk factors were tested in preparing this year's report: coma (COMAI), catastrophic sequelae of AMI (AMISEQUI), fracture (FRACTURI), and ischemic bowel or liver (VASINSUI). All of these risk factors represent acute complications of an AMI, and were therefore identified only from index AMI hospitalizations. Coma and ischemic bowel or liver are manifestations of decreased perfusion of vital organs, resulting from poor cardiac output. Catastrophic sequelae are major structural failures involving the heart muscle; the code for congenital ventricular septal defects is included here because of suspected miscoding. A fracture may represent either a marker of poor health status (e.g., a patient with cancer and pathologic fractures), a comorbid factor (e.g., a patient who fell during an AMI and suffered a fracture), or a complication of care (e.g., a patient who fell out of bed in the hospital). Because these four new risk factors may have developed either before or after the date and time of admission, they were tested only in Model B.

Definitions of Clinical Risk Factors

Table 7.3 shows the definitions of all clinical risk factors used in any of the final risk-adjustment models for AMI mortality. Table 7.4 shows the prevalence of these risk factors in the sample of patients with one or more prior admissions. Table 7.5 shows the prevalence of these risk factors in the sample of patients with no prior admissions. Table 7.6 shows additional risk factors evaluated but ultimately not used in the AMI risk-adjustment models, for reasons described in Chapter Eight.

Some risk factors have been redefined or dropped since last year's report. A couple of these changes reflect the advent of new ICD-9-CM codes or national coding guidelines. For example, *Coding Clinic for ICD-9-CM* recently instructed coders not to use hypotension (458.9) in AMI patients. Because of this new guideline, hypotension was dropped from the list of potential clinical risk factors.

Several risk factors were redefined in this year's report because data from prior hospitalizations were available for a 6 month period rather than the limited 8 week period used in earlier reports. For example, the definition of chronic renal failure (CHRRENAB) was expanded to include dialysis access or revision procedures performed during prior hospitalizations, or complications thereof during either prior or index hospitalizations. The definition of cardiac pacemaker (PRPACE) was expanded to include insertion

procedures performed during prior hospitalizations, unless the pacemaker was removed before the index AMI admission. The definition of old myocardial infarction (OLDAMIB) was expanded to include AMI diagnosis codes (410.xx) documented during prior hospitalizations. Finally, the definition of other cerebrovascular disease (OTHCVAP) was expanded to include a wider variety of manifestations of cerebrovascular disease, including transient ischemic attacks, during prior hospitalizations. Lengthening the period for ascertaining data from prior hospitalizations led to eliminating two risk factors that often respond quickly to treatment and might therefore have resolved by the time of the index AMI: acute ulcer disease (ACULCERP) and anemia (ANEMNOSP).

Most changes in the definitions of risk factors resulted from reexamining the unadjusted associations between specific ICD-9-CM codes and mortality. Special analyses were performed to ensure that each of the ICD-9-CM codes used to define a risk factor had a similar effect on the risk of death. For example, cancers of the colon, rectum, nasal cavities, and larynx were moved from the high-risk (HRSECMAB) to the low-risk (LRPMALIB) category, making the cancer diagnoses assigned to each risk factor more homogeneous. "Oliguria and anuria" was added to the definition of acute renal failure (ACRENALB) because of evidence that it is frequently used to describe the same clinical situation, and has a similar effect on mortality. "Arthropathy, unspecified" was added to the definition of osteoarthritis (OSTARTHB). Other risk factors were redefined by removing uncommon ICD-9-CM diagnoses with marginal clinical significance. For example, viral pneumonia and influenza with pneumonia were removed from the definition of pneumonia (PNEUMONI). Bacteremia was removed from the definition of sepsis (SEPSISI). Essential tremor and benign dystonias such as torticollis and writer's cramp were removed from the definition of central nervous system diseases (CNSDISB). Coagulopathy (COAGULI) was redefined to focus on acute coagulopathies rather than congenital factor deficiencies, which have less impact on short-term mortality. Atherosclerosis (ATHEROSB) was redefined to focus on peripheral vascular disease and aortic aneurysms; aneurysms involving other arteries and non-atherosclerotic arterial strictures were removed from the definition.

Clinical Risk Factors Dropped Due to Suspected Miscoding

Numerous risk factors were deemed ineligible for use in this year's risk-adjustment models, even though some had demonstrated clinically meaningful and statistically significant effects on mortality in previous reports. Most of these risk factors were not reliably coded by California hospitals, according to OSHPD's AMI validation study. Detailed results from this validation study were presented in the 1996 AMI report. Table 7.7 lists and defines the specific risk factors that failed to meet minimum standards for coding reliability. Note that some of these definitions differ from those used in previous years, because attempts were made to maximize the reliability of each risk factor by focusing on the most precise ICD-9-CM codes and incorporating recent coding guidelines. All of the risk factors in Table 7.7 had

weighted coding sensitivities, using reabstracted records as a gold standard, of less than 30 percent **and** weighted kappa statistics less than 0.4.

Two additional risk factors were deemed ineligible for use in risk-adjustment models because of implausible variation in prevalence from year to year. The prevalence of intermediate coronary syndrome, also known as unstable angina (INCORSYI), decreased from 8.5 percent to 5.0 percent between 1991 and 1993. Conversely, the prevalence of hyperlipidemia (HYPERLIB) increased from 10.6 percent to 14.9 percent during this period. These temporal trends have no clinical explanation and suggest changes in diagnostic or coding practices. Intermediate coronary syndrome also demonstrated poor specificity and positive predictive value in OSHPD's AMI validation study. This finding confirmed that unstable angina is not uniformly defined in California hospitals. Uncomplicated hypertension and subendocardial site also demonstrated consistent increases in prevalence between 1991 and 1993, from 35.4 percent to 40.2 percent and from 30.2 percent to 34.0 percent, respectively. These variables were retained because hospitals that clearly underreported these risk factors in 1991 and 1992 were excluded from the analysis, as described in Chapter Six.

It was important to drop poorly coded clinical risk factors from the risk-adjustment models, because unreliable coding creates a problem known to epidemiologists as information bias. In other words, inaccurate information about the presence or absence of risk factors leads to incorrect estimates of their effect on mortality. If the regression models in this report were biased, the estimated probabilities generated by the models would also be biased. The resulting classifications of hospital mortality could be adversely affected. In practice, however, removing these poorly coded variables had little impact on the performance of the risk-adjustment models and the hospital-specific results.

Table 7.1: Demographic characteristics of acute myocardial infarction cases (after exclusions)

Characteristic	*Number*	*Percent*
Total	116,174	100.0
Sex		
Male	73,330	63.1
Female	42,844	36.9
Race		
White	90,997	78.3
Black	7,072	6.1
Hispanic	10,836	9.3
Native American	254	0.2
Asian	5,384	4.6
Other	1,073	0.9
Missing/Unknown	558	0.5
Age		
Mean	67.3	
Std Dev	13.3	

Table 7.2: Hospitalization characteristics of acute myocardial infarction cases (after exclusions)

Characteristic	*Number*	*Percent*
Total	116,174	100.0
Admission Type		
Emergency	70,200	60.4
Urgent	43,162	37.2
Elective	2,754	2.4
Missing/unknown	58	0.0
Admission Source		
Routine	16,551	14.2
Emergency Room	98,976	85.2
Other Facility	463	0.4
Home Health	73	0.1
Other	111	0.1
Payment Source		
Medicare	60,487	52.1
Medi Cal	7,027	6.0
Worker's Comp	429	0.4
Title V	1	0.0
Other Government	806	0.7
Blue C/S	2,773	2.4
Insurance Co	13,498	11.6
HMO/PHP	23,913	20.6
Self Pay	5,207	4.5
No Charge	55	0.0
Other Non Govt	367	0.3
Section 17000	1,610	1.4
Missing/unknown	1	0.0
Discharge Disposition		
Routine	61,989	53.4
Acute Hospital	24,728	21.3
Intermediate Care Facility	131	0.1
Skilled Nursing Facility	4,760	4.1
Other Facility	1,467	1.3
Left Against Med Advice	1,180	1.0
Home Health	8,311	7.2
Died	13,608	11.7

Table 7.3: ICD-9-CM codes for clinical risk factors for death after acute myocardial infarction

Code	*ICD-9-CM Description*	*Source of Data**
	Aspiration pneumonia (ASPPNEUI)	Index only[1]
507.0	Aspiration pneumonia	
	Catastrophic sequelae of AMI (AMISEQUI)	Index only[1]
429.5	Rupture of chordae tendineae	
429.6	Rupture of papillary muscle	
429.71	Acquired cardiac septal defect	
745.4	Ventricular septal defect	
	Central nervous system disease (CNSDISB)	Index or prior
331.1-331.9	Other cerebral degenerations (except Alzheimer's disease)	
332.x	Parkinson's disease	
333.0	Other degenerative diseases of the basal ganglia	
333.2	Myoclonus	
333.3	Tics of organic origin	
333.4	Huntington's chorea	
333.5	Other choreas	
333.6	Idiopathic torsion dystonia	
333.7	Symptomatic torsion dystonia	
340	Multiple sclerosis	
341.x	Other demyelinating diseases of central nervous system	
344.x	Other paralytic syndromes	
	Cerebrovascular disease, other (OTHCVAI)	Index only[1]
430	Subarachnoid hemorrhage	
431	Intracerebral hemorrhage	
432.x	Other and unspecified intracranial hemorrhage	
434.x	Occlusion of cerebral arteries	
436	Acute but ill-defined cerebrovascular disease	
437.1	Other generalized ischemic cerebrovascular disease	
	Coma (COMAI)	Index only[1]
780.0x	Alteration of consciousness	
250.2x	Diabetes with hyperosmolarity (hyperosmolar coma)	
250.3x	Diabetes with other coma	
572.2	Hepatic coma	
	Complete atrioventricular block (COATRBLI)	Index only[1]
426.0	Complete atrioventricular block	
	Congestive heart failure (CHFB)	Index or prior
425.x	Cardiomyopathy	
428.x	Heart failure	
	Diabetes, complicated (DBTCMPB)	Index or prior
250.1x-250.9x	Diabetes with mention of complication	
357.2	Polyneuropathy in diabetes	
362.0x	Diabetic retinopathy	
	High-risk or secondary malignant neoplasm (HRSECMAB)	Index or prior
141.x-152.x	Malignant neoplasm of oral cavity, pharynx, esophagus, stomach, small intestine	

Table 7.3: ICD-9-CM codes for clinical risk factors for death after acute myocardial infarction, *continued*

Code	*ICD-9-CM Description*	*Source of Data**
155.x-159.x	Malignant neoplasm of liver, gall bladder, pancreas, peritoneum	
162.x-171.x	Malignant neoplasm of lung, pleura, heart, thorax, bone, connective tissue	
196.x-199.x	Secondary malignant neoplasm	
	Hypertension (HTB)	Index or prior[2]
401.x	Essential hypertension	
402.x0	Hypertensive heart disease	
403.x0	Hypertensive renal disease	
404.x0	Hypertensive heart and renal disease	
405.xx	Secondary hypertension	
	Infarction site, anterior wall (SITE_ANT)	Index only[3]
410.0x	Anterior wall	
410.1x	Other anterior wall	
410.2x	Inferolateral	
410.5x	Other lateral	
	Infarction site, inferior wall (SITE_INF)	Index only[4]
410.3x	Inferoposterior wall	
410.4x	Other inferior wall	
410.6x	Posterior wall	
	Infarction site, other (SITE_OI)	Index only[5]
410.8x	Other unspecified sites	
410.9x	Unspecified sites	
	Infarction site, subendocardial (SUBENDOI)	Index only
410.7x	Subendocardial	
	Ischemic bowel or liver (VASINSUI)	Index only[1]
557.x	Vascular insufficiency of intestine	
570	Acute and subacute necrosis of liver	
	Paroxysmal ventricular tachycardia (PVENTACI)	Index only[1]
427.1	Paroxysmal ventricular tachycardia	
	Prior coronary artery bypass graft (PRCABG)	
996.03	Mechanical complication due to coronary bypass graft	Index or prior[6]
V45.81	Aortocoronary bypass status	Index or prior[7]
36.1x	Bypass anastomosis for heart revascularization	Prior only
	Pulmonary edema (PULEDEMI)	Index only[1]
514	Pulmonary congestion and hypostasis	
518.4	Acute edema of lung, unspecified	
518.5	Pulmonary insufficiency following trauma and surgery	
518.81	Respiratory failure	
518.82	Other pulmonary insufficiency, not elsewhere classified	
	Renal failure, acute or unspecified (ACRENALI)	Index only[1,8]
584.x	Acute renal failure	
586	Renal failure, unspecified	
788.5	Oliguria and anuria	

Table 7.3: ICD-9-CM codes for clinical risk factors for death after acute myocardial infarction, *continued*

Code	*ICD-9-CM Description*	*Source of Data**
	Renal failure, chronic (CHRRENAB)	
585	Chronic renal failure	Index or prior
403.x1	Hypertensive renal disease (malignant, benign, or unspecified), with renal failure	Index or prior
404.x2	Hypertensive heart and renal disease (malignant, benign, or unspecified), with renal failure	Index or prior
404.x3	Hypertensive heart and renal disease (malignant, benign, or unspecified), with congestive heart and renal failure	Index or prior
996.73	Other complications due to renal dialysis device, implant, and graft	Index or prior[9]
39.27	Arteriovenostomy for renal dialysis	Prior only
39.42	Revision of arteriovenous shunt for renal dialysis	Index or prior[10]
39.93	Insertion of vessel-to-vessel cannula	Prior only
39.94	Replacement of vessel-to-vessel cannula	Index or prior[10]
V45.1	Renal dialysis status	Index or prior
	Seizure disorder (EPILEPB)	Index or prior[1]
345.xx	Epilepsy	
780.3	Convulsions	
	Sepsis (SEPSISI)	Index only[1]
038.xx	Sepsis	
112.5	Disseminated candidiasis	
	Shock (SHOCKI)	Index only[1]
785.5x	Shock without mention of trauma	
	Skin ulcer (SKNULCRP)	Prior only
707.x	Chronic skin ulcer	
	Thyroid disease (THYROIDB)	Index or prior
243.x-244.x	Hypothyroidism	

* **Index only:** variable ascertained only from index AMI hospitalizations (including linked hospitalizations when patients were transferred from one facility to another). These variables represent acute complications of AMI that may be important for risk-adjustment if present at admission.
Prior only: variable ascertained only from prior hospitalizations. These variables represent conditions that may be either acute or chronic, so they are counted as risk factors only if they were present during a prior admission.
Index or prior: variable ascertained from either index or prior hospitalizations. These variables represent conditions that are very unlikely to occur acutely, and therefore almost certainly represent comorbidities.

