ANNALS OF SURGERY
Vol. 222, No. 4, 593–599

The Need for Accurate Risk-Adjusted Measures of Outcome in Surgery

Lessons Learned Through Coronary Artery Bypass

Bartley P. Griffith, M.D., Brack G. Hattler, M.D., Robert L. Hardesty, M.D., Robert L. Kormos, M.D., Si M. Pham, M.D., and Henry T. Bahnson, M.D.

From the Division of Cardiothoracic Surgery, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania

Objective
The authors review the Pennsylvania Health Care Cost Containment Council reports on coronary artery surgery and compare this reporting structure to others, including the Society for Thoracic Surgeons database, currently used by their own program. The authors review the growing likelihood of a need for outcome measures for all of the surgical subspecialties.

Summary and Background Data
Pressure from consumers and insurers will require surgical specialties to be graded by objective outcome measures. Practitioners must be prepared and become involved in the process.

Methods
The authors reviewed the data, which grades all of Pennsylvania's hospitals at which coronary artery bypass is performed. Apparently, the major risk factors commonly employed in most other risk adjustment schemes for cardiac surgery have been deleted, and the practitioners might be judged unfairly. The Pennsylvania system appears to be insurance driven to reward low-cost providers who operate on patients with the lowest risk.

Results
Review of data suggests that the Pennsylvania Health Care Cost Containment Council's annual publication, *A Consumer's Guide for Coronary Artery Bypass Surgery*, misrepresents fair risk adjustment in favor of lower-risk patients, thereby encouraging better score cards for those institutions with patients who are less ill. Data regarding charges for the procedure have not been risk adjusted or related to a regional economic index.

Conclusions
Surgeons must prepare to better understand relevant models that evaluate outcome. Cardiothoracic surgery is one of the first specialties to feel the pressures of mandated evaluations, and the lessons learned in Pennsylvania should be applicable to other states and their practitioners.

Surgeons who perform coronary artery bypass have been among the first physicians to feel the effects of financially and politically driven regulations to reign in costs. These regulations have shown inadequate regard for quality of care.[1] An overabundance of surgeons and hospitals, large volumes of patients and costs, and relatively uniform procedures have made use of coronary artery bypass an easy target. In 1987, the Health Care Financing Administration (HCFA) and the Veterans Affairs hospitals publicly disclosed survival rates after operation.[2,3] The process was deemed "scorecard surgery" by the press when such data were made available in New York and Pennsylvania.[4,5] Surgeons are accustomed to a review of outcomes and results of their efforts and are aware of many factors that might affect rates of survival. The system introduced in Pennsylvania was politically imposed without input from hospital administrators or cardiac surgeons. This system has not proven valuable as an instrument of quality improvement. Our purpose in this report was to comment on the program of Pennsylvania's Health Care Cost Containment Council and on our finding that it is better to evaluate quality through use of the Society of Thoracic Surgeons National Database for cardiac surgery. We hope that discussion of the attack currently aimed at cardiac surgery will stimulate preparedness of others.

THE HEALTH CARE COST CONTAINMENT COUNCIL OF PENNSYLVANIA

Experience with outcome measurement in Pennsylvania began in 1987, when labor unions and insurance companies encouraged the state legislature to create the Health Care Cost Containment Council with the mission of addressing costs and quality of health care. The 21-member council consists of 12 business and labor representatives (6 each), 1 physician, 1 hospital representative, and 1 health care consumer. The law mandated the collection of data to include information on the severity of illness and associated morbidity of all hospital admissions. The council contracted Mediqual Systems Inc. of Westborough, Massachusetts, for use of its MedisGroup Severity of Illness System (now known as Atlas Outcomes) to obtain information on mortality of patients and severity of illness. The council has published its estimate of hospital effectiveness based on 41 hospitals performing coronary artery bypass since 1990.[6] In 1992, faced with an appropriations deadline, the council rushed to print its first annual physician-specific report, entitled "1990 Coronary Artery Bypass Graft Surgery," (Fig. 1) whose purpose was to "provide data on the treatment, effectiveness, and average hospital charge for coronary artery bypass surgical cases performed in Pennsylvania in 1990 and was designed for health care purchasers and consumers to use as a guide to selecting a cardiac surgeon and a hospital."[7] Public interest was high. More than 14,000 requests for the document were made, 8,000 of which were from individuals. The mortality risk adjustment depended entirely on a patient's "admission severity group," as determined by the use of the MedisGroup program.[8] This program was designed to assess severity of illness for all admitted patients regardless of diagnosis. Its use as a predictor of death after coronary artery bypass had not been validated by experts or practitioners.[9,10] A score for admission severity is calculated from the hospital record by hospital personnel, who search for key clinical findings at specified times during the course of hospitalization. Missing data are recorded

