# CHAPTER THIRTEEN:
# VALIDATION STUDY OF ACUTE MYOCARDIAL INFARCTION, METHODOLOGY

This chapter describes the methods used to validate the results of the 1993 California Hospital Outcomes Project for acute myocardial infarction. It details the sampling, data collection, and data analysis methodologies that were used.

## OVERVIEW

The AMI validation study was a retrospective cross-sectional study based on a stratified two-stage probability sample of AMI hospitalizations at medium to high volume, acute care hospitals in California. All AMI admissions between July 31, 1990 and May 31, 1991 that were included in OSHPD's 1993 California Hospital Outcomes Project report were eligible for sampling. At OSHPD's request, participating hospitals submitted a complete copy of each sampled record. Each record was exhaustively reviewed by both a medical records professional and a clinician (intensive care nurse or physician), who then entered all abstracted information into a computerized data entry system with built-in error checks and branching logic. To maximize the reliability and validity of abstraction, reviewers were given detailed written guidelines, received special training and on-site supervision, and were monitored through 5% overreading.

The data were cleaned and missing values of critical variables were filled in whenever possible. A variety of univariate, bivariate, and multivariate data analyses were performed, as described in the next chapter. Weighted analyses were performed when appropriate, to compensate for the oversampling of outlier hospitals and patients who died.

The entire study protocol was approved by the Human Subjects Review Committee at the University of California, Davis. Appropriate safeguards were established to ensure that all records are stored safely and are not accessible to persons outside the California Hospital Outcomes Project staff.

## POWER ANALYSIS

OSHPD's power analysis was based on Question 3 described in Chapter Twelve. The null hypothesis for this analysis is that key risk factors, such as congestive heart failure and anterior wall involvement, are coded with equal sensitivity at hospitals with high, average, and low risk-adjusted mortality. The alternative (two-tailed) hypothesis is that these risk factors are coded differently, corresponding to a hospital's risk-adjusted mortality classification. To achieve 80% power to detect a 20% difference in coding sensitivity (e.g., 60% versus 80%) with a type I error rate of 5%, each of the three hospital mortality classes

must have at least 78 patients with the risk factor of interest. Coding differences of less than 20% would be unlikely to cause significant bias in inter-hospital comparisons.

The sample size necessary to get 78 patients with the risk factor of interest in each comparison group varies according to the prevalence of the condition. Some risk factors, such as congestive heart failure (CHF) and anterior wall AMI, have high prevalences (29.8% and 31.0%, respectively). Other risk factors, such as complicated diabetes and prior coronary artery bypass grafting (CABG), are somewhat less frequent (7.5% and 7.4%, respectively). A sample of 300-350 patients in each of the three hospital mortality classes, or 1,000 patients overall, has more than adequate power to detect differences in the coding of high-prevalence risk factors. Although a sample of this size lacks the power to detect differences in the coding of low-prevalence risk factors, oversampling of deaths boosts the frequency of many of these factors.

## HOSPITAL SAMPLING

The 394 hospitals included in OSHPD's 1993 AMI report (model B) were stratified according to patient volume. Low volume hospitals with less than 50 AMI patients were not included in the sampling frame, because such hospitals virtually never meet the statistical thresholds to be labeled as mortality outliers. The remaining 228 hospitals were divided into two volume categories, which were designed to include equal numbers of outlier hospitals (approximately 17 in each category). Ninety-two hospitals with 118 or more AMI patients were designated as high-volume; 136 hospitals with 50-117 AMI patients were designated as medium-volume. Both groups were then stratified according to their risk-adjusted mortality classification (better than expected, worse than expected, or neither at $p<0.05$), as shown in Table 13.1. Each eligible hospital in a stratum was assigned a random number; six medium-volume and four high-volume hospitals were sampled from each of the three outcome strata. Fewer high-volume hospitals were sampled because more patients were available for sampling at each of these hospitals.