1. These conditions may represent complications of hospital care rather than comorbidities or pre-existing diagnoses. They were therefore included only in Model B, which was specifically designed to adjust for clinical conditions that could have arisen after a patient was admitted to the hospital.
2. If HTHRTFB = 0 in Table 7.6 and no diagnoses indicative of hypertensive renal failure are present (403.x1, 404.x2, or 404.x3).
3. If no diagnoses indicative of subendocardial site are present.
4. If no diagnoses indicative of subendocardial or anterior site are present.
5. If no diagnoses indicative of subendocardial, anterior, or inferior site are present.
6. unless 36.1x occurred on the same or prior admission during the same series of transfer hospitalizations
7. unless 36.1x occurred on a prior admission during the same series of transfer hospitalizations
8. If CHRRENAB = 0
9. Unless 39.27 or 39.93 occurred during the same or a prior admission in the same series of transfer hospitalizations.
10. Unless 39.27 or 39.93 occurred on the same or a prior day during the same series of transfer hospitalizations.

Table 7.4: Clinical characteristics of AMI patients with one or more prior admissions (N=19,882)*

Characteristic	*Number*	*Percent*
Aspiration pneumonia (ASPPNEUI)	261	1.3
Central nervous system disease (CNSDISB)	469	2.4
Cerebrovascular disease, other (OTHCVAI)	463	2.3
Coma (COMAI)	218	1.1
Congestive heart failure (CHFB)	10,949	55.1
Diabetes, complicated (DBTCMPB)	3,714	18.7
High-risk or secondary malignant neoplasm (HRSECMAB)	665	3.3
Hypertension (HTB)	9,157	46.1
Infarction site, anterior wall (SITE_ANT)	5,422	27.3
Infarction site, inferior wall (SITE_INF)	3,933	19.8
Infarction site, other (SITE_OI)	2,349	11.8
Paroxysmal ventricular tachycardia (PVENTACI)	1,547	7.8
Prior coronary artery bypass graft (PRCABG)	3,385	17.0
Pulmonary edema (PULEDEMI)	1,605	8.1
Renal failure, acute or unspecified (ACRENALI)	742	3.7
Renal failure, chronic (CHRRENAB)	2,296	11.5
Seizure disorder (EPILEPB)	688	3.5
Sepsis (SEPSISI)	367	1.8
Shock (SHOCKI)	1,293	6.5
Skin ulcer (SKNULCRP)	470	2.4

* Characteristics in this table were ascertained from either index admissions or prior admissions or both, as noted in Table 7.3. Only variables included in either of the final risk-adjustment models are shown.

Table 7.5: Clinical characteristics of AMI patients with no prior admissions (N=96,292)*

Characteristic	*Number*	*Percent*
AMI sequela (AMISEQUI)	240	0.2
Aspiration pneumonia (ASPPNEUI)	1,165	1.2
Central nervous system disease (CNSDISB)	894	0.9
Cerebrovascular disease, other (OTHCVAI)	1,884	2.0
Coma (COMAI)	827	0.9
Complete atrioventricular block (COATRIBLI)	3,101	3.2
Congestive heart failure (CHFB)	29,484	30.6
Diabetes, complicated (DBTCMPB)	7,054	7.3
High-risk or secondary malignant neoplasm (HRSECMAB)	649	0.7
Hypertension (HTB)	35,016	36.4
Infarction site, anterior wall (SITE_ANT)	32,636	33.9
Infarction site, inferior wall (SITE_INF)	27,399	28.5
Infarction site, other (SITE_OI)	6,611	6.9
Ischemic bowel or liver (VASINSUI)	236	0.2
Paroxysmal ventricular tachycardia (PVENTACI)	8,083	8.4
Prior coronary artery bypass graft (PRCABG)	7,159	7.4
Pulmonary edema (PULEDEMI)	5,578	5.8
Renal failure, acute or unspecified (ACRENALI)	2,760	2.9
Renal failure, chronic (CHRRENAB)	2,566	2.7
Seizure disorder (EPILEPB)	1,378	1.4
Shock (SHOCKI)	5,311	5.5
Thyroid disease (THYROIDB)	3,212	3.3

* Characteristics in this table were ascertained from either index admissions or prior admissions or both, as noted in Table 7.3. Only variables included in either of the final risk-adjustment models are shown.

Table 7.6: ICD-9-CM codes for clinical risk factors evaluated but not included in final AMI models

Code	*ICD-9-CM Description*	*Source of Data**
	Arterial embolism (ARTEREMI)	Index only[1]
	(ARTEREMP)	Prior only
444.xx	Arterial embolism and thrombosis	
785.4	Gangrene	
	Atrial fibrillation (ATRFIBB)	Index or prior[1]
	(ATRFIBP)	Prior only
427.31	Atrial fibrillation	
	Atrioventricular block, first degree (DEG1AVBB)	Index or prior[1]
426.11	First degree atrioventricular block	
	Atrioventricular block, second degree (DEG2AVBB)	Index or prior[1]
	(DEG2AVBP)	Prior only
426.10	Atrioventricular block, unspecified	
426.12	Mobitz (Type) II atrioventricular block	
426.13	Other second degree atrioventricular block	
	Cardiac pacemaker (PRPACE)	
996.01	Mechanical complication due to cardiac pacemaker	Index or prior [3]
V45.0	Cardiac pacemaker in situ	Index or prior [4]
V53.3	Cardiac pacemaker reprogramming	Index or prior[4]
37.70	Initial insertion of lead (electrode), not elsewhere specified	Prior only[5]
37.71	Initial insertion of transvenous lead (electrode) into ventricle	Prior only[5]
37.72	Initial insertion of transvenous lead (electrode) into atrium and ventricle	Prior only[5]
37.73	Initial insertion of transvenous lead (electrode) into atrium	Prior only[5]
37.74	Initial insertion of transvenous lead (electrode) into epicardium	Prior only[5]
37.75	Revision of lead (electrode)	Index or prior[6]
37.76	Replacement of transvenous atrial and/or ventricular lead(s) (electrode)	Index or prior[6]
37.77	Removal of lead(s) (electrode) without replacement	Index only[7]
37.78	Insertion of temporary transvenous pacemaker system	Prior only[5]
37.79	Revision or relocation of pacemaker pocket	Index or prior[6]
37.80	Insertion of permanent pacemaker, initial or replacement, type of device not specified	Prior only[5]
37.81	Insertion of single-chamber device, not specified as rate responsive	Prior only[5]
37.82	Insertion of single-chamber device, rate responsive	Prior only[5]
37.83	Initial insertion of dual-chamber device	Prior only[5]
37.85	Replacement of any type pacemaker device with single-chamber device, not specified as rate responsive	Index or prior[6]
37.86	Replacement of any type pacemaker device with single-chamber device, rate responsive	Index or prior[6]
37.87	Replacement of any type pacemaker device with dual-chamber device	Index or prior[6]
37.89	Revision or removal of pacemaker device	Index only[7]
	Cerebrovascular disease, other (OTHCVAP)	Prior only
430	Subarachnoid hemorrhage	
431	Intracerebral hemorrhage	
432.x	Other and unspecified intracranial hemorrhage	

Table 7.6: ICD-9-CM codes for clinical risk factors evaluated but not included in final AMI models, *continued*

Code	*ICD-9-CM Description*	*Source of Data**
433.x	Occlusion and stenosis of precerebral arteries	
434.x	Occlusion of cerebral arteries	
435.x	Transient cerebral ischemia	
436	Acute but ill-defined cerebrovascular disease	
437.0	Cerebral atherosclerosis	
437.1	Other generalized ischemic cerebrovascular disease	
437.8	Other cerebrovascular disease	
437.9	Unspecified cerebrovascular disease	
	Chronic obstructive pulmonary disease (COPDB)	Index or prior
491.xx	Chronic bronchitis	
492.x	Emphysema	
494	Bronchiectasis	
496	Chronic airway obstruction, NEC	
500-505	Pneumoconioses and other lung diseases due to external agents	
	Coagulation defects (COAGULI)	Index only[1]
286.6	Defibrination syndrome	
286.7	Acquired coagulation factor deficiency	
286.9	Other and unspecified coagulation defects	
287.4	Secondary thrombocytopenia	
287.5	Thrombocytopenia, unspecified	
287.9	Unspecified hemorrhagic conditions	
	Collagen vascular disease (COLLVASB)	Index or prior
710.x	Diffuse diseases of connective tissue	
714.xx	Rheumatoid arthritis and other inflammatory polyarthropathies	
	Congestive heart failure (CHFI)	Index only[1]
	(CHFP)	Prior only
425.x	Cardiomyopathy	
428.x	Heart failure	
	Dementia (DEMENTB)	Index or prior
290.xx	Senile and presenile organic psychotic conditions	
294.x	Other organic psychotic conditions	
310.x	Specific nonpsychotic mental disorders due to organic brain damage	
331.0	Alzheimer's disease	
	Diabetes, uncomplicated (DBTUNCMB)	Index or prior[2]
250.0x	Diabetes, uncomplicated	
	Drug and alcohol abuse (DRUGALCB)	Index or prior
291.x	Alcoholic psychoses	
292.0	Drug withdrawal syndrome	
292.82	Drug-induced dementia	
303.xx	Alcohol dependence syndrome	
304.xx	Drug dependence	
305.0x	Alcohol abuse	
305.2x-305.9x	Nondependent abuse of drugs, other	
357.5	Alcoholic polyneuropathy	
425.5	Alcoholic cardiomyopathy	

Table 7.6: ICD-9-CM codes for clinical risk factors evaluated but not included in final AMI models, *continued*

Code	*ICD-9-CM Description*	*Source of Data**
535.3x	Alcoholic gastritis	
571.0	Alcoholic fatty liver	
571.1	Acute alcoholic hepatitis	
571.2	Alcoholic cirrhosis of liver	
571.3	Alcoholic liver damage, unspecified	
980.0	Toxic effect of alcohol, ethyl alcohol	
980.9	Toxic effect of alcohol, unspecified alcohol	
V11.3	Personal history of alcoholism	
	Fracture (FRACTURI)	Index only[1]
733.1x	Pathologic fracture	
800.xx-829.x	Fracture	
	Gastrointestinal hemorrhage (GIHEMORI)	Index only[1]
531.0x	Acute gastric ulcer, with hemorrhage	
531.2x	Acute gastric ulcer, with hemorrhage and perforation	
531.4x	Chronic or unspecified gastric ulcer, with hemorrhage	
531.6x	Chronic or unspecified gastric ulcer, with hemorrhage and perforation	
532.0x	Acute duodenal ulcer, with hemorrhage	
532.2x	Acute duodenal ulcer, with hemorrhage and perforation	
532.4x	Chronic or unspecified duodenal ulcer, with hemorrhage	
532.6x	Chronic or unspecified duodenal ulcer, with hemorrhage and perforation	
533.0x	Acute peptic ulcer, with hemorrhage	
533.2x	Acute peptic ulcer, with hemorrhage and perforation	
533.4x	Chronic or unspecified peptic ulcer, with hemorrhage	
533.6x	Chronic or unspecified peptic ulcer, with hemorrhage and perforation	
534.0x	Acute gastrojejunal ulcer, with hemorrhage	
534.2x	Acute gastrojejunal ulcer, with hemorrhage and perforation	
534.4x	Chronic or unspecified gastrojejunal ulcer, with hemorrhage	
534.6x	Chronic or unspecified gastrojejunal ulcer, with hemorrhage and perforation	
535.x1	Gastritis and duodenitis, with hemorrhage	
537.83	Angiodysplasia of stomach and duodenum, with hemorrhage	
562.02	Diverticulosis of small intestine, with hemorrhage	
562.03	Diverticulitis of small intestine, with hemorrhage	
562.12	Diverticulosis of colon, with hemorrhage	
562.13	Diverticulitis of colon, with hemorrhage	
569.85	Angiodysplasia of intestine, with hemorrhage	
578.x	Gastrointestinal hemorrhage	
	Nephritis (CHRGLOMB)	Index or prior
582.x	Chronic glomerulonephritis	
583.81	Nephritis and nephropathy, not specified as acute or chronic, in diseases classified elsewhere	
	Osteoarthrosis (OSTARTHB)	Index or prior[1]
274.xx	Gout	
715.xx	Osteoarthrosis and allied disorders	
716.9	Arthropathy, unspecified	

Table 7.6: ICD-9-CM codes for clinical risk factors evaluated but not included in final AMI models, *continued*

Code	*ICD-9-CM Description*	*Source of Data**
	Peripheral vascular disease (ATHEROSB)	Index or prior
440.x	Atherosclerosis	
441.x	Aortic aneurysm	
443.1	Thromboangiitis obliterans (Buerger's disease)	
443.8x	Other specified peripheral vascular diseases	
443.9	Peripheral vascular disease, unspecified	
	Pneumonia (PNEUMONI)	Index only[1]
481	Pneumococcal pneumonia	
482.xx	Other bacterial pneumonia	
483.x	Pneumonia due to other specified organism	
484.x	Pneumonia in infectious diseases classified elsewhere	
485	Bronchopneumonia, organism unspecified	
486	Pneumonia, organism unspecified	
	Psychosis (PSYCHOSB)	Index or prior
295.xx	Schizophrenic disorder	
296.xx	Affective psychoses	
297.x	Paranoid states	
298.x	Other nonorganic psychoses	
299.xx	Psychoses with origin specific to childhood	
	Seizure disorder (EPILEPP)	Prior only
345.xx	Epilepsy	
780.3	Convulsions	
	Skin ulcer (SKNULCRB)	Index or prior[1]
	(SKNULCRI)	Index only[1]
707.x	Chronic skin ulcer	
	Urinary tract infection (URINTRCI)	Index only[1]
590.1x	Acute pyelonephritis	
590.2	Renal and perinephric abscess	
590.8x	Other pyelonephritis or pyonephrosis	
590.9	Infection of kidney, unspecified	
595.0	Acute cystitis	
595.9	Cystitis, unspecified	
599.0	Urinary tract infection	

* **Index only:** variable ascertained only from index AMI hospitalizations (including linked hospitalizations when patients were transferred from one facility to another). These variables represent acute complications of AMI that may be important for risk-adjustment if present at admission.
Prior only: variable ascertained only from prior hospitalizations. These variables represent conditions that may be either acute or chronic, so they are counted as risk factors only if they were present during a prior admission.
Index or prior: variable ascertained from either index or prior hospitalizations. These variables represent conditions that are very unlikely to occur acutely, and therefore almost certainly represent comorbidities.

1. These conditions may represent complications of hospital care rather than comorbidities or pre-existing diagnoses. They were therefore included only in Model B, which was specifically designed to adjust for clinical conditions that could have arisen after a patient was admitted to the hospital.
2. If DBTCMPB = 0 in Table 7.3
3. If index record, unless 37.70-37.74, 37.78, 37.80-37.83, or 39.64 occurred during the same or a prior admission in the same series of transfer hospitalizations. If prior record, unless 37.77 or 37.89 occurred on the same or a subsequent admission during the same series of prior hospitalizations.

Table 7.6: ICD-9-CM codes for clinical risk factors evaluated but not included in final AMI models, *continued*

4. If index record, unless 37.70-37.74, 37.78, or 37.80-37.83, occurred during a prior admission in the same series of transfer hospitalizations. If prior record, unless 37.77 or 37.89 occurred on the same or a subsequent admission during the same series of prior hospitalizations.
5. Unless 37.77 or 37.89 occurred on the same or a subsequent day during the same series of prior hospitalizations.
6. If index record, unless 37.70-37.74, 37.78, or 37.80-37.83 occurred on the same or a prior day during the same series of transfer hosptializations. If prior record, unless 37.77 or 37.89 occurred on the same or a subsequent day during the same series of prior hospitalizations.
7. Unless 37.70-37.74, 37.78, or 37.80-37.83 occurred on the same or a prior day during the same series of transfer hospitalizations.