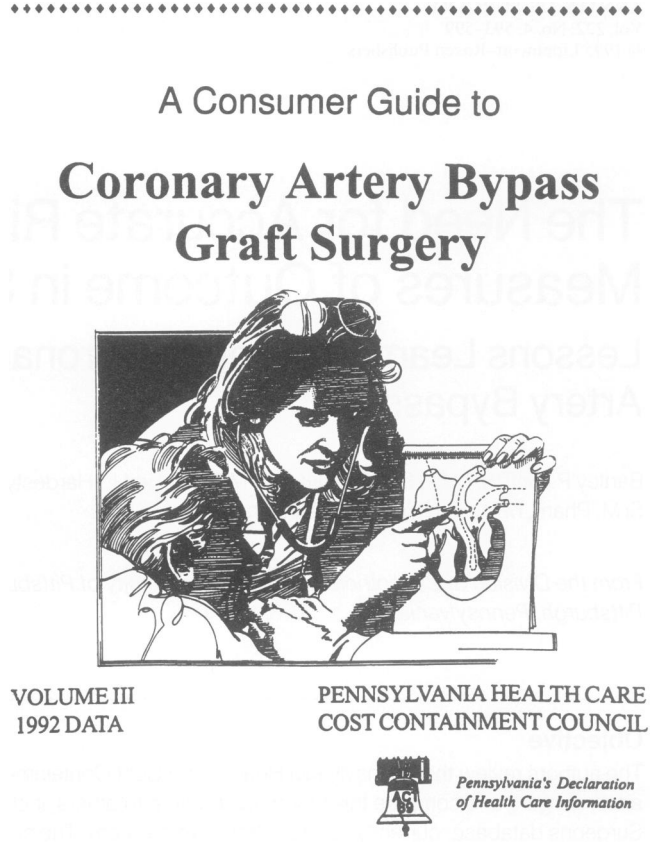

Figure 1. Front cover from a consumer-oriented Pennsylvania Health Care Cost Containment Council–sponsored annual publication.

Presented at the 115th Annual Meeting of the American Surgical Association, April 6–8, 1995, Chicago, Illinois.

Address reprint requests to Bartley P. Griffith, M.D., University of Pittsburgh Medical Center, Division of Cardiothoracic Surgery, C 700 Presbyterian University Hospital, 200 Lothrop Street, Pittsburgh, PA 15213-2582.

Accepted for publication April 10, 1995.

Table 1. 1992 TEST VARIABLES AND SIGNIFICANT RISK FACTORS

Test Variables
Acute MI
Admission severity group
Admit type
Age
Age squared
Angina
Cardiogenic shock
CHF
Diabetes
Dialysis
Gender
Hypertension
Previous CABG
PTCA
Renal failure
Transfer-in
Significant risk factors
Acute MI
Cardiogenic shock
Dialysis dependency
Previous CABG
Admission severity group
Age square
CHF
Female
Renal failure

MI = myocardial infarction; CHF = congestive heart failure; CABG = coronary artery bypass graft; PTCA = percutaneous transterminal coronary angioplasty.