## PATIENT SAMPLING

The second stage of sampling involved sampling patients within each randomly selected hospital. Patients were stratified by outcome status (i.e., in-hospital death within 30 days of admission). Deaths were oversampled to ensure that each hospital stratum included a sufficient number of high-risk individuals, for whom coding errors and unmeasured risk factors might significantly affect the predicted probability of death. The statewide death rate for AMI during the study period was 13%; the target death rate in OSHPD's validation sample was twice that rate, or 26%. A sample of approximately 180 patients was drawn randomly from each of the six hospital volume-outcome strata, such that approximately 26% or 47 were deaths and 74% or 133 were survivors.

Calculating the number of cases to be selected from each hospital was done in two stages. First, the numbers of deaths and survivors in each stratum were totaled. Sampling fractions for the stratum were calculated by dividing the "target" number of sampled deaths for the stratum (i.e., 47) by the observed number of deaths. For example, a stratum with 1000 patients and 300 deaths has a death sampling fraction of 47/300 or 0.16. The same calculation yields a survivor sampling fraction of 133/700 or 0.19. These stratum-specific sampling fractions were applied to each sampled hospital to determine the target number of deaths and survivors to select from that hospital. For example, in a stratum for which the sampling fraction for deaths is 0.16 and that for survivors is 0.19, a hospital with 20 deaths and 80 survivors would have a target number of deaths equal to $n_d = 20*0.16 = 3.2$, and a target number of survivors equal to $n_s = 80*0.19 = 15.2$. All target numbers were rounded to the nearest integer.
After the target numbers of deaths and survivors were calculated for each hospital, all eligible cases at that hospital were sorted by a random number and the desired numbers of deaths and survivors were chosen for the study. The net effect of this procedure was to draw $n_i$ patients from the ith hospital, where $n_i = 180*(N_i/N_t)$, $N_i$ equals the total number of eligible patients in the hospital, and $N_t$ equals the total number of patients at all sampled hospitals in the stratum. Thus, the number of cases contributed by each hospital within a stratum was proportional to its volume.

## DATA COLLECTION

The medical record abstraction instrument included the following sets of data elements:

1. Core demographic variables for validating matches between the reabstract and the original discharge abstract (e.g., dates of admission and discharge, date of birth, social security number, gender), coded according to OSHPD guidelines;

2. ICD-9-CM diagnosis and procedure codes, coded in a consistent manner according to the guidelines published in *Coding Clinic for ICD-9-CM*, *Official Coding Guidelines*, and the American Hospital Association's *ICD-9-CM Coding Handbook*;

3. The date of each ICD-9-CM procedure and the date on which each ICD-9-CM diagnosis was first established, together with a dichotomous variable indicating whether that diagnosis was documented in a paramedic, emergency room, or admission note;

4. Additional clinical risk factors obtained from the admission histories and physical examinations, laboratory studies, radiographic and cardiac imaging studies, and electrocardiograms; and

5. The use of various therapies for acute coronary artery disease, including intravenous or intracoronary thrombolysis, aspirin, beta blockers,

subcutaneous or intravenous heparin, percutaneous angioplasty, and coronary arterial bypass grafting.

Extensive consultation with the AMI Clinical Advisory Panel took place during the design of this data collection instrument. The panel suggested various historical and physiological risk factors, and "process of care" variables, to retrieve from the medical records. Additional risk factors were identified by reviewing the biomedical literature for original, English-language papers describing the factors associated with short-term (inpatient or 30-day) mortality after AMI in large cohorts of hospitalized patients in the US or Western Europe. See Chapter Two for more information about this literature review.

A draft instrument and set of accompanying guidelines were written and distributed to all members of the research group and the AMI Clinical Advisory Panel. These guidelines were extremely detailed and included definitions of every medical term. They specified the allowable data sources for each question, the hierarchy of data sources if there was conflicting information, the allowable synonyms (e.g., crackles and rales), and any constraints on the timing of physical findings and laboratory values. The panel was asked to comment upon the utility of each proposed data element, the extent to which it would be available in the medical record, and the difficulty of abstracting it.