Table 7.7: ICD-9-CM codes for clinical risk factors ineligible for inclusion in AMI models

Code	*ICD-9-CM Description*	*Source of Data*[*]
	Acidosis (ACIDOSI)	Index only[1]
276.2	Acidosis	
276.4	Mixed acid-base balance disorders	
	Alkalosis (ALKALOSI)	Index only[1]
276.3	Alkalosis	
276.8	Hypopotassemia	
276.9	Electrolyte and fluid disorders NEC	
493.xx	**Asthma (ASTHMAB)**	Index or prior
	Bundle branch block (BBBLKB)	Index or prior[1]
	(BBBLKP)	Prior only
426.3	Other left bundle branch block	
426.4	Right bundle branch block	
426.51	Right bundle branch block and left posterior fascicular block	
426.52	Right bundle branch block and left anterior fascicular block	
426.53	Other bilateral bundle branch block	
426.54	Trifascicular block	
	Cardiomegaly (CARDMEGB)	Index or prior[1]
	(CARDMEGP)	Prior only
429.3	Cardiomegaly	
	Cerebrovascular disease, late effects (LATECVAB)	
342.x	Hemiplegia and hemiparesis	Prior only
438	Late effects of cerebrovascular disease	Index or prior
784.3	Aphasia	Prior only
	Chronic liver disease (CHRLIVEB)	Index or prior
456.0-456.2x	Esophageal varices	
571.2	Alcoholic cirrhosis of liver	
571.5	Cirrhosis of liver without mention of alcohol	
571.6	Biliary cirrhosis	
571.8	Other chronic nonalcoholic liver disease	
571.9	Unspecified chronic liver disease without mention of alcohol	
572.2	Hepatic coma	
572.3	Portal hypertension	
572.4	Hepatorenal syndrome	
572.8	Other sequelae of chronic liver disease	
573.0	Chronic passive congestion of liver	
573.8-573.9	Other specified and unspecified disorders of liver	
	Chronic pulmonary heart disease (CHRPULHB)	Index or prior
416.x	Chronic pulmonary heart disease	
	Hyperlipidemias (HYPERLIB)	Index or prior
272.0	Pure hypercholesterolemia	
272.1	Pure hyperglyceridemia	
272.2	Mixed hyperlipidemia	
272.3	Hyperchylomicronemia	
272.4	Other and unspecified hyperlipidemia	
	Hyperosmolality (HYPERMOI)	Index only[1]
276.0	Hyperosmolality/hypernatremia	
276.5	Volume depletion	
276.7	Hyperpotassemia	

Table 7.7: ICD-9-CM codes for clinical risk factors ineligible for inclusion in AMI models, *continued*

Code	*ICD-9-CM Description*	*Source of Data*[*]
	Hypertensive heart failure (HTHRTFB)	Index or prior
402.x1	Hypertensive heart disease (malignant, benign, or unspecified), with congestive heart failure	
404.x1	Hypertensive heart and renal disease (malignant, benign, or unspecified) with congestive heart failure	
404.x3	Hypertensive heart and renal disease (malignant, benign, or unspecified) with congestive heart and renal failure	
	Hyposomolality (HYPOSMOI)	Index only[1]
276.1	Hyposomolality/hyponatremia	
276.6	Fluid overload	
	Intermediate coronary syndrome (INCORSYI)	Index only[1]
411.1	Intermediate coronary syndrome	
	Malignant neoplasm history (HISMALIB)	Index or prior[2]
V10.00-V10.82, V10.84-V10.99	Personal history of malignant neoplasm, except of skin	
	Malignant neoplasm, low-risk primary (LRPMALIB)	Index or prior[3]
153.x	Malignant neoplasm of colon	
154.x	Malignant neoplasm of rectum	
160.x	Malignant neoplasm of nasal cavities	
161.x	Malignant neoplasm of larynx	
172.x	Malignant melanoma of skin	
174.x-195.x	Malignant neoplasm of female breast, uterus, cervix, ovary, placenta, prostate, testis, other reproductive organs, bladder, kidney, eye, brain, other nervous system, thyroid, other endocrine glands, other sites	
200.x-208.x	Lymphosarcoma and reticulosarcoma	
238.7	Neoplasm of uncertain behavior of other and unspecified sites and tissues, other lymphatic and hematopoietic tissues	
	Mitral valve disorders (MITVALVB)	Index or prior
	(MITVALVP)	Prior only
396.2	Mitral valve insufficiency and aortic valve stenosis	
396.3	Mitral valve insufficiency and aortic valve insufficiency	
396.8	Multiple involvement of mitral and aortic valves	
424.0	Mitral valve disorders	
	Nutritional disorders (NUTRITB)	Index or prior
260-263.x	Nutritional deficiencies	
799.4	Cachexia	
	Obesity (OBESITYB)	Index or prior
278.0	Obesity	
	Pleural effusion (PLEUREFI)	Index only[1]
511.1	Pleurisy, with effusion, with mention of bacterial cause other than tuberculosis	
511.8	Other specified forms of effusion	
511.9	Unspecified pleural effusion	
	Premature beats (PREBEATB)	Index or prior[1]
427.6x	Premature beats	
	Previous myocardial infarction (OLDAMIB)	
410.xx	Acute myocardial infarction	Prior only

Table 7.7: ICD-9-CM codes for clinical risk factors ineligible for inclusion in AMI models, *continued*

Code	*ICD-9-CM Description*	*Source of Data*[*]
412	Old myocardial infarction	Index or prior
	Supraventricular tachycardia (SUPVTACB) (SUPVTACP)	Index or prior[1] Prior only
427.0	Paroxysmal supraventricular tachycardia	
427.2	Paroxysmal tachycardia, unspecified	
427.32	Atrial flutter	
427.89	Other cardiac dysrhythmias	
427.9	Cardiac dysrhythmia, unspecified	
	Syncope (SYNCOPEI)	Index only[1]
780.2	Syncope	
	Valve disorders, other (OTHVALVE)	
394.x-397.x	Rheumatic valve disorders	Index or prior
424.1-424.9x	Nonrheumatic disorders involving valves other than mitral	Index or prior
996.02	Mechanical complication due to heart valve prothesis	Index or prior
996.71	Other complications due to heart valve prothesis	Index or prior
V42.2	Organ or tissue replaced by transplant, heart valve	Index or prior
V42.3	Organ or tissue replaced by other means, heart valve	Index or prior
35.2x	Replacement of heart valve	Prior only

* **Index only:** variable ascertained only from index AMI hospitalizations (including linked hospitalizations when patients were transferred from one facility to another). These variables represent acute complications of AMI that may be important for risk-adjustment if present at admission.
Prior only: variable ascertained only from prior hospitalizations. These variables represent conditions that may be either acute or chronic, so they are counted as risk factors only if they were present during a prior admission
Index or prior: variable ascertained from either index or prior hospitalizations. These variables represent conditions that are very unlikely to occur acutely, and therefore almost certainly represent comorbidities.

1. These conditions may represent complications of hospital care rather than comorbidities or pre-existing diagnoses. They were therefore included only in Model B, which was specifically designed to adjust for clinical conditions that could have arisen after a patient was admitted to the hospital.
2. If HRSECMAB = 0 in Table 7.3 and LRPMALIB = 0.
3. If HRSECMAB = 0 in Table 7.3.

Chapter 8

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. Development of the risk-adjustment models followed a series of ten steps. Table 8.1 shows the results of risk-factor evaluations in Steps 1, 2, and 3.

New in 1997

- Because this year's report includes AMI data from 1991-1993, no risk factors had to be eliminated from the analysis because of low frequency.
- Because six months of data were used to identify prior hospitalizations for each case in this report, the number of prior admissions and the number of weeks since the most recent prior admission were tested as risk factors for the first time.
- The variable selection methods used in this report are more reliable than those used in previous reports because 100 bootstrap samples were generated instead of just 10.
- In this report, a more thorough search was performed than ever before for possible two-way interactions among clinical risk factors (especially in Model B).

Step 1: Review of Potential Clinical Risk Factors

The potential clinical risk factors listed in Chapter Seven were reviewed to identify two important subsets, which were analyzed in somewhat different ways from the remaining risk factors. Note that the risk factors listed in Table 7.7 were not considered because OSHPD's AMI validation study showed that they were not reliably coded. These risk factors are labeled "ineligible" in Table 8.1.

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. The stepwise methods later used to select variables might otherwise have eliminated crucial predictors. 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, age, 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.

Before January 1, 1997, California patient discharge abstracts did not distinguish between comorbidities that were present at admission and complications that developed during an inpatient stay. This distinction is so important that 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. Model A includes only clinical risk factors with at least a 67 percent likelihood of having been present at admission, according to OSHPD's AMI validation study. Model B includes all of the clinical risk factors in Model A, plus clinical risk factors less likely to have been present at admission. Model B thereby gives hospitals the benefit of the doubt by adjusting for associated conditions that may actually represent complications of care.

AMI risk factors considered for Model B but not for Model A were acute renal failure, catastrophic sequelae of AMI, aspiration pneumonitis, coagulopathy, complete atrioventricular block, coma, fracture, gastrointestinal hemorrhage, pneumonia, pulmonary edema, paroxysmal ventricular tachycardia, sepsis, shock, and vascular insufficiency. These risk factors were based on ICD-9-CM codes from the index AMI hospitalization (and subsequent transfer hospitalizations), because they were considered unlikely to affect short-term mortality if they were diagnosed and resolved during a prior episode of care.

Diagnoses from prior hospitalizations were available for 17.1 percent of AMI cases. Several risk factors were considered for Model A only if they appeared on the discharge abstract from a prior hospitalization (proving that they were actually present at admission), but were considered for Model B no matter which discharge abstract listed the diagnosis. These AMI risk factors included arterial embolism or thrombosis, atrial fibrillation, epilepsy, other cerebrovascular disease, and skin ulcer.

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 aggregated as appropriate.

Binary risk factors present in less than 1 percent of all cases were examined carefully. Whenever possible, these risk factors were aggregated with physiologically related risk factors that showed similar associations with mortality. No AMI risk factors were eliminated strictly because of low frequency.

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 percent 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 even marginally associated with death ($p > 0.10$) were eliminated from further consideration. This cutoff was selected to screen out risk factors unlikely to contribute significantly to a multivariate model. The risk factors eliminated for this reason are shown in Table 8.1.

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 death were examined. Risk factors that appeared to lower the risk of death after AMI, when previous literature and clinical experience suggested the opposite effect, were eliminated from the analysis. Studies using reabstraction[11,12] or data linkage[13] have demonstrated substantial underreporting of several such conditions. Counterintuitive risk-outcome associations could be explained by selective underreporting among patients with poor outcomes.[14,15] The risk factors eliminated for this reason are shown in Table 8.1.

2.4 Multi-level clinical risk factors were summarized as multiple dummy (dichotomous) variables.

11. 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.
12. Romano PS, Mark DH. Bias in the coding of hospital discharge data and its implications for quality assessment. *Medical Care* 1994; 32:81-90.
13. 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.
14. 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.
15. 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.

Three clinical risk factors could be divided into multiple 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 could be classified as complicated if it was associated with ketoacidosis, coma, or end-organ disease (e.g., neuropathy, retinopathy, nephropathy). Hypertension could be classified as complicated if it was associated with kidney or heart disease. Infarct site could be classified as anterior wall, lateral wall, inferior wall, posterior wall, subendocardial, or other/unspecified.

To determine how to model the effects of these multi-level clinical risk factors, the unadjusted association between each factor and death was summarized using relative risk estimates with 95 percent confidence limits and p-values derived from the chi-square distribution. These analyses confirmed that multiple dummy variables would be preferable to a single ordinal variable for modelling each of the three risk factors. Two similar levels were combined into one dummy variable if they were associated with similar risk, such as inferior wall and posterior wall.

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

3.1 The distributions of age and other continuous or ordinal 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 one year and five year increments, so that each age group had a sufficient number of observations for analysis. Piecewise components of the age-mortality relationship were tested using various age cutoffs. As a result of this analysis, 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. In the subset of cases with no prior hospitalizations, a separate term was used to characterize the decreasing risk of death with advancing age up to 35 years.

The same approach was applied to examine the relationships between the *month, quarter, or year of admission* (ordered sequentially from the beginning to the end of the study period) and mortality. The monthly and quarterly analyses demonstrated a surprising cyclical pattern, with the highest risk of death in November and December of each year. Because

three years of data were insufficient to estimate this seasonal effect reliably, year of admission was selected as the most appropriate predictor. A set of dummy variables for year of admission was forced into each model so that each hospital's results could be stratified by year without introducing bias.

Among cases with one or more prior hospitalizations, the *number of prior admissions and the number of days or weeks from the most recent prior admission to the index AMI admission* were evaluated as potential predictors of death. The logit risk of death increased linearly with the number of prior admissions up to five, which was selected as the upper truncation point because only 1.3 percent (n=266) of those with any prior admissions had more than five. The logit risk of death decreased linearly with the number of days or weeks from the most recent prior admission to the index AMI admission, after approximately ten days. This effect was modeled using the interval in weeks (rounded downward to the nearest integer), so that cases admitted within the previous week would be assigned a value of 0 and would therefore fall into the reference group.

3.2 The distribution of categorical non-clinical variables and the associations between these variables and mortality 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) 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 are concentrated at certain hospitals. Adjusting for HMO insurance 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 percent of these

cases should have been reported as ER admissions, and because their risk of death was similar to that of ER 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 modeling. Males were selected as the reference group in all models. In all models that included race, white was the reference group. In all models that included source of payment, insurance other than Medicare and MediCal was the reference group. In all 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 percent of the original cases (without replacement) for the estimation sample and setting aside the remaining 40 percent 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 percent/40 percent 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) and by the presence or absence of prior hospitalizations, to ensure that the overall probability of death was the same in the corresponding 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 robust, significant predictors of death for Model A. To this end, 100 bootstrap 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 predictors, while minimizing the risk of overfitting the model to the peculiarities of a particular sample. This procedure was applied separately to cases with and without prior admissions, because of differences in the list of available predictors.

5.1 One hundred random subsamples were generated, without replacement and with a sampling fraction of 50 percent, from the 60 percent estimation sample.

Sampling without replacement means that the same case could not have been selected more than once for a single subsample. Sampling with replacement has the theoretical, but minor, 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 set of risk factors for each subsample was determined by stepwise regression.

For each subsample, a multivariate logistic regression model was fit using stepwise forward selection with the significance level tolerance set to 0.01, 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. Backward elimination procedures were tested in previous years and were found to generate identical results, so only forward selection was used in this year's report.

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

All risk factors that were significant at $p < 0.01$ in 50 or more of the 100 subsamples were retained in the construction of Model A. In fact, only one risk factor in the model for cases with one or more prior admissions, and one risk factor in the model for cases without prior admissions, entered more than 35 but fewer than 88 subsamples. This finding suggests a relatively clear dichotomy between robust and non-robust predictors. The risk factors that were eliminated at this stage are shown in Table 8.1.