as normal. The score is not revised after the specified time should a patient develop a problem, such as a myocardial infarction or shock, before the need for surgery. The insensitivity of the model to comorbid conditions was criticized, because it did not include ejection fractions, left main occlusive disease, emergently performed operations for failed percutaneous coronary arterial angioplasty, emergencies associated with acute cardiac catheterization, and the requirement for preoperative inotropic drugs or intra-aortic balloon pump.[11] The University of Pittsburgh Medical Center filed an official grievance with the council. We found that MedisGroup did not accurately assess high-risk patients who underwent coronary artery bypass, and we pushed for the addition of omitted but widely acknowledged comorbid conditions. Consequently, the council now includes a review of 15 additional potential factors, which are included with the MedisGroup admission severity score.[12] The 16 variables were tested in 1991 and 1992 by regression analysis for determination of their individual effects on outcome. Nine additional comorbid factors were added in 1991, and each of these, with the exception of diabetes, was also found to be significant in 1992 (Table 1). The statistical methods used are now available, and only the MedisGroup admission severity weight remains proprietary.[12]

Other complaints received by the council have included the following: (1) The use of death as the sole adverse outcome regardless of postoperative functional status, (2) the fact that the report was at least 2 years behind current practice when published, and (3) the finding that ICD-9 coding rules were not interpreted uniformly by abstractors. In 1992, average charges for coronary artery bypass in the Commonwealth of Pennsylvania ranged from $25,000 to $97,000.[13] Institutions such as our own with relatively high charges complained that the publication of a rank order of charges was a poor indication of the average payment received by hospitals and that charges were not related to the complexity of cases or regional economic indexes. Institutional charges were unfairly related to the number of patient deaths.

SOCIETY OF THORACIC SURGEONS NATIONAL CARDIAC SURGERY DATABASE

In 1991, our group opted to enroll all cardiac surgical patients in the Society of Thoracic Surgeons (STS) National Database.[14] We hoped that this would provide us with comparative data that had been validated. This data base had evolved from a proposal from the STS Standard and Ethics Committee in 1987 to develop a national data base for cardiac and thoracic surgery. Approximately 1400 surgeons from more than 700 hospitals have since entered more than 520,000 patients into this voluntary system. Twenty-four variables for each patient are analyzed with an algorithm for risk stratification that uses the Bayes theorem.[15-17] The large enrollment allows for an accurate statistical model and is predictive of probability of operative death. At our institution, the data base had previously been useful in quantification of risk to life and morbidity associated with the treatment of severely ill, high-risk patients who required coronary artery bypass.[18] Between 1991 and 1993, 1044 consecutive patients undergoing coronary artery bypass at the University of Pittsburgh Medical Center were entered into the STS data base. The predicted and observed mortalities were compared. The possible effects of comorbid conditions were evaluated (Table 2). The observed mortality rates correlated well with predicted rates ($p < 0.005$). Observed and expected risk correlated with the number of postoperative complications and length of hospital stay. Those patients at lowest risk for death averaged 0.29 complications, with a corresponding hospital stay of 8.9 days, whereas those with predicted and observed risk of death greater than 10% to 20% had 1.63 complications and required an average of 18.2 days of hospitalization.

Table 2. UPMC CABG OUTCOME BY STS RISK

STS expected mortality (%)	0–2.5	2.5–5	5–10	10–20
UPMC observed mortality (%)	2.3	4.9	11.1	16.9
No. of complications/patient	0.29	0.75	0.94	1.63
Hospital days	8.9	12.7	11.6	18.2
No. of CABG patients	607	225	117	59

CABG = coronary artery bypass graft; UPMC = University of Pittsburgh Medical Center; STS = Society of Thoracic Surgeons.