The California Peer Review Organization (California Medical Review, Inc. or CMRI) performed the actual abstraction of records. CMRI staff reviewed the abstraction tool to determine how often specific information is available in the medical record, where in the record it is typically located, and how difficult it is to abstract. The draft instrument and guidelines were pretested on 15 non-randomly sampled records from different hospitals. All problems recognized during pretesting were addressed. Finally, the instrument was programmed for direct data entry by chart reviewers. The program included skip patterns based on branching logic, default values, precoded response options, range checks on dates and physiologic variables, and numerous relational logic checks. As a result, reviewers were precluded from entering clearly invalid data. Extra fields were created for reviewers to write comments when they had difficulty reading or interpreting the medical record, or when they experienced any other problems. Coders and reviewers participated in a full day of training conducted by the contractors. At this training session, a final set of data entry guidelines was distributed, the data entry program was demonstrated, and questions were answered.

Upon receipt, each chart was logged in and verified as the correct record based on the patient's date of birth, gender, and social security number, and the dates of admission and discharge. Each chart was then reabstracted by skilled Accredited Records Technicians who were also Certified Coding Specialists with at least ten years of experience. The coders were blinded to the original discharge abstract. The coders wrote the ICD-9-CM diagnosis and procedure codes on a hard copy form, which remained with the chart. A clinician (e.g., intensive care nurse or physician) reviewer then verified and entered the ICD-9-CM codes and abstracted all of the clinical data elements. The coding and

review team spent an average of 90 to 120 minutes per chart. Supervisors overread 5% of the records. All coders and record reviewers maintained 95% or greater accuracy throughout the project.

## MEDICAL RECORD REQUESTS

In July 1994, the administrator and director of medical records of each sampled hospital were contacted by mail. This letter described the purpose of the validation study and the procedures for maintaining patient confidentiality, guaranteed anonymity for the hospital in all OSHPD publications and data releases, outlined OSHPD's expectations, and explained that participation was entirely voluntary. Each letter was accompanied by a letter of support from the California Association of Hospital and Health Systems (CAHHS). A follow-up telephone call was made within one week in order to obtain the administrator's verbal consent.

Once a hospital agreed to participate, a list of all sampled records was sent to the director of medical records. Each list contained the patients' admission and discharge dates, date of birth, gender, principal diagnosis, and social security number (if known). Medical record departments were given three weeks to locate, photocopy and mail all components of the requested records, including but not limited to: ambulance records, emergency room notes, admission notes, physician and nursing progress notes, nursing flow sheets or graphic records, physician orders, laboratory and radiology reports, electrocardiograms, medication administration records, demographic data forms, and consultation reports. Upon receipt of the records by CMRI, photocopying and mailing costs were reimbursed.

One of the six selected hospitals in the medium volume, better than expected mortality stratum declined to participate. The hospital's Institutional Review Board apparently reviewed and rejected OSHPD's request, despite the approval of the Human Subjects Review Committee at the University of California, Davis. There was only one potential replacement hospital in that stratum. Fortunately, the administrator of this hospital agreed to participate.

## RECORD SUBMISSION

Overall, 1005 of the 1065 requested records (94.4%) were submitted by participating hospitals. This percentage is slightly lower than that achieved in previous reabstraction studies by OSHPD and HCFA, presumably because OSHPD was unable to provide patient names or medical record numbers. The 60 unretrieved records were at 15 different hospitals, although 35 (58%) were concentrated at three hospitals ($p<0.01$). Twenty-two (37%) of the unretrieved records were from the first 60 days of the 10-month study period ($p=0.02$), suggesting that hospitals experienced more difficulty locating older records.

The number of records and submission rate in each sampling stratum is shown in table 13.2. The submission rate was unrelated to hospital volume but was better at hospitals with high risk-adjusted mortality than at hospitals with low or intermediate risk-adjusted mortality ($p<0.001$). This difference is reflected in the case weights used in reestimating risk adjustment models (see Chapter Fourteen). The submission rate was unrelated to race, insurance status, or in-hospital death, but records with missing social security numbers were less likely to have been submitted than other records (87% versus 95%, $p=0.002$).