5.4 The variables confirmed as robust predictors of adverse outcomes were tested in a stepwise logistic regression model on the entire 60 percent 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 50-sample bootstrap criterion were tested in a stepwise logistic regression using the full 60 percent estimation sample (p-to-enter < 0.01). In fact, this procedure eliminated no risk factors from either AMI model.

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 selected approach was based on the premise that only interactions involving the most statistically or clinically important main effects would contribute meaningfully to risk-adjustment models. Therefore, only interactions involving age, infarct site (e.g., anterior wall, inferior wall, other or unspecified), prior coronary artery bypass surgery, and congestive heart failure were evaluated.

All of these interactions were tested using the 100 randomly generated subsamples described above. For each bootstrap subsample, a multivariate logistic regression model was fit using stepwise forward selection with the significance level tolerance set to 0.01, 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. All interactions that were significant at $p < 0.01$ in 50 or more of the 100 subsamples were retained in the construction of Model A, and then confirmed in a stepwise logistic regression using the full 60 percent estimation sample (p-to-enter < 0.01).

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 percent estimation sample were compared to the corresponding parameter estimates derived by fitting the same model to the 40 percent validation sample. Model specification was considered adequate if a parameter estimate from the 60 percent estimation sample fell within the corresponding 95 percent confidence intervals from the 40 percent validation sample. The calibration of each risk-adjustment model was assessed with the Hosmer-Lemeshow goodness of fit test, as described in Chapter Ten. Specifically, the risk-adjustment model developed on the 60 percent estimation sample was applied to the 40 percent 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.

These procedures generally confirmed the internal validity of Model A. All parameter estimates based on the 60 percent estimation samples were within, or just slightly outside, the corresponding 95 percent confidence intervals based on the 40 percent validation samples. Although a few statistically significant variables in the estimation sample were not significant in the validation sample, none changed signs (e.g., showed an adverse effect in the estimation sample and a protective effect in the validation sample, or vice versa). The Hosmer-Lemeshow goodness-of-fit test showed that the two

models estimated using the 60 percent sample actually fit as well or better using the 40 percent validation sample (see Table 10.1 for details). Therefore, no further changes to the risk-adjustment models were necessary.

Step 8: Selection of Additional Main Effects Risk Factors for Model B

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 percent 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.

To select additional risk factors for Model B, a procedure was applied similar to that used to select Model A risk factors in Step 5. One hundred random bootstrap subsamples were generated without replacement, and with a sampling fraction of 50 percent, from the 60 percent analytic sample. Stepwise forward selection, forcing in all of the main effect and interaction variables from Model A, was used to select covariates. Model A covariates were forced into this model to ascertain the independent effects of additional demographic and clinical factors, adjusting for those included in Model A. Candidate risk factors that were significantly associated with mortality at the $p < 0.01$ level in 50 or more of the 100 subsamples were retained in Model B, except that race and at least one category of expected payment source were forced in to adjust for any socioeconomic effect, no matter how insignificant. All risk factors that met the 50-sample bootstrap criterion were tested in a stepwise logistic regression using the full 60 percent estimation sample (p-to-enter < 0.01). The entire analysis was stratified by the presence or absence of prior admissions.

In fact, only one risk factor in the model for cases with one or more prior admissions, and none in the model for cases without prior admissions, entered more than 40 but fewer than 90 subsamples. This finding suggests a

relatively clear dichotomy between robust and non-robust predictors. The risk factors that were eliminated by the bootstrap subsample procedure are shown in Table 8.1. No risk factors were eliminated in the final stepwise regression using the full 60 percent sample.

Step 9: Selection of Additional Risk Factor Interactions for Model B

Although the version of Model B estimated at the end of Step 8 had excellent discrimination, its calibration was poor. In other words, the model significantly overestimated the probability of death among high-risk patients. This problem was felt to be secondary to unidentified interactions. However, the number of Model B risk factors was too large to consider all two-way interactions, let alone three-way and higher order interactions. The selected approach was based on the premise that only interactions involving the most statistically or clinically important main effects would contribute meaningfully to risk-adjustment models. Therefore, all interactions between the additional risk factors in Model B and age, infarct site (e.g., anterior wall, inferior wall, other or unspecified), prior coronary artery bypass surgery, congestive heart failure, and shock were tested in Step 9.1.

9.1 Among cases with prior admissions, the most significant and robust interactions between Model B risk factors and several variables in Model A were selected using bootstrap subsamples.

All of these candidate interactions were tested using the 100 randomly generated subsamples described above. For each bootstrap subsample, a multivariate logistic regression model was fit using stepwise forward selection with the significance level tolerance set to 0.01, forcing in all of the important main effects identified in Step 8, plus the Model A main effects and interactions identified previously. Probability values to enter and remove variables were based on the likelihood ratio statistic. All interactions that were significant at $p < 0.01$ in 50 or more of the 100 subsamples were retained in the construction of Model A, and then confirmed in a stepwise logistic regression using the full 60 percent estimation sample (p-to-enter < 0.01).

This procedure identified only one significant and robust interaction (between congestive heart failure and shock). Adding this interaction to Model B did not improve calibration of the model. After exploratory analyses, it was recognized that powerful interactions among the additional clinical risk factors in Model B were causing substantial bias. However, the bootstrap subsample approach proved impractical for identifying these interactions because of its computational demands.

9.2 Interactions among the additional clinical risk factors in Model B were identified by stepwise regression using the full 60 percent estimation sample.

Among cases with one or more prior admissions, all of the two-way interactions between shock and both demographic (e.g., age, sex) and clinical risk factors were created. Among cases with no prior admissions, all of the two-way interactions between all additional eleven clinical risk factors from Model B and both demographic (e.g., age, sex) and clinical risk factors were created. Multivariate logistic regression models were fit using stepwise forward selection on the full 60 percent estimation sample, with the significance level tolerance set to 0.01. All of the important main effects identified in Step 8, plus the Model A main effects and interactions identified previously, were forced in. Probability values to enter and remove variables were based on the likelihood ratio statistic.

Among cases with one or more prior admissions, this analysis was performed in one step and identified six additional interactions involving shock (as shown in Table 9.4). Among cases with no prior admissions, this analysis had to be performed in a series of steps. Two-way interactions involving each of the eleven additional clinical risk factors in Model B were entered in a block, with the corresponding main effect. After all eleven blocks of interaction variables were entered and tested, superfluous interaction terms were removed by backward elimination (p-to-exit < 0.01). This procedure identified 22 additional interactions, as shown in Table 9.3.

Step 10: Re-Estimation of Model Parameters Using All Cases

The 60 percent estimation sample and the 40 percent validation sample were re-combined into the full dataset. Model A and Model B were reestimated by fitting the models developed in Steps 1 through 9 to the complete (100 percent) data set. The purpose of this step was to generate the most reliable possible estimate of each parameter, using all available data. The final models reestimated in this step were used to estimate the probability of death for each case in the analysis. These estimated probabilities were used in all subsequent analyses of hospital mortality.

Table 8.1: Results of risk factor evaluation for AMI mortality models

		AMI mortality model				
		Priors		*No Priors*		*ICD-9-CM*
Risk factor	*Description*	*Model A*	*Model B*	*Model A*	*Model B*	*Table*
ACIDOSI	Acidosis	Ineligible[1]				7.7
ACRENALI	Renal failure, acute or unspecified	Not evaluated	Included	Not evaluated	Included	7.3
ALKALOSI	Alkalosis	Ineligible[1]				7.7
AMISEQUI	Catastrophic sequelae of AMI	Not evaluated	Excluded[6]	Not evaluated	Included	7.3
ARTEREMI	Arterial embolism (index only)	Not evaluated	Excluded[6]	Not evaluated	Excluded[6]	7.6
ARTEREMP	Arterial embolism (prior only)	Excluded[6]	Not evaluated	Not evaluated		7.6
ASPPNEUI	Aspiration pneumonia	Not evaluated	Included	Not evaluated	Included	7.3
ASTHMAB	Asthma	Ineligible[1]				7.7
ATHEROSB	Peripheral vascular disease	Excluded[6]	Not evaluated	Excluded[6]	Not evaluated	7.6
ATRFIBB	Atrial fibrillation (index or prior)	Not evaluated	Excluded[6]	Not evaluated	Excluded[6]	7.6
ATRFIBP	Atrial fibrillation (prior only)	Excluded[6]	Not evaluated	Not evaluated		7.6
BBBLKB	Bundle branch block (index or prior)	Ineligible[1]				7.7
BBBLKP	Bundle branch block (prior only)	Ineligible[1]		Not evaluated		7.7
CARDMEGB	Cardiomegaly (index or prior)	Ineligible[1]				7.7
CARDMEGP	Cardiomegaly (prior only)	Ineligible[1]		Not evaluated		7.7
CHFB	Congestive heart failure (index or prior)	Included	Included	Included	Included	7.3
CHFI	Congestive heart failure (index only)	Excluded[3]		Excluded[3]		7.6
CHFP	Congestive heart failure (prior only)	Excluded[3]		Not evaluated		7.6
CHRGLOMB	Nephritis	Excluded[4]		Excluded[6]	Not evaluated	7.6
CHRLIVEB	Chronic liver disease	Ineligible[1]				7.7
CHRPULHB	Chronic pulmonary heart disease	Ineligible[1]				7.7
CHRRENAB	Renal failure, chronic	Included	Included	Included	Included	7.3
CNSDISB	Central nervous system disease	Included	Included	Included	Included	7.3
COAGULI	Coagulation defects	Not evaluated	Excluded[6]	Not evaluated	Excluded[6]	7.6
COATRBLI	Complete atrioventricular block	Not evaluated	Excluded[6]	Not evaluated	Included	7.3

Table 8.1: Results of risk factor evaluation for AMI mortality models, *continued*

		AMI mortality model				
		Priors		*No Priors*		*ICD-9-CM*
Risk factor	*Description*	*Model A*	*Model B*	*Model A*	*Model B*	*Table*
COLLVASB	Collagen vascular disease	Excluded[4]		Excluded[4,5]		7.6
COMAI	Coma	Not evaluated	Included	Not evaluated	Included	7.3
COPDB	Chronic obstructive pulmonary disease	Excluded[6]	Not evaluated	Excluded[6]	Not evaluated	7.6
DBTCMPB	Diabetes, complicated	Included	Included	Included	Included	7.3
DBTUNCMB	Diabetes, uncomplicated	Excluded[5]		Excluded[6]	Not evaluated	7.6
DEG1AVBB	Atrioventricular block, first degree	Excluded[4,5]		Excluded[5]		7.6
DEG2AVBB	Atrioventricular block, second degree (index or prior)	Excluded[4]		Excluded[4]		7.6
DEG2AVBP	Atrioventricular block, second degree (prior only)	Excluded[4,5]		Not evaluated		7.6
DEMENTB	Dementia	Excluded[6]	Not evaluated	Excluded[6]	Not evaluated	7.6
DRUGALCB	Drug and alcohol abuse	Excluded[5]		Excluded[5]		7.6
EPILEPB	Seizure disorder (index or prior)	Not evaluated	Included	Not evaluated	Included	7.3
EPILEPP	Seizure disorder (prior only)	Excluded[6]	Not evaluated	Not evaluated		7.6
FRACTURI	Fracture	Not evaluated	Excluded[6]	Not evaluated	Excluded[6]	7.6
GIHEMORI	Gastrointestinal hemorrhage	Not evaluated	Excluded[6]	Not evaluated	Excluded[6]	7.6
HISMALIB	Malignant neoplasm history	Ineligible[1]				7.7
HRSECMAB	High-risk or secondary malignant neoplasm	Included	Included	Included	Included	7.3
HTB	Hypertension	Included	Included	Included	Included	7.3
HTHRTFB	Hypertensive heart failure	Ineligible[1]				7.7
HYPERLIB	Hyperlipidemias	Ineligible[2]				7.7
HYPERMOI	Hyperosmolality	Ineligible[1]				7.7
HYPOSMOI	Hyposomolality	Ineligible[1]				7.7
INCORSYI	Intermediate coronary syndrome	Ineligible[2]				7.7
LATECVAB	Cerebrovascular disease, late effects	Ineligible[1]				7.7
LRPMALIB	Malignant neoplasm, low-risk primary	Ineligible[1]				7.7
MITVALVB	Mitral valve disorders (index or prior)	Ineligible[1]				7.7

		AMI mortality model				
		Priors		No Priors		ICD-9-CM
Risk factor	*Description*	*Model A*	*Model B*	*Model A*	*Model B*	*Table*
MITVALVP	Mitral valve disorders (prior only)	Ineligible[1]		Not evaluated		7.7
NUTRITIB	Nutritional disorders	Ineligible[1]				7.7
OBESITYB	Obesity	Ineligible[1]				7.7
OLDAMIB	Previous myocardial infarction	Ineligible[1]				7.7
OSTARTHB	Osteoarthrosis	Excluded[4,5]		Excluded[5]		7.6
OTHCVAI	Cerebrovascular disease, other (index only)	Not evaluated	Included	Not evaluated	Included	7.3
OTHCVAP	Cerebrovascular disease, other (prior only)	Excluded[6]	Not evaluated	Not evaluated		7.6
OTHVALVE	Valve disorders, other	Ineligible[1]				7.7
PLEUREFI	Pleural effusion	Ineligible[1]				7.7
PNEUMONI	Pneumonia	Not evaluated	Excluded[6]	Not evaluated	Excluded[6]	7.6
PRCABG	Prior coronary artery bypass graft	Included	Included	Included	Included	7.3
PREBEATB	Premature beats	Ineligible[1]				7.7
PRPACE	Cardiac pacemaker	Excluded[6]	Not evaluated	Excluded[6]	Not evaluated	7.6
PSYCHOSB	Psychosis	Excluded[6]	Not evaluated	Excluded[4]		7.6
PULEDEMI	Pulmonary edema	Not evaluated	Included	Not evaluated	Included	7.3
PVENTACI	Paroxysmal ventricular tachycardia	Not evaluated	Included	Not evaluated	Included	7.3
SEPSISI	Sepsis	Not evaluated	Included	Not evaluated	Excluded[6]	7.3
SHOCKI	Shock	Not evaluated	Included	Not evaluated	Included	7.3
SITE_ANT	Infarction site, anterior wall	Included	Included	Included	Included	7.3
SITE_INF	Infarction site, inferior wall	Included	Included	Included	Included	7.3
SITE_OI	Infarction site, other	Included	Included	Included	Included	7.3
SKNULCRB	Skin ulcer (index or prior)	Not evaluated		Not evaluated	Excluded[6]	7.6
SKNULCRI	Skin ulcer (index only)	Not evaluated	Excluded[6]	Not evaluated		7.6
SKNULCRP	Skin ulcer (prior only)	Included	Included	Not evaluated		7.3
SUPVTACB	Supraventricular tachycardia (index or prior)	Ineligible[1]				7.7
SUPVTACP	Supraventricular tachycardia (prior only)	Ineligible[1]		Not evaluated		7.7
SYNCOPEI	Syncope	Ineligible[1]				7.7
THYROIDB	Thyroid disease	Excluded[4]		Included	Included	7.3
URINTRCI	Urinary tract infection	Excluded[4]		Excluded[4]		7.6
VASINSUI	Ischemic bowel or liver	Not evaluated	Excluded[6]	Not evaluated	Included	7.3

Table 8.1: Results of risk factor evaluation for AMI mortality models, *continued*

		AMI mortality model				
		Priors		*No Priors*		*ICD-9-CM*
Risk factor	*Description*	*Model A*	*Model B*	*Model A*	*Model B*	*Table*

Reason for ineligibility or exclusion:

1. Poorly coded according to the 1996 AMI validation study, with weighted sensitivity less than 30% and weighted kappa statistic less than 0.4.
2. Temporal analysis demonstrated an implausible change in prevalence during the study period, suggesting coding variability.
3. Consolidated into another, more prevalent risk factor.
4. No association with mortality in bivariate analyses ($p > 0.10$).
5. Counterintuitive association with decreased mortality in bivariate and/or multivariate analyses.
6. Significant in less than 50 of 100 random bootstrap samples (as described under “Step 5: Selection of Main Effects Risk Factors for Model A” and “Step 8: Selection of Additional Main Effects Risk Factors for Model B”).