Table 4. UPMC CABG OUTCOME, 1044 CABG

	Age	LVEF	LVEDP	PAS	PAWP
Alive	64.4	50.3	13.1	32.4	12.5
Dead	67.2	42.1	17.2	40.6	16.9
p value	<0.10	<0.001	<0.025	<0.005	<0.025

UPMC = Univeristy of Pittsburgh Medical Center; CABG = coronary artery bypass graft; LVEF = left ventricular ejection fraction; LVEDP = left ventricular end diastolic pressure; PAS = pulmonary artery systolic; PAWP = pulmonary artery wedge pressure.

Univariate analysis of risk factors demonstrated female sex, prior operation, cardiogenic shock, and congestive heart failure to be strongly associated with risk of death (Table 3). Average age did not different between survivors and nonsurvivors. The ejection fraction was lower, and left ventricular end diastolic pressure and pulmonary artery systolic pressure were higher in nonsurvivors (Table 4).

DISCUSSION

Unadjusted statistics on hospital mortality rates from claims made to Medicare for cardiac surgery were first released to the public in 1988, when the *New York Times* filed a request under the Freedom of Information Act.[2] In 1991, *Newsweek* sued the New York State Department of Health to obtain the data base for information regarding not only coronary artery bypass, but also cardiac surgeons who performed the procedure.[19] The Supreme Court of New York agreed that the public was entitled to the information so as to make informed decisions about health care. Soon thereafter, Pennsylvania enacted a law mandating public disclosure of outcomes for hospital and surgeon. There was a flurry of public outcry from cardiac surgeons, who felt unfairly scrutinized, and from patients.[20] Fortunately, turmoil has settled and cardiac surgeons have begun to understand the need for and requirements of organized outcome measurements. A national symposium on the status of outcome measurement for cardiac surgery was held in 1994 by the Veterans Affairs Office of Quality Management.[21] There was strong support for programs of continuous quality improvement, and various successful models were discussed. One requirement agreed on was physician involvement.

Table 3. STS UNIVARIABLE ANALYSIS: SIGNIFICANT PREOPERATIVE RISK VARIABLES

	Mortality With Risk Present	RR	p Value
Female	25/273 (9.16)	2.4	<0.005
Hypertension	44/686 (6.41)	2.1	<0.05
Pulmonary hypertension	7/50 (14)	2.9	<0.025
Prior CABG	11/93 (11.8)	2.6	<0.01
Prior other cardiac	4/17 (23.5)	4.7	<0.005
Prior MI <21 days	23/248 (9.27)	3.2	<0.005
Cardiogenic shock	9/37 (24.3)	5.3	<0.005
ACE inhibitor	17/170 (10)	2.3	<0.005
IV nitrates	25/285 (8.77)	2.2	<0.005
Inotropes	6/33 (18.2)	3.1	<0.005
Diuretics	28/259 (10.81)	3.1	<0.005
NYHA class IV	29/328 (8.84)	2.4	<0.005
Nonelective	35/346 (10.12)	3.5	<0.005
IABP	16/88 (18.2)	4.5	<0.005
EF 30	9/71 (12.68)	2.7	<0.01
LVED p > 18	10/93 (10.75)	3.0	<0.025

STS = Society of Thoracic Surgeons; RR = relative risk; CABG = coronary artery bypass graft; MI = myocardial infarction; ACE = angiotensin converting enzyme; IV = intravenous; NYHA = New York Heart Association; IABP = intra-aortic balloon pump; EF = ejection fraction; LVED = left ventricular end diastolic pressure.
Values in parentheses are percentages.

The HCFA has stopped reporting hospital mortality rates, primarily because data were produced without an informed market for the information, but it has enthusiastically initiated a new program entitled the "Health Care Quality Improvement Program."[22] Still underfunded at approximately 0.1% of Medicare dollars, this program is meant to identify quality of care, not just outcome, through recruitment of a community of interested parties committed to the task. The HCFA views the previous reports of raw mortality rates as flawed attempts to "cull bad apples." Medicare intermediaries have been contracted to look for indicators of quality of care and to recruit a few hospitals to provide test sites where the indicators can be evaluated. The Cooperative Cardiovascular Project of the HCFA is an early quality-improvement initiative that studies patient care after acute myocardial infarction. The goal is to develop a group of indicators that demonstrate a consistent process of care. A similar study of coronary artery bypass is to follow. It