There was a continuing problem throughout the project with missing or illegible electrocardiograms (ECGs). CMRI devoted considerable effort to contacting hospitals and requesting better copies of ECGs. After CMRI completed its data collection, the primary research team again contacted 18 hospitals to ask for better copies of 158 ECGs. Of these, 98 (62%) were supplied. This percentage probably reflects that the original ECGs could not be located and that copies could not be improved due to the deterioration of the ink on the original tracing. Two physicians from the primary research team reviewed these ECGs to complete the items on electrocardiography.

## DATA CLEANING AND ANALYSIS

Despite the edits built into the data entry program, some final data cleaning was necessary. The univariate distribution of each variable was examined; illogical values were checked against the original records and recoded when appropriate. Illogical combinations, such as cardiopulmonary resuscitation without a cardiopulmonary arrest and "routine" discharge within 24 hours of admission, were corrected. Discrepancies between ICD-9-CM procedure codes and corresponding data elements in the abstraction instrument, such as whether percutaneous coronary angioplasty or coronary bypass surgery was performed, were identified and reconciled. All cases with less than eight weeks between the index AMI admission and a reported prior AMI were reviewed; several of these reported prior AMIs were not supported by the medical record, so recoding was performed. All cases with major contraindications to thrombolytic and aspirin therapy were also reviewed; several cases of aspirin allergy were not supported by the medical record, so recoding was performed.

Special efforts were also made to fill in missing values wherever possible. For example, the upper limit of normal for CK was missing in numerous records, but was filled in using the value for other persons of the same sex in the same hospital during the same time period. Electrocardiograms were reviewed manually when necessary to correct illogical (e.g., normal sinus rhythm with a rate over 160) or missing interpretations. Missing fields related to chest radiography were recoded based on the arguable assumption that all positive findings were described in the radiologists' dictated notes.

Because of the complex sample structure, weighted analyses were performed when appropriate. The weight was defined as the inverse of the sampling probability, which was in turn calculated by multiplying the probability of sampling a specific hospital by the probability of sampling an individual within that hospital. The weights were adjusted to reflect both nonsubmitted records and records that were later classified as ineligible for the study. Because of the sampling design, survivors were weighted more heavily than deaths and cases from non-outlier hospitals were weighted more heavily than those from low-mortality or high-mortality hospitals. These weights were used to estimate the statewide prevalence of various characteristics in the target population of AMI patients admitted to medium-to-high volume California hospitals. Special statistical software is available to estimate the confidence intervals surrounding weighted rates and proportions, but p values cannot be estimated from weighted multivariate analyses. Therefore, unweighted analyses were also performed.

Table 13.1: Number of hospitals selected for AMI validation study by sampling strata

| | Hospital Risk-Adjusted Mortality | | | |
|---|---|---|---|---|
| Hospital Volume | Better than expected | Neither better nor worse | Worse than expected | Total |
| Medium (50-117 pts) | | | | |
| Selected | 6 | 6 | 6 | 18 |
| Eligible | 7 | 119 | 10 | 136 |
| Large (118+ pts) | | | | |
| Selected | 4 | 4 | 4 | 12 |
| Eligible | 11 | 75 | 6 | 92 |
| Total | | | | |
| Selected | 10 | 10 | 10 | 30 |
| Eligible | 18 | 194 | 16 | 228 |

Table 13.2: Number of records received for AMI validation study by sampling strata

| | Hospital Risk-Adjusted Mortality | | | |
|---|---|---|---|---|
| Hospital Volume | Better than expected | Neither better nor worse | Worse than expected | Total |
| Medium | | | | |
| Received | 151 | 175 | 174 | 500 |
| Requested | 168 | 180 | 179 | 527 |
| % Received | 89.9 | 97.2 | 97.2 | 94.9 |
| Large | | | | |
| Received | 174 | 153 | 178 | 505 |
| Requested | 179 | 180 | 179 | 538 |
| % Received | 97.2 | 85.0 | 99.4 | 93.9 |
| Total | | | | |
| Received | 325 | 328 | 352 | 1005 |
| Requested | 347 | 360 | 358 | 1065 |
| % Received | 93.7 | 91.1 | 98.3 | 94.4 |