Chapter 9

Presentation and Interpretation of Final Models

In this chapter, the final risk-adjustment models developed through the process described in Chapter Eight are presented. These models represent a best effort to elucidate the relationship between AMI mortality and various demographic and clinical risk factors.

New in 1997

- This report includes minor changes to all of the risk-adjustment models. These changes reflect the cumulative impact of the methodologic enhancements described in previous chapters.

- The most important of these enhancements were the use of total 30-day mortality instead of inpatient 30-day mortality as the outcome, the use of six months instead of eight weeks of data from prior hospitalizations to ascertain risk factors, and the redefinition and exclusion of several risk factors based on findings from OSHPD's AMI validation study.

The Four Models

The risk-adjustment models for AMI mortality were classified according to whether one or more hospitalizations occurred during the 8 weeks before the index admission. If there were prior hospitalizations, then more information about possible comorbidities was available. For example, cerebrovascular disease could be used as a risk factor in Model A if it was diagnosed during a prior hospitalization. If no records from prior hospitalizations were available, cerebrovascular disease could not be used as a risk factor in Model A because it could have represented an in-hospital complication of the AMI. Overall, 19,882 (17.1 percent) of the 116,174 study cases had one or more prior hospitalizations.

Table 9.1 shows the AMI Model A parameters for cases with no prior admissions; Table 9.2 shows the Model A parameters for cases with one or more prior admissions. Table 9.3 shows the Model B parameters for cases with no prior admissions; Table 9.4 shows the Model B parameters for cases with one or more prior admissions. Each risk variable in these tables is defined in Chapter Seven.

The columns in these tables provide the following information:

1. **The parameter estimate** is a measure of the risk associated with a covariate. A negative parameter estimate indicates that the covariate has

a protective effect (reduces risk); a positive parameter estimate indicates that the covariate has a harmful effect (increases risk). The further this parameter estimate is from zero, the greater the impact of this covariate on the risk of AMI death. These numbers are maximum likelihood estimates, meaning that they are more consistent with the observed data than any other possible set of parameter estimates.

The relationship between these parameter estimates and the estimated probability of death can be expressed as:

$$\log\left(\frac{\hat{p}}{1-\hat{p}}\right) = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \ldots + \hat{\beta}_q x_q$$

where $\hat{p}$ represents the estimated probability of death within 30 days after an AMI, $\hat{\beta}_0$ represents the intercept term, $x_1, \ldots, x_q$ represent risk variables, and $\hat{\beta}_1, \ldots, \hat{\beta}_q$ represent the associated parameter estimates. Solving for the estimated probability of death, this formula can be rewritten as:

$$\hat{p} = \left[1 + e^{-\left(\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \ldots + \hat{\beta}_q x_q\right)}\right]^{-1}$$

2. **The p-value** is a measure of the statistical significance of a parameter estimate. It is based on the Wald statistic, which approximately follows a chi-square distribution. A small p-value (less than 0.05) indicates that the observed data are *not* consistent with the null hypothesis that the true value of the parameter is zero.

3. **The estimated odds ratio** associated with a covariate is another measure of risk, which may be easier to interpret than the parameter estimate. It equals the odds of death ($\hat{p}/[1-\hat{p}]$, where $\hat{p}$ is the probability of death) among patients with a risk factor, divided by the odds of death among patients without that characteristic, adjusted for all of the other factors in the model. When the outcome is relatively infrequent, this odds ratio approximates the relative risk. An odds ratio less than one indicates that the covariate has a protective effect; an odds ratio greater than one indicates that the covariate has a harmful effect.

   The estimated odds ratios were derived by exponentiating the corresponding parameter estimates. For example, the odds ratio of 1.42 for CHRRENAB in Table 9.2 is equal to $e^{0.3502}$. This odds ratio represents the odds of death among AMI patients with chronic renal failure, divided by the odds of death among similar patients without chronic renal failure.

   Note that the odds ratio for age, which is a continuously distributed variable, must be interpreted differently from other odds ratios. In this case, the estimated odds ratio represents the odds of death among patients of a certain age, divided by the odds of death among patients who are one year younger. The odds ratio associated with a ten-year age difference can be computed by raising the one-year odds ratio to the tenth power.

If a risk factor is involved in a two-way interaction with any other risk factor, these odds ratios may be misleading. With a statistically significant ($p<0.05$) interaction, the effect of a particular risk factor on outcomes varies according to the level of a second risk factor. For example, the odds ratio associated with risk factor A may equal 4 if risk factor B is absent, but may equal 2 if that risk factor is present. To calculate the odds ratio for one variable conditioned on a specific value of a second (interacting) variable, use this formula:

$$\begin{aligned} \mathrm{OR}(x_1|x_2=\delta) &= \frac{\mathrm{odds}(x_1=a|x_2=\delta)}{\mathrm{odds}(x_1=b|x_2=\delta)} \\ &= \frac{e^{(\hat{\beta}_1 a+\hat{\beta}_3 a\delta)}}{e^{(\hat{\beta}_1 b+\hat{\beta}_3 b\delta)}} \end{aligned}$$

where x_1 and x_2 represent the two interacting risk factors, $\hat{\beta}_1$ and $\hat{\beta}_2$ represent the corresponding parameter estimates ($\hat{\beta}_2$ drops out of the above formula because x_2 is fixed equal to δ), $\hat{\beta}_3$ represents the parameter estimate for the two-way interaction, and a and b represent two possible values of the first risk factor (x_1).

4. **The upper and lower confidence limits for the odds ratio** are an expression of confidence in the estimated odds ratio. There is a 95 percent probability that the true value of the odds ratio is between the lower confidence limit and the upper confidence limit. If the interval between these confidence limits includes one, then the null hypothesis that the covariate has no effect on the outcome cannot be rejected.

   The confidence limits for the odds ratio were computed by exponentiating the upper and lower confidence limits for the corresponding parameter estimate. These confidence limits were computed by adding 1.96 times the estimated standard error of the parameter estimate to its original value (upper limit), and subtracting 1.96 times the estimated standard error of the parameter estimate from its original value (lower limit). These standard errors are not shown, but are available upon request from OSHPD.

   If a risk factor is involved in a two-way interaction with another risk factor, these confidence limits may be misleading. To calculate the confidence limits for one variable conditioned on a specific value of a second (interacting) variable, one must refer to the covariance matrix of the parameter estimates (available upon request from OSHPD).

Table 9.1 shows that the following factors were associated with a significantly increased risk of death among AMI cases **without** prior hospitalizations: congestive heart failure (CHF), high-risk or metastatic malignancy,

complicated diabetes, chronic kidney disease, chronic central nervous system disease, female sex, anterior wall site, and other or unspecified site. Uncomplicated hypertension, hypothyroidism, and prior coronary artery bypass grafting were associated with a significantly decreased risk of death among these AMI cases. The relationship between age and mortality followed a spline function, with decreasing risk of death up to 35 years of age and increasing risk of death above that age. Relative to AMIs in 1993, AMIs in 1991 were associated with 8 percent (significantly) higher mortality and those in 1992 were associated with 4 percent (nonsignificantly) higher mortality. The interaction terms reveal that the independent effect of CHF on mortality decreased with age (reaching zero at 115 years) and varied by site; this effect was greatest among subendocardial infarctions and smallest among infarctions of unspecified or other site. In addition, the impact of CHF was greater among men than among women. The protective effect of prior coronary bypass grafting disappeared among cases with CHF. The incremental risk associated with female sex and other or unspecified site declined with age, whereas the incremental risk associated with inferior site increased with age.

Table 9.2 shows that the following factors were associated with a significantly increased risk of death among AMI cases **with** prior hospitalizations: CHF, high-risk or metastatic malignancy, complicated diabetes, chronic kidney disease, chronic central nervous system disease, skin ulcer (if diagnosed on a prior hospitalization), female sex, age, anterior wall site, inferior wall site, and other or unspecified site. Uncomplicated hypertension was associated with a significantly decreased risk of death among these AMI cases. Relative to AMIs in 1993, AMIs in 1991 and 1992 were associated with 5 percent (nonsignificantly) higher mortality. The number of weeks between the AMI admission and the most recent prior admission was inversely related to AMI mortality. The interaction terms reveal that the independent effect of CHF on mortality decreased with age (reaching zero at 117 years) and varied by site; this effect was greatest among subendocardial infarctions and smallest among infarctions of unspecified or other site. In addition, the impact of CHF was greater among men than among women. This model includes fewer predictors than the model in Table 9.1 because of its smaller sample size.

The following additional risk factors in Model B were associated with a significantly increased risk of death among AMI cases **without** prior hospitalizations (Table 9.3): pulmonary edema, shock, cerebrovascular disease, paroxysmal ventricular tachycardia, coma, aspiration pneumonia, acute kidney disease, complete atrioventricular block, ischemic bowel or liver, catastrophic sequelae of AMI, and epilepsy. All of these risk factors except epilepsy were derived exclusively from the index record. Uninsured patients had a higher risk of death than privately insured or MediCal patients. Black patients had lower risk than white or Hispanic patients. Among AMI cases with shock, other risk factors such as CHF, acute kidney disease, aspiration pneumonia, coma, cerebrovascular disease, pulmonary edema, complete atrioventricular block, ventricular tachycardia, and catastrophic AMI sequelae, conferred little additional risk. Several other interaction terms

accounted for the fact that combinations of Model B risk factors generally showed less than multiplicative effects on the odds of death.

The following additional risk factors in Model B were associated with a significantly increased risk of death among AMI cases **with** prior hospitalizations (Table 9.4): pulmonary edema, shock, cerebrovascular disease, paroxysmal ventricular tachycardia, coma, sepsis, aspiration pneumonia, acute kidney disease, and epilepsy. All of these risk factors except epilepsy were derived exclusively from the index record. Payer source was not associated with the risk of death, but black patients had lower risk than white or Hispanic patients. Among AMI cases with shock, other risk factors such as CHF, acute kidney disease, aspiration pneumonia, coma, cerebrovascular disease, pulmonary edema, and sepsis, conferred little additional risk.

Table 9.1: Acute myocardial infarction mortality Model A, cases with no prior admissions (N=96,292)

*Variable**	*Parameter Estimate*	*p value*	*Lower CI for Odds Ratio*	*Odds Ratio*	*Upper CI for Odds Ratio*
INTERCEP	-8.0072	0.0001			
FEMALE	1.0697	0.0001	2.22	2.91	3.82
IAGEYRS	0.0692	0.0001	1.07	1.07	1.08
IAGE35	0.1677	0.0001	1.12	1.18	1.24
IADM91	0.0662	0.0083	1.02	1.07	1.12
IADM92	0.0292	0.2423	0.98	1.03	1.08
CHFB	2.8105	0.0001	12.42	16.62	22.24
CHRRENAB	0.4372	0.0001	1.40	1.55	1.71
CNSDISB	0.3314	0.0001	1.18	1.39	1.65
DBTCMPB	0.4108	0.0001	1.41	1.51	1.61
HRSECMAB	0.7353	0.0001	1.74	2.09	2.50
HTB	-0.4645	0.0001	0.60	0.63	0.66
PRCABG	-0.2655	0.0001	0.68	0.77	0.86
SITE_ANT	1.4050	0.0001	3.75	4.08	4.43
SITE_INF	0.1592	0.3403	0.85	1.17	1.63
SITE_OI	3.4170	0.0001	20.80	30.48	44.66
THYROIDB	-0.6111	0.0001	0.48	0.54	0.61
I_AGECHF	-0.0246	0.0001	0.97	0.98	0.98
I_AGEFEM	-0.0106	0.0001	0.99	0.99	0.99
I_AGEINF	0.0116	0.0001	1.01	1.01	1.02
I_AGESIT	-0.0175	0.0001	0.98	0.98	0.99
I_ANTCHF	-0.5019	0.0001	0.54	0.61	0.68
I_CHFCAB	0.4242	0.0001	1.30	1.53	1.80
I_CHFFEM	-0.3004	0.0001	0.68	0.74	0.80
I_INFCHF	-0.2823	0.0001	0.67	0.75	0.85
I_SITCHF	-0.8588	0.0001	0.37	0.42	0.49

* For the full name and ICD-9-CM description of each clinical variable, see Table 7.3. For the description of each demographic variable, see the Appendix in the *User's Guide*.