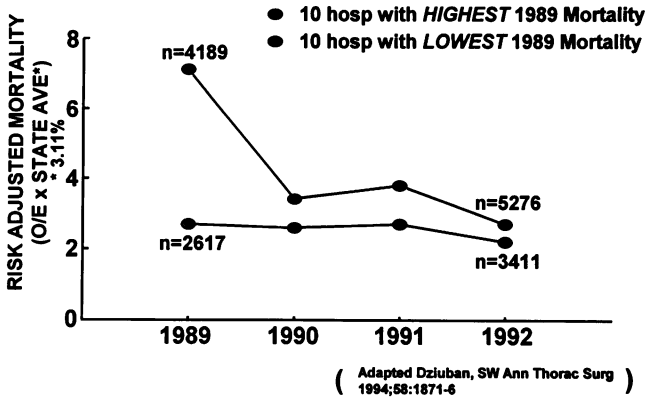

Figure 2. New York State Cardiac Surgery Report shows improvement in ten centers previously cited as low performers.

projects that by 1996, 40% of Medicare discharges will have complete sets of indicators of quality.[22]

The Department of Veterans Affairs has also moved primarily from oversight of cardiac centers to promotion of continuous quality improvement through the use of risk-adjusted outcomes.[23] In 1987, the Veterans Cardiac Surgery Consultants Committee, vested with oversight, requested the addition of an evaluation of patient risk to better judge quality of care. Earlier efforts, like those of the HCFA, had involved consideration of raw mortality rates only. This forced some surgeons to refuse to operate on high-risk patients to avoid censure. To help improve care, the committee's program now uses risk-adjusted mortality rates that are reviewed semiannually and discussed cooperatively with program directors and hospital administrators. Hospital administration is involved in morbidity and mortality evaluation. Occasionally, operative techniques and supervision of surgical residents are criticized along with structures of care, including such diffuse items as the availability of blood gas analysis, alarms for cardiopulmonary bypass pumps, and numbers and training of personnel in the intensive care unit. Consultants point to a significant reduction in observed *versus* expected mortality rate.

The New York State Cardiac Surgery Reporting System appears to be on the right track as well. The Department of Health and the Cardiac Advisory Committee, made up of expert physicians, have cooperated in the process since 1989. Significant overall improvement in the mortality rate from all providers has been demonstrated, from 2.7% to 2.19%.[24] Remarkably, the providers who performed poorly initially are now indistinguishable from the highest performers (Fig. 2). Although we cannot weigh the relative value of the state's reporting system on improvement, there is evidence that the personnel and procedural changes made by individual hospitals may have reduced the number of preventable deaths.[25] We must stress that elimination of "bad apples" has not been the only consequence of public disclosure in New York: risk-adjusted mortality rates have improved for the highest- and lowest-ranking providers, and evidence of case shifting has not been found.[24]

In our center, we have relied more on the STS data base than on the commonwealth's mandated program, based on MedisGroup, to stratify risks and to assess outcomes after coronary artery bypass. This is problematic in that outcomes in the public domain may not be those followed closely intramurally. Reassured by the STS data base, we have comfortably pursued our differences with the Health Care Cost Containment Council. As an academic medical center, we believe we have a responsibility to work with the council to improve and to understand its basis of risk assessment. Data released in 1993 and 1994 showed improvement, due in part to suggestions made by our group and by others statewide. It remains troublesome that the council lacks significant physician involvement, and clearly it is an instrument of business, labor, and insurance companies, which may not give priority to issues other than cost. We question whether the council's annual budget of more than $3 million is cost-efficient. Politically imposed systems do have the advantage of nonvoluntary participation and should, if properly composed, be useful. The New York State and Veterans Affairs hospital programs appear to be capable of improving quality through the engagement of experts who are committed to the process. We recognize that the STS data base is flawed because it is voluntary and risks gaming by inflation of risk-adjustment variables. Annual revision of the Bayes condition probability model, based on changing population and severity of illness, has been proven sound, but without internal audit the data will be considered suspect at some level.