Table 9.2: Acute myocardial infarction mortality Model A, cases with one or more prior admissions (N=19,882)

*Variable**	*Parameter Estimate*	*p value*	*Lower CI for Odds Ratio*	*Odds Ratio*	*Upper CI for Odds Ratio*
INTERCEP	-6.4169	0.0001			
FEMALE	0.2744	0.0001	1.15	1.32	1.50
IAGEYRS	0.0533	0.0001	1.05	1.05	1.06
IADM91	0.0255	0.5870	0.94	1.03	1.12
IADM92	0.0409	0.3795	0.95	1.04	1.14
PRIOLAGW	-0.0192	0.0001	0.98	0.98	0.99
CHFB	2.3058	0.0001	5.88	10.03	17.13
CHRRENAB	0.3502	0.0001	1.27	1.42	1.59
CNSDISB	0.5381	0.0001	1.39	1.71	2.12
DBTCMPB	0.1772	0.0004	1.08	1.19	1.32
HRSECMAB	0.8153	0.0001	1.89	2.26	2.70
HTB	-0.2909	0.0001	0.69	0.75	0.81
PRCABG	-0.0248	0.6427	0.88	0.98	1.08
SITE_ANT	1.1703	0.0001	2.93	3.22	3.55
SITE_INF	0.9264	0.0001	2.26	2.53	2.82
SITE_OI	2.0317	0.0001	6.32	7.63	9.20
SKNULCRP	0.5063	0.0001	1.34	1.66	2.06
I_AGECHF	-0.0199	0.0001	0.97	0.98	0.99
I_CHFFEM	-0.4019	0.0001	0.57	0.67	0.79
I_SITCHF	-0.5424	0.0001	0.47	0.58	0.72

* For the full name and ICD-9-CM description of each clinical variable, see Table 7.3. For the description of each demographic variable, see the Appendix in the *User's Guide.*

Table 9.3 Acute myocardial infarction mortality Model B, cases with no prior admissions (N=95,755)

*Variable**	*Parameter Estimate*	*p value*	*Lower CI for Odds Ratio*	*Odds Ratio*	*Upper CI for Odds Ratio*
INTERCEP	-8.8307	0.0001			
FEMALE	0.7101	0.0001	1.50	2.03	2.76
IAGEYRS	0.0711	0.0001	1.07	1.07	1.08
IAGE35	0.1277	0.0003	1.06	1.14	1.22
RACE_BLK	-0.1365	0.0130	0.78	0.87	0.97
RACE_HIS	0.0121	0.7699	0.93	1.01	1.10
PAY_MCAL	0.1460	0.0043	1.05	1.16	1.28
PAY_UNIN	0.3696	0.0001	1.30	1.45	1.61
IADM91	0.1058	0.0002	1.05	1.11	1.17
IADM92	0.0585	0.0350	1.00	1.06	1.12
ACRENALI	1.6881	0.0001	4.53	5.41	6.46
AMISEQUI	1.5342	0.0001	3.16	4.64	6.80
ASPPNEUI	1.2056	0.0001	2.76	3.34	4.04
ATYP_ER	0.3571	0.0001	1.36	1.43	1.50
CHFB	2.1374	0.0001	6.09	8.48	11.80
CHRRENAB	0.5091	0.0001	1.49	1.66	1.85
CNSDISB	0.2474	0.0089	1.06	1.28	1.54
COATRBLI	0.8439	0.0001	2.07	2.33	2.61
COMAI	4.7353	0.0001	44.13	113.90	293.97
DBTCMPB	0.2893	0.0001	1.24	1.34	1.44
EPILEPB	3.1868	0.0001	11.47	24.21	51.12
HRSECMAB	0.8563	0.0001	1.93	2.35	2.87
HTB	-0.3586	0.0001	0.67	0.70	0.73
OTHCVAI	1.1666	0.0001	2.85	3.21	3.62
PRCABG	-0.2110	0.0019	0.71	0.81	0.93
PULEDEMI	1.5072	0.0001	3.94	4.51	5.18
PVENTACI	0.4160	0.0001	1.40	1.52	1.64
SHOCKI	3.2618	0.0001	23.26	26.10	29.28
SITE_ANT	1.1963	0.0001	3.02	3.31	3.62
SITE_INF	-0.1458	0.4351	0.60	0.86	1.25
SITE_OI	3.4306	0.0001	20.10	30.89	47.49
THYROIDB	-0.5249	0.0001	0.52	0.59	0.68
VASINSUI	0.9027	0.0001	1.78	2.47	3.41
I_ACRCHF	-0.4866	0.0001	0.50	0.61	0.75
I_ACROTH	-0.9332	0.0001	0.26	0.39	0.60
I_AGECHF	-0.0171	0.0001	0.98	0.98	0.99
I_AGEFEM	-0.0062	0.0028	0.99	0.99	1.00
I_AGEINF	0.0119	0.0001	1.01	1.01	1.02
I_AGESIT	-0.0204	0.0001	0.97	0.98	0.99
I_ANTCHF	-0.4895	0.0001	0.54	0.61	0.69
I_CHFCAB	0.4119	0.0001	1.26	1.51	1.81
I_CHFFEM	-0.2485	0.0001	0.71	0.78	0.86
I_COAPVE	-0.4805	0.0005	0.47	0.62	0.81
I_COMAGE	-0.0281	0.0001	0.96	0.97	0.99
I_COMASP	-1.1561	0.0001	0.19	0.31	0.52
I_COMCHF	-0.7839	0.0001	0.33	0.46	0.64
I_COMEPI	-0.8070	0.0022	0.27	0.45	0.75
I_COMOTH	-0.4938	0.0380	0.38	0.61	0.97
I_COMPUL	-0.9710	0.0001	0.26	0.38	0.55
I_EPIAGE	-0.0307	0.0001	0.96	0.97	0.98
I_INFCHF	-0.2661	0.0001	0.67	0.77	0.88
I_PULANT	-0.1404	0.0450	0.76	0.87	1.00

Table 9.3 Acute myocardial infarction mortality Model B, cases with no prior admissions (N=95,755), *continued*

*Variable**	*Parameter Estimate*	*p value*	*Lower CI for Odds Ratio*	*Odds Ratio*	*Upper CI for Odds Ratio*
I_PULASP	-0.7070	0.0001	0.37	0.49	0.65
I_PULCHF	-0.7712	0.0001	0.40	0.46	0.54
I_SHOACR	-0.5076	0.0001	0.48	0.60	0.75
I_SHOAMI	-1.2452	0.0001	0.16	0.29	0.52
I_SHOASP	-1.2394	0.0001	0.21	0.29	0.41
I_SHOCHF	-0.8068	0.0001	0.39	0.45	0.52
I_SHOCOA	-1.0001	0.0001	0.30	0.37	0.46
I_SHOCOM	-1.1018	0.0001	0.20	0.33	0.54
I_SHOOTH	-1.2576	0.0001	0.20	0.28	0.41
I_SHOPUL	-1.1302	0.0001	0.27	0.32	0.38
I_SHOPVE	-0.4353	0.0001	0.54	0.65	0.78
I_SITCHF	-0.7491	0.0001	0.40	0.47	0.55

* For the full name and ICD-9-CM description of each clinical variable, see Table 7.3. For the description of each demographic variable, see the Appendix in the *User's Guide.*

Table 9.4: Acute myocardial infarction mortality Model B, cases with one or more prior admissions (N=19,803)

*Variable**	*Parameter Estimate*	*p value*	*Lower CI for Odds Ratio*	*Odds Ratio*	*Upper CI for Odds Ratio*
INTERCEP	-6.9519	0.0001			
FEMALE	0.3028	0.0001	1.17	1.35	1.57
IAGEYRS	0.0531	0.0001	1.05	1.05	1.06
RACE_BLK	-0.2079	0.0142	0.69	0.81	0.96
RACE_HIS	-0.0293	0.6883	0.84	0.97	1.12
PAY_MCAL	-0.0180	0.8413	0.82	0.98	1.17
PAY_UNIN	0.2068	0.1786	0.91	1.23	1.66
IADM91	0.0500	0.3270	0.95	1.05	1.16
IADM92	0.0584	0.2472	0.96	1.06	1.17
PRIOLAGW	-0.0197	0.0001	0.98	0.98	0.99
ACRENALI	1.1565	0.0001	2.60	3.18	3.88
ASPPNEUI	0.9056	0.0001	1.83	2.47	3.34
ATYP_ER	0.3558	0.0001	1.31	1.43	1.55
CHFB	1.7825	0.0001	3.29	5.94	10.75
CHRRENAB	0.4645	0.0001	1.41	1.59	1.80
CNSDISB	0.3750	0.0017	1.15	1.45	1.84
COMAI	2.1932	0.0001	6.15	8.96	13.06
DBTCMPB	0.1061	0.0555	1.00	1.11	1.24
EPILEPB	0.6391	0.0001	1.56	1.89	2.31
HRSECMAB	0.8460	0.0001	1.92	2.33	2.82
HTB	-0.2788	0.0001	0.70	0.76	0.82
OTHCVAI	1.0091	0.0001	2.20	2.74	3.42
PRCABG	0.0140	0.8100	0.90	1.01	1.14
PULEDEMI	0.7715	0.0001	1.88	2.16	2.48
PVENTACI	0.3158	0.0001	1.20	1.37	1.57
SEPSISI	0.9800	0.0001	2.04	2.66	3.48
SHOCKI	2.9925	0.0001	15.31	19.94	25.97
SITE_ANT	0.9863	0.0001	2.42	2.68	2.97
SITE_INF	0.7476	0.0001	1.87	2.11	2.38
SITE_OI	1.9002	0.0001	5.45	6.69	8.21
SKNULCRP	0.4977	0.0001	1.30	1.64	2.08
I_AGECHF	-0.0142	0.0005	0.98	0.99	0.99
I_CHFFEM	-0.3858	0.0001	0.57	0.68	0.81
I_CHFSHO	-0.8356	0.0001	0.32	0.43	0.59
I_SHOACR	-0.5108	0.0166	0.40	0.60	0.91
I_SHOASP	-1.5731	0.0001	0.10	0.21	0.43
I_SHOCOM	-1.5178	0.0011	0.09	0.22	0.54
I_SHOOTH	-0.8453	0.0460	0.19	0.43	0.98
I_SHOPUL	-0.8968	0.0001	0.30	0.41	0.56
I_SHOSEP	-1.5758	0.0001	0.12	0.21	0.37
I_SITCHF	-0.5388	0.0001	0.46	0.58	0.73

* For the full name and ICD-9-CM description of each clinical variable, see Table 7.3. For the description of each demographic variable, see the Appendix in the *User's Guide.*

Chapter 10

Testing the Internal Validity of Risk-Adjustment Models

For this study, the internal validity of a risk adjustment model is defined as how well it controls for differences in patient characteristics that would otherwise confound outcome comparisons across hospitals. A model that does not adequately control for such differences may generate biased and misleading estimates of risk-adjusted mortality rates. The internal validity of the risk-adjustment models presented in Chapter Nine was assessed in four basic ways: content validity, construct validity, discrimination, and calibration.

New in 1997

- Model A continues to demonstrate better calibration but poorer discrimination than Model B. In other words, Model A provides unbiased estimates of patients' probability of death while Model B provides better separation between patients who died and survivors.

- In this report, the calibration of Model B was improved by a special adjustment procedure. This procedure makes Model B almost as well calibrated as Model A.

Content Validity

In previous years, all risk-adjustment models were carefully reviewed with members of the AMI clinical advisory panel and outside consultants. The advisory panel included several cardiologists, one nurse researcher, and one coding professional with specialized expertise in the topic. They advised project staff about whether the models included appropriate covariates and whether the parameter estimates were consistent with previous research and experience in the field. The advisory panel was not reconvened this year because the risk-adjustment procedure was thoroughly refined and validated in 1995 and 1996. However, the same criteria previously advocated by the AMI clinical advisory panel were applied by project staff in 1997 to ensure the face validity of all models. After removing several variables that were counterintuitively associated with lower mortality, the remaining risk factors demonstrated the expected effects on AMI mortality.

Discrimination

A model that distinguishes well between individuals who have poor outcomes and those who have good outcomes has excellent discrimination. A model with perfect discrimination would assign to every patient an expected

probability of either zero or one; all persons with an expected probability of one, but no one with an expected probability of zero, would experience the outcome of interest. No model has perfect discrimination in the real world, but good models show substantial spread in the expected probability of the outcome (death) between those who actually experienced it and those who did not.

The most commonly used measure of discrimination is the c statistic, which represents the proportion of all randomly selected pairs of observations with different outcomes (e.g., one death and one survivor) in which the patient who died had a higher expected probability of death than the survivor.[16] The c statistic takes on values between 0 and 1.0; higher values indicate greater discrimination but there is no cutoff that distinguishes "adequate" from "inadequate" models. A value of 0.5 can be obtained by random selection.

Table 10.1 shows that the primary risk models for AMI mortality have c statistics of 0.773 for cases with no prior admissions and 0.749 for cases with one or more prior admissions.[17] These c statistics are based on Model A, which omitted demographic and clinical risk factors that may reflect quality of care. As expected, Model B shows greater discrimination than Model A, with c statistics of 0.854 for cases with no prior admissions and 0.821 for cases with one or more prior admissions. This difference between the results for Model A and Model B is largely attributable to two powerful predictors that were used only in Model B: shock and pulmonary edema. These predictors were omitted from Model A because they may represent either in-hospital complications or associated conditions present on admission.

It is difficult to compare the performance of these risk models with that of models developed by other agencies evaluating hospital outcomes, because of obvious differences in methods. Pennsylvania's Health Care Cost Containment Council reported a c statistic of 0.88, using MedisGroups plus administrative data elements in a specially designed model to predict acute myocardial infarction mortality.[18] Cleveland Health Quality Choice has a very detailed data set with extensive clinical data; it reported c statistics of 0.85 to 0.92 from five risk-adjusted mortality models (including 0.89 for AMI cases).[19] Using clinical data on coronary bypass patients from New York's Cardiac Surgery Reporting System, Hannan et al reported a c statistic of 0.787.[20] By comparison, the best he could achieve using administrative data for the same patients was c = 0.74.[21] Using Medicare claims from 84 randomly

16. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; 143:29-36. The c statistic is equivalent to the area under a receiver operating characteristic curve, which represents a plot of sensitivity versus 1-specificity at various cutoff values for the predicted probability.
17. These statistics are based on the complete 100 percent sample. A stricter test of model discrimination comes from applying a regression equation estimated using 60 percent of the cases to the remaining 40 percent validation sample. The resulting c statistics, shown in Table 10.1, are virtually identical to those based on the total sample.
18. Focus on Heart Attack in Pennsylvania, Research Methods and Results. Harrisburg, PA: The Pennsylvania Health Care Cost Containment Council, April 1996.
19. Quality Information Management Corporation. *Cleveland-Area Hospital Quality Outcome Measurements and Patient Satisfaction Report*. Volume II. Cleveland, OH: Spring 1994.
20. Hannan EL, Kilburn H, Racz M, Shields E, Chassin MR. Improving the outcomes of coronary artery bypass surgery in New York State. *JAMA* 1994; 271:761-766.
21. Hannan EL, Kilburn H Jr, Lindsey ML, Lewis R. Clinical versus administrative data bases for CABG surgery: Does it matter? *Medical Care* 1992; 30:892-907.

selected US hospitals to predict 30-day mortality, Krakauer et al reported a c statistic of 0.84.[22] This model was similar to that used by the Health Care Financing Administration to generate its reports on Medicare hospital mortality.

Three recent studies have compared the ability of various severity indices to predict in-hospital mortality for AMI patients. The oldest study included 775 patients treated either medically or surgically at 12 New Orleans hospitals in 1985. The following c statistics were reported: 0.66 for Disease Staging (a proprietary system that uses administrative data to stage the complexity of illness), 0.74 for Patient Management Categories (another proprietary system that uses administrative data to predict resource utilization), 0.70 for APACHE II, and 0.73 for MedisGroups.[23] Researchers at Queens University[24] used commercial risk adjustment systems to predict 30-day and 60-day mortality among Medicare beneficiaries from six states in 1984-1985. For AMI patients, they reported the following c statistics: 0.750 for Computerized Severity Index, 0.725 for MedisGroups, 0.663 for Patient Management Categories, 0.623 for initial APACHE II, 0.512 for Disease Staging, and 0.464 for Acuity Index Method. By artificially dichotomizing each severity scale, these researchers underestimated scale performance. The best study included 11,880 adults managed medically for acute myocardial infarction at 108 acute care hospitals in 1991.[25,26] The following c statistics were reported: 0.862 for Disease Staging, 0.842 for All-Patient Refined Diagnosis Related Groups (a proprietary modification of the Health Care Financing Administration's DRG system), 0.834 for MedisGroups, and 0.832 for the Acute Physiology Score from APACHE III.

Finally, OSHPD's AMI validation study found that adding nine clinical variables derived from chart abstraction (e.g., systolic blood pressure and heart rate at presentation, shock at presentation, cardiac arrest within the previous 24 hours, do-not-resuscitate order written on the day of admission, peak CK ratio, pulmonary rales or loud systolic murmur at presentation, prior history of stroke) to the risk-adjustment model published by the California Hospital Outcomes Project in 1993 would improve the Model A c statistic from 0.782 to 0.854 and the Model B c statistic from 0.837 to 0.877. However, eliminating risk factors that were actually diagnosed after admission lowered the peak c statistic to 0.859.

These findings are very consistent with data from another recent study involving 30 Cleveland-area hospitals. Pine and colleagues[27] compared the performance of two models based on administrative data, similar to OSHPD's

22. Krakauer H, Bailey RC, Skellan KJ, et al. Evaluation of the HCFA model for the analysis of mortality following hospitalization. *Health Services Research* 1992; 27:317-335.
23. Alemi F, Rice J, Hankins R. Predicting in-hospital survival of myocardial infarction. *Medical Care* 1990; 28:762-775.
24. Case Mix Research, Queens University. Patient Classification Systems: An Evaluation of the State of the Art. Volume I. Springfield, VA: National Technical Information Service. 1991.
25. Iezzoni LI, Ash AS, Shwartz M, Daley J, Hughes JS, Mackiernan YD. Predicting who dies depends on how severity is measured: Implications for evaluating patient outcomes. *Annals of Internal Medicine* 1995; 123:763-770.
26. Iezzoni LI, Ash AS, Shwartz M, Daley J, Hughes JS, Mackiernan YD. Judging hospitals by severity-adjusted mortality rates: The influence of the severity-adjustment method. American Journal of Public Health 1996; 86:1379-1387.
27. Pine M, Norusis M, Jones B, Rosenthal GE. Predictions of hospital mortality rates: A comparison of data sources. *Annals of Internal Medicine* 1997; 126:347-354.

Model A and Model B, with two models that include additional clinical and laboratory data. The latter models included only ICD-9-CM defined risk factors that were likely to have been present at admission. For AMI cases, the Cleveland version of OSHPD's Model A had a c statistic of 0.76, whereas their version of Model B had a c statistic of 0.91. Their best model using additional clinical and laboratory data, similar to OSHPD's AMI validation model, had a c statistic of 0.88.

This summary demonstrates that the risk models developed as part of the California Hospital Outcomes Project compare favorably with others based on administrative data, but are probably inferior to those based on more detailed clinical data (e.g., Cleveland Health Quality Choice). This comparison is complicated by the fact that some risk-adjustment models, including MedisGroups and AMI Model B in this report, include clinical characteristics that represent complications of care rather than comorbidities.

Calibration and Bias

Calibration is the extent to which observed outcome rates correspond to predicted rates across a set of defined strata. A well-calibrated model demonstrates excellent fit across a broad range of patient characteristics. Calibration may be a more relevant measure than discrimination when the purpose of a model is to predict outcome rates for groups of persons with similar characteristics (e.g., inpatients at the same hospital). By contrast, discrimination is more important if a model is being used to predict an individual's outcome and to make treatment decisions. The most commonly used measure of calibration is Hosmer and Lemeshow's chi square test,[28] which compares observed with predicted outcomes across several strata (e.g., 10) that are defined by increasing levels of risk.

Table 10.1 shows that the primary risk model for AMI mortality, Model A, has a non-significant Hosmer-Lemeshow statistic among cases with no prior admissions (χ^2=8.67, p=0.37) and a marginally significant Hosmer-Lemeshow statistic among cases with one or more prior admissions (χ^2=16.97, p=0.030). These statistics are based on the complete 100 percent sample. A better test of model calibration comes from applying a regression equation estimated using 60 percent of the cases to the remaining 40 percent validation sample. This procedure generated nonsignificant Hosmer-Lemeshow statistics (Table 10.1), demonstrating that the model does not suffer from systematic lack of fit.

Model B demonstrates poorer calibration, as the Hosmer-Lemeshow chi square statistics are 46.20 (p<0.0001) among cases with no prior admissions and 30.41 (p=0.0002) among cases with one or more prior admissions. These statistics are based on the complete 100 percent sample. When the regression equation estimated using 60 percent of the cases was applied to

28. Hosmer DW, Lemeshow S. Applied Logistic Regression. New York: John Wiley & Sons, 1989.

the remaining 40 percent validation sample, the Hosmer-Lemeshow statistics were marginally significant among cases with no prior admissions (p=0.022) and nonsignificant among cases with one or more prior admissions (p=0.31). Model B has consistently poor calibration among cases with no prior admissions because it overestimates the probability of death among the lowest-risk and highest-risk patients. Attempts were made to correct this problem by testing additional interaction terms, as described in Chapter Eight, but this effort had limited success. Although Model B's calibration is now substantially better than it was last year (total sample χ^2 = 46.20 instead of 65.22), it still fails to meet conventional standards.

This problem was remedied by post hoc adjustment of the linear predictor from Model B, using a quadratic function. A second-order, patient-level logistic model was estimated, in which the dependent variable was the observed outcome (e.g., death) and the independent variables were the logit and logit squared of the estimated probability of death:

$$\log\left(\frac{\hat{p}^*}{1-\hat{p}^*}\right) = \hat{\gamma}_0 + \hat{\gamma}_1 z + \hat{\gamma}_2 z^2$$

In this equation, $\hat{p}^*$ is the adjusted estimate of the probability of death and z is the logit of the original estimate of that probability, $\hat{p}$, derived from the logistic models shown in Tables 9.3 and 9.4. As shown on page 76,

$$z = \log\left(\frac{\hat{p}}{1-\hat{p}}\right) = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \ldots + \hat{\beta}_q x_q$$

A perfectly calibrated model would have $\hat{\gamma}_0 = 0$, $\hat{\gamma}_1 = 1$ and $\hat{\gamma}_2 = 0$,[29] so that:

$$\log\left(\frac{\hat{p}^*}{1-\hat{p}^*}\right) = z = \log\left(\frac{\hat{p}}{1-\hat{p}}\right)$$

When this second-order model was estimated using the probability estimates from Model B, the calibration adjustment coefficients shown in Table 10.1 were generated. The intercept $\hat{\gamma}_0$, was not statistically different from zero, but the linear term $\hat{\gamma}_1$ was significantly less than one (95% confidence interval, 0.863 to 0.959 among cases with one or more prior admissions and 0.904 to 0.958 among cases with no prior admissions), and the quadratic term $\hat{\gamma}_2$ was significantly less than zero (95% confidence interval, -0.071 to -0.030 among cases with one or more prior admissions and -0.036 to -0.019 among cases without prior admissions). As a result, the estimated probability of death was adjusted slightly downward among especially low-risk and high-risk patients, but slightly upward among intermediate risk patients. The

29. Miller ME, Langefeld CD, Tierney WM, Hui SL, McDonald CJ. Validation of probabilistic predictions. *Medical Decision Making* 1993; 13:49-58.

definition of intermediate risk can be determined by solving both quadratic equations: 11.6 percent to 57.5 percent among cases with one or more prior admissions; 7.3 percent to 51.8 percent among cases with no prior admissions.

Adjusting the Model B probability estimates in this manner substantially improved calibration, although the Hosmer-Lemeshow chi-square statistic remained significant at 25.21 ($p = 0.0014$) among cases with no prior admissions. As noted above, the Model A probability estimates did not require adjustment because the calibration of Model A was excellent. The consistent difference in calibration between Model A and Model B is probably due to multi-way interactions involving the major clinical risk factors in Model B (e.g., shock, pulmonary edema) that are complex and difficult to model.

Bias tests also were performed for a variety of other patient characteristics that were deliberately omitted from the risk-adjustment models or specified in a particular manner. None of these models shows significant ($p<0.01$) bias related to race, year, or quarter of admission. AMI Model A shows significant ($p<0.000001$) bias related to the source and type of admission, and the expected source of payment, because these variables were deliberately omitted from Model A. AMI Model B shows significant ($p<0.000001$) bias across 10-year age categories, reflecting underprediction of mortality among persons 35 to 44 years of age. This finding is probably explained by complex interaction or confounding effects involving the spline function used to model the association between age and mortality.

Bias testing therefore confirmed that, with minor exceptions, the risk-adjustment models developed for the California Hospital Outcomes Project are relatively free from bias due to temporal and demographic factors. Substantial bias due to unmeasured clinical factors is likely, and was further characterized in the 1996 *Technical Appendix*, which presented results from the AMI Validation Study.

Table 10.1: Goodness-of-fit tests for AMI mortality models

	Priors		*No Priors*	
	Model A	*Model B*	*Model A*	*Model B*
Number of cases	19,882	19,803	96,292	95,755
Number of deaths	4,037	4,021	12,797	12,731
Death rate, percent	20.37	20.31	13.29	13.30
Model chi square	2,408.86	4,854.24	10,415.38	25,126.77
df	19	40	25	63
p value	<0.0001	<0.0001	<0.0001	<0.0001
C statistic				
Estimation sample	0.748	0.822	0.773	0.853
Cross validation sample	0.751	0.819	0.774	0.855
Total sample	0.749	0.821	0.773	0.854
Calibration adjustment coefficients				
B_0 (intercept)		0.032		0.005
B_1 (intercept)		0.910		0.931
B_2 (intercept)		-0.052		-0.028
Hosmer Lemeshow statistic				
Estimation sample (df=8)	20.90*	19.78	8.28	20.14*
Cross validation sample (df=10)	10.35	11.57	11.96	20.82
Total sample (df=8)	16.97	30.41**	8.67	46.20**
Total sample, adjusted (df=8)		16.18		25.21*

*p<0.01
**p<0.001

Chapter 11

Calculation of Hospital Outcome Measures

The risk adjustment models described in Chapter Nine were used to calculate several hospital outcome measures. The actual values of these measures, by year and overall, are reported in the Detailed Statistical Results. The User's Guide classifies all hospitals treating AMI patients as "significantly better than expected," "significantly worse than expected," or "not significantly different than expected" based on the exact probability of the observed number of AMI deaths (or a more extreme number) at that hospital. It also includes a chart showing each hospital's risk-adjusted death rate with 98 percent confidence limits, based on aggregated 1991-1993 data. Each of these outcome measures is described below, along with the methods used to calculate it.

New in 1997

- This report includes both 98 percent and 95 percent confidence limits for hospitals' risk-adjusted death rates; the 98 percent confidence limits based on aggregated 1991-1993 data appear in the *User's Guide* while the 95 percent confidence limits based on individual years of data appear in the *Detailed Statistical Results*.

- In previous reports, confidence intervals were constructed by estimating the variance of observed mortality at each hospital. In this report, the variance of expected mortality based on risk-adjustment models was also estimated, but was found not to affect the classification of hospital performance.

Number of Observed Deaths and Observed Death Rate

The number of observed deaths at a hospital is simply the total number of deaths within 30 days of admission, among qualifying AMI patients. The death may have occurred at the index hospital, a transfer hospital, or outside the hospital setting. The observed death rate at a hospital equals the number of observed deaths, divided by the total number of qualifying patients at that hospital. This quantity was multiplied by 100 to yield a percentage.

The distribution of observed death rates among eligible hospitals with at least one expected death is shown in Figure 11.1. The height of each bar represents the number of hospitals with observed death rates in the specified range.

Number of Expected Deaths and Expected Death Rate

The number of expected deaths at a hospital equals the sum of the estimated probabilities of death for all of its qualifying patients. These estimated probabilities were calculated using the logistic formulas in Chapter Nine; Model B estimates were adjusted slightly to improve calibration, as described in Chapter Ten. For example, the number of expected AMI deaths would be 5 if a hospital had 10 patients, each of whom had a 50 percent risk of death, or if a hospital had 100 patients, each of whom had a 5 percent risk of death.

The expected death rate at a hospital equals the number of expected deaths, divided by the total number of qualifying patients at that hospital. This quantity was multiplied by 100 to yield a percentage. The expected death rate also represents the mean estimated probability of death for all patients at a hospital, which is a measure of average severity of illness. If a hospital's expected death rate is higher than the statewide death rate, then patients at that hospital tend to be higher risk than the statewide average. If a hospital's expected death rate is lower than the statewide death rate, then patients at that hospital tend to be lower risk than the statewide average.

The distribution of expected death rates among eligible hospitals with at least one expected death is shown in Figure 11.2. The height of each bar represents the number of hospitals with expected death rates in the specified range.

Risk-Adjusted Death Rate

The risk-adjusted (or indirectly standardized) death rate at a hospital equals the statewide rate, multiplied by the ratio of the number of observed deaths to the number of expected deaths at that hospital:[30]

$$I_i = s\left(\frac{\sum_{j=1}^{n_i} o_j}{\sum_{j=1}^{n_i} \hat{p}_j}\right) = s\frac{O_i}{\pi_i}$$

where I_i is the indirectly standardized outcome rate for the ith hospital, s is the statewide outcome rate, o_j is the observed value of the adverse outcome (0 or 1) for the jth patient, and $\hat{p}_j$ is the estimated probability of the adverse outcome for the jth patient. The latter two variables are summed over all patients at the ith hospital.

This risk-adjusted death rate provides a basis for comparing the performance of different hospitals, because each hospital's rate is adjusted to reflect what

30. Williams RL. Measuring the effectiveness of perinatal medical care. *Medical Care* 1979; 17:95-110.

its death rate would be if its patients were about as ill as the statewide average. The ratio of the number of observed deaths to the number of expected deaths at a hospital provides a quick assessment of that hospital's performance. For a hospital with fewer observed than expected deaths, this ratio is less than one; for a hospital with more observed than expected deaths, this ratio is greater than one.

The distribution of risk-adjusted death rates among eligible hospitals with at least one expected death is shown in Figure 11.3. The height of each bar represents the number of hospitals with risk-adjusted death rates in the specified range. The distribution of risk-adjusted death rates is tighter than the distribution of observed death rates, indicating that risk-adjustment reduces some of the apparent variability in hospital performance.

Confidence Limits for Risk-Adjusted Death Rates

The 95 percent and 98 percent confidence limits reflect the level of confidence in a hospital's risk-adjusted death rate. In general, when the upper and lower confidence limits are far apart, the estimated risk-adjusted death rate is unreliable. Assuming that the risk model is correct, there is a 95 percent chance that a hospital's true risk-adjusted death rate falls within the 95 percent confidence limits, and a 98 percent chance that this value falls within the 98 percent confidence limits. The narrower 95 percent confidence limits are used in the *Detailed Statistical Results*, for the benefit of individual hospitals and physician groups that wish to evaluate their own performance. Wider 98 percent confidence limits are used in the *User's Guide*, because of the large number of hospitals evaluated in the study increases the risk of mislabeling a hospital as an outlier.

These 95 percent and 98 percent confidence limits were constructed from the standard deviation of the number of observed deaths at each hospital:

$$\text{Lower CI}(I_i) = \frac{s}{\pi_i}\text{Max}\left(0, O_i - z_{1-\alpha/2}\left(\sum_{j=1}^{n_i} p_j(1-p_j)\right)^{1/2}\right)$$

$$\text{Upper CI}(I_i) = \frac{s}{\pi_i}\text{Min}\left(n_i, O_i + z_{1-\alpha/2}\left(\sum_{j=1}^{n_i} p_j(1-p_j)\right)^{1/2}\right)$$

where I_i, O_i, p_i, and $\hat{p}_j$ are defined as before. For the 95 percent and 98 percent confidence limits, respectively, $z_{1-a/2}$ takes on values of 1.9600 and 2.3263. The lower confidence limits are constrained so they do not fall below 0 percent; and likewise the upper confidence limits are constrained not to exceed 100 percent. As a result of these constraints, the confidence limits for very small hospitals may not truly provide 95 percent or 98 percent coverage.

In calculating these confidence limits, the estimated probability of death for each case, and hence the expected death rate at each hospital, were treated as fixed quantities. The estimated probabilities were derived from logistic regression models that included all eligible patients in California. With such large samples, prediction error is generally negligible relative to the random

error that results from the small number of cases at each hospital.[31] The statewide death rate was also treated as a fixed quantity. Therefore, the confidence intervals were constructed around the observed death rate, which was treated as a random variable. Because cases at the same hospital show considerable variability in their probabilities of death, the above formula includes the estimated probabilities for individual patients rather than a hospital's average probability.

The assumption that prediction error is negligible was specifically tested this year by estimating this component of variance, using an adaptation of a recently reported technique.[32] Let i be the index for hospitals; $j = 1,\ldots,n_i$ be the index for heart attacks within hospital i; $\pi_i = \sum_j \hat{p}_j$ where $\hat{p}_j$ is the predicted probability of mortality; $\mathbf{b}$ be the $q \times 1$ vector of parameter estimates, $(\hat{\beta}_1,\ldots,\hat{\beta}_q)$; $\mathbf{X}_i$ be the $n_i \times q$ matrix of risk factor values; and $\mathbf{v}_i$ be an $n_i \times 1$ vector with the jth value as $\hat{p}_j(1-\hat{p}_j)$, so that $\mathbf{V}_i = \mathbf{I}_{n_i \times n_i}\mathbf{v}_i$. Then, using a Taylor series approximation of the variance of a function of a random variable, the variance of π_i is

$$\mathrm{Var}(\pi_i) = \left(\frac{\partial \pi_i}{\partial \mathbf{b}}\right)' \mathrm{Cov}(\mathbf{b}) \left(\frac{\partial \pi_i}{\partial \mathbf{b}}\right)$$

Using the fact that

$$\frac{\partial \hat{p}_j}{\partial \beta_k} = \frac{\partial}{\partial \beta_k}\left\{\frac{\exp(\mathbf{x}_j'\mathbf{b})}{1+\exp(\mathbf{x}_j'\mathbf{b})}\right\} = x_{jk}\hat{p}_j(1-\hat{p}_j)$$

gives

$$\frac{\partial \pi_i}{\partial \mathbf{b}} = \left(\sum_j \frac{\partial \hat{p}_j}{\partial \beta_1},\ldots,\sum_j \frac{\partial \hat{p}_j}{\partial \beta_q}\right)'$$

So now the variance of π can be written as

$$\mathrm{Var}(\pi_i) = \mathbf{v}_i'\mathbf{X}_i(\mathbf{X}_i'\mathbf{V}_i\mathbf{X}_i)^{-1}\mathbf{X}_i'\mathbf{v}_i$$

Now let $O_i = \sum_j o_j$.

So

$$\begin{aligned}
\mathrm{Var}(O_i/\pi_i) &= \mathrm{Var}\{\exp[\log(O_i/\pi_i)]\} \\
&= \{\exp[\log(O_i/\pi_i)]\}^2 \mathrm{Var}[\log(O_i/\pi_i)] \\
&= (O_i/\pi_i)^2 \mathrm{Var}[\log(O_i) - \log(\pi_i)] \\
&= (O_i/\pi_i)^2 \{\mathrm{Var}[\log(O_i)] + \mathrm{Var}[\log(\pi_i)]\} \\
&= (O_i/\pi_i)^2 \left[\frac{\mathrm{Var}(O_i)}{O_i^2} + \frac{\mathrm{Var}(\pi_i)}{\pi_i^2}\right]
\end{aligned}$$

31. Health Care Financing Administration. *Medicare Hospital Mortality Information, 1988-1989-1990, Volume 55*. Washington, D.C.: US Government Printing Office.
32. Hosmer DW, Lemeshow S. Confidence interval estimates of an index of quality performance based on logistic regression models. *Statistics in Medicine* 1995; 14:2161-2172.

where $\mathrm{Var}(O_i) = \mathbf{1}_{1\times n_i}\,\mathbf{v}_i$ [33]

This component of variance was indeed found to be very small, and to have minimal effect on the confidence intervals estimated using the above formula. Indeed, none of the hospitals that are labeled as AMI mortality outliers in this report became non-outliers when these expanded confidence intervals were applied. There were no differences in the classification of hospital performance, using either 95% or 98% confidence limits. Therefore, the simpler and more straightforward formula on page 93 was retained.

Exact Probability of the Number of Observed Deaths

The exact probability of the number of observed deaths (or a more extreme number) occurring by chance, given the number of expected deaths at a hospital, was used to identify the outlier hospitals labeled with stars or circles in the *User's Guide*. This approach differs from the more widely used normal approximation in that it gives better estimates for hospitals with relatively few expected deaths.[34]

If the number of observed deaths exceeded the number of expected deaths, an upper probability (p) value was computed. If the number of observed deaths was less than or equal to the number of expected deaths, a lower probability (p) value was computed.

The upper p-value for a hospital is the probability that the observed number of deaths or more occurred by chance. The upper p-value represents a "test" of whether a hospital has systematically worse outcomes than the statewide average. A very small p-value of 0.001 means that one would expect to see this many deaths or more only 1 time in 1000, by chance. This finding leads one to reject the null hypothesis that the hospital's performance is equivalent to the statewide average. A more likely explanation would be a difference in quality of care, or some other systematic factor.

The lower p-value for a hospital is the probability that the observed number of deaths or fewer occurred by chance. The lower p-value represents a "test" of whether a hospital has systematically better outcomes than the statewide average. A very small p value again leads one to reject the null hypothesis that the hospital's performance is equivalent to the statewide average.

The classification of hospitals' AMI death rates as "significantly better than expected," "significantly worse than expected," or "not significantly different than expected" in the *User's Guide* was based on a p value threshold of 0.01. Hospitals classified as significantly better than expected had fewer deaths than expected and a lower p value less than 0.01. Hospitals classified as

33. Zhou H, Romano PS. Letters to the editor: Confidence interval estimates of an index of quality performance based on logistic regression models. *Statistics in Medicine* 1997; 16:in press.
34. Luft HS, Brown BW Jr. Calculating the probability of rare events: Why settle for an approximation? *Health Services Research* 1993; 28:419-439.

significantly worse than expected had more deaths than expected and an upper p value less than 0.01. The p value threshold of 0.01 was chosen instead of the more commonly used threshold of 0.05 (or 0.025) because the number of hospital outliers at the 0.05 (or 0.025) level did not significantly exceed the number expected under the null hypothesis that all hospitals have equivalent risk-adjusted death rates.

This report includes 418 eligible hospitals that contributed one or more years of data. Using Model A, 26 of these hospitals were classified as "significantly better than expected" and 21 were classified as "significantly worse than expected" based on their AMI mortality from 1991 through 1993. Using Model B, 22 hospitals were classified as "significantly better than expected" and 20 were classified as "significantly worse than expected" based on their AMI mortality from 1991 through 1993. Thirteen hospitals were rated "significantly better than expected" using both models, whereas thirteen achieved this rating using only Model A and nine achieved this rating using Model B. Ten hospitals were rated "significantly worse than expected" using both models, whereas eleven achieved this rating using only Model A and ten achieved this rating using Model B.

Figure 11.1: Distribution of Observed Death Rates Across California Hospitals

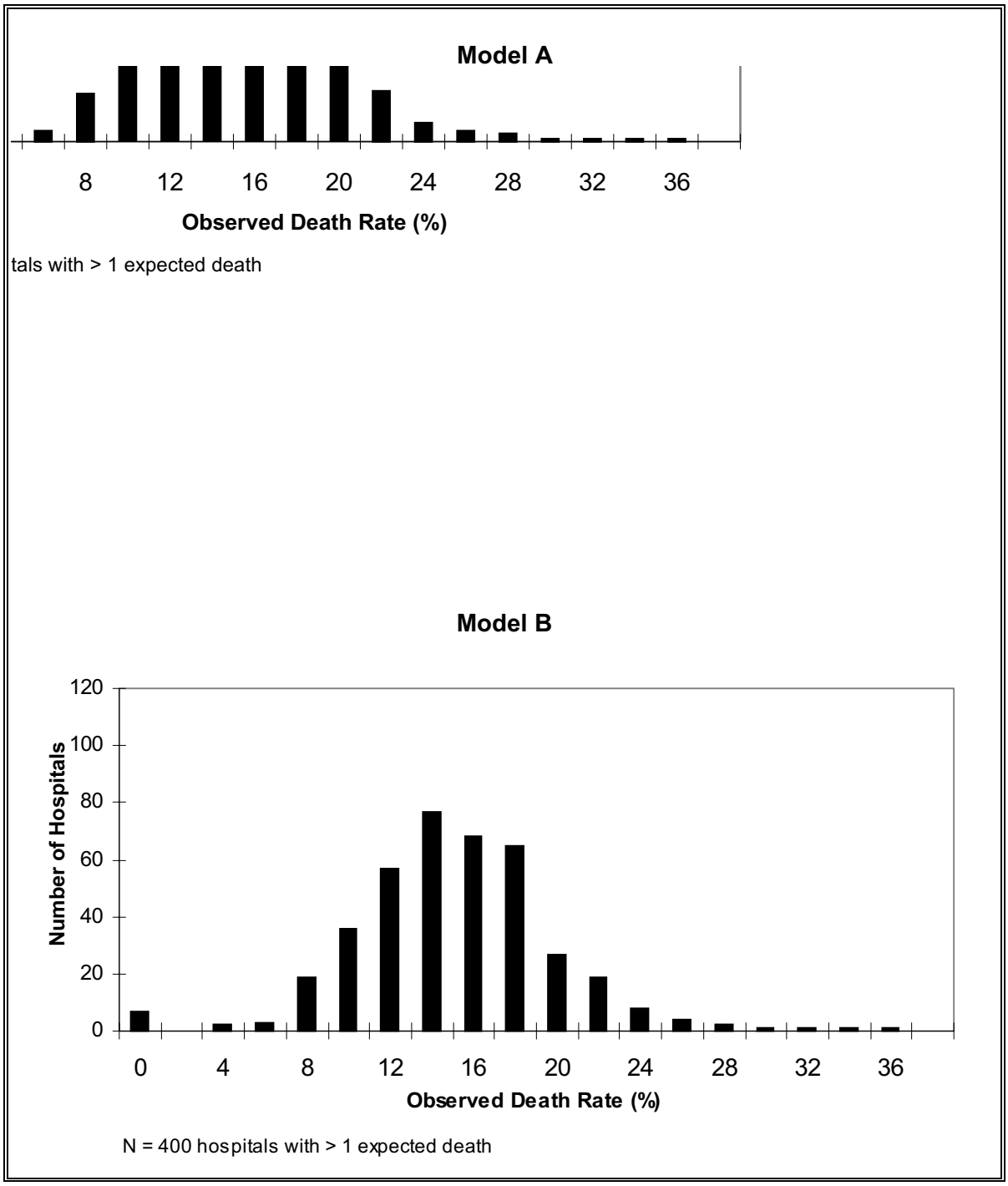

Figure 11.2: Distribution of Expected Death Rates Across California Hospitals

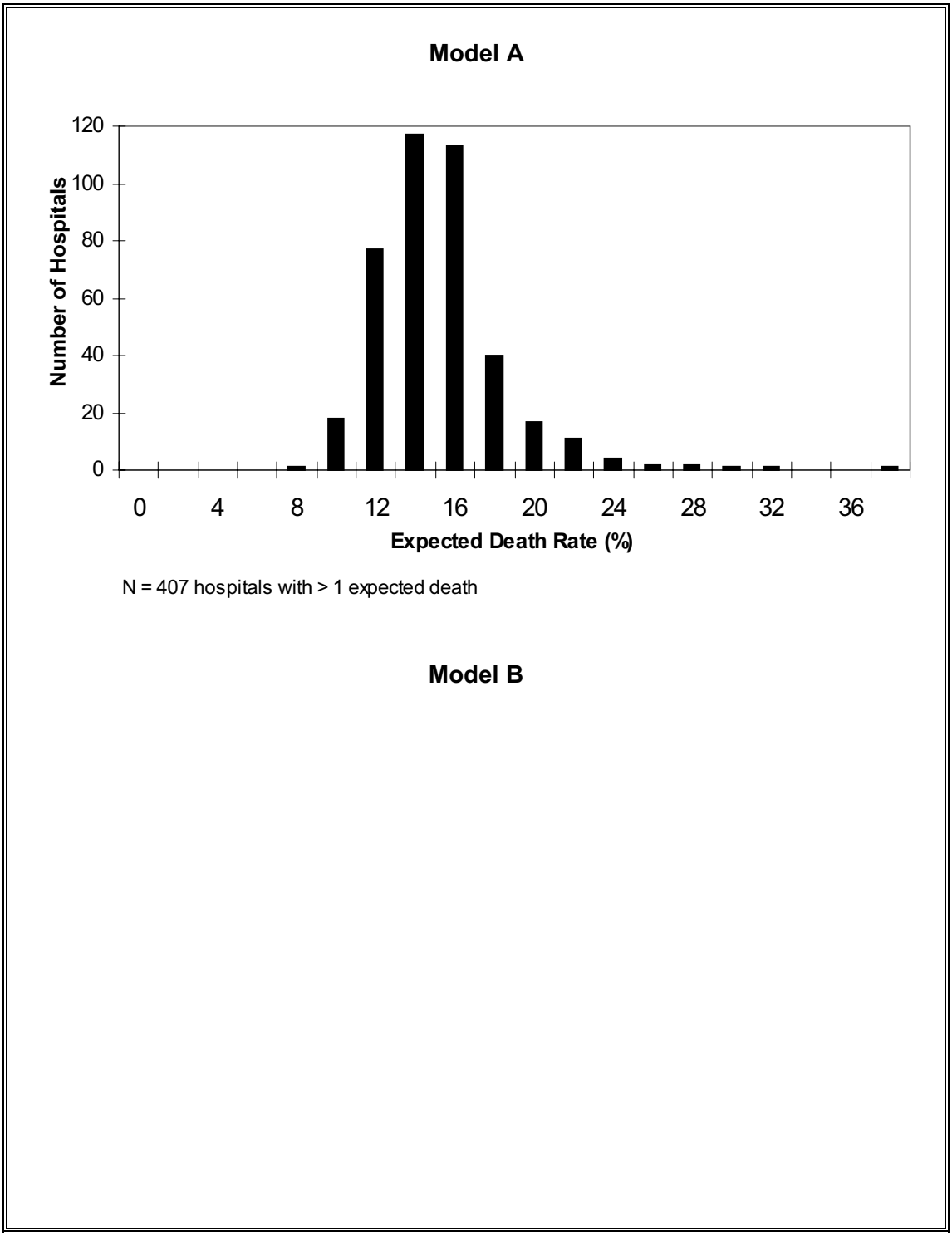

Figure 11.3: Distribution of Risk Adjusted Death Rates Across California Hospitals