We have asked members of the Society of Surgical Chairmen to respond to a simple questionnaire designed to determine the extent to which measures of outcome were recorded by surgical departments. Of 142 inquiries, 92 responses were received. Ninety-two percent (82/89) of cardiac surgical services were recording data on outcomes. Ten of 52 states mandated assessment of outcomes for cardiac surgery. Only Pennsylvania and New York publish annual physician-specific reports. Although it appeared that general surgical services were not diligent in recording outcomes, 56% (51/89) were involved to some extent in the process. The American College of Surgeons has sensed the need to respond to the challenge brought on us by managed care, which prioritizes cost and soft estimates (*i.e.*, not scientific) of patient satisfaction as primary measures of outcome. In response to a 1994 research committee panel, the American College of Surgeons has scheduled meeting of representatives from each of the specialty advisory councils for June 1995. The agenda is not fixed, but a major focus

will be on whether the American College of Surgeons should become a manager and central depository for outcome data (Jonnasson O, American College of Surgeons. Unpublished data, 1995). Although efforts to reduce costs are laudatory, those suggestions not associated with significant input by practitioners are hazardous. Perhaps one of the greatest challenges for surgeons in the next few years will be to present a validated plan for risk-adjusted outcomes that is sensitive to the concerns of patients and payors. As discussed in this presentation, models exist at the federal, state, regional, and specialty levels that have been reasonably successful in dealing with cardiac surgery. To underestimate the importance of the issue is tantamount to capitulation to shortsighted forces concerned with cost and not with dissemination of information required for continuous improvement of quality.

References

1. Relman AS. Assessment and accountability: The third revolution in medical care. N Engl J Med 1988; 319:1220–1222.
2. Medicare Hospital Mortality Information, 1986. Washington, DC: U.S. Department of Health and Human Services, 1987, Publication 01-002.
3. Takaro T, Ankeney JL, Lanning RC, Peduzzi PN. Quality control for cardiac surgery in the Veterans Administration. Ann Thorac Surg 1986; 42:37–44.
4. "Surgical Scorecards," American Broadcasting Company, June 4, 1992. Prime Time Transcript #248.
5. Topol ET, Califf RM. Scorecard cardiovascular medicine: Its impact and future directions. Ann Intern Med 1994; 120:65–70.
6. Hospital Effectiveness Report: Pennsylvania Health Care Cost Containment Council 1990, Publication HE-4-88.
7. Coronary Artery Bypass Graft Surgery Guideline for Consumers. HC4 1992, Harrisburg, PA.
8. Coronary Artery Bypass Graft Surgery: A Technical Report. The Pennsylvania Health Care Cost Containment Council, Harrisburg, PA 1992.
9. Jezzoni LI, Moskowitz MA. A clinical assessment of MedisGroups. JAMA 1968; 260:3159–3163.
10. Blumberg MS, Dias ED. Estimates of expected acute myocardial infarction mortality using MedisGroups admission severity groups. JAMA 1991; 265:2965–2970.
11. Report of the ad hoc committee on physician-specific mortality rates for cardiac surgery. Ann Thorac Surg 1993; 56:1200–1202.
12. Coronary Artery Bypass Graft Surgery Technical Report. Vol III. Pennsylvania Health Care Cost Containment Council, 1992, Harrisburg, PA.
13. Coronary Artery Bypass Graft Surgery, Guide for Consumer HC4, 1994, Harrisburg, PA.
14. The Society of Thoracic Surgeons National Cardiac Surgery Database Manual for Data Managers. Summit Medical Systems, August 1993, Minneapolis, MN.
15. Edwards FN, Clark RE, Schwartz M. Coronary artery bypass grafting: the Society of Thoracic Surgeons National Database Experience. Ann Thorac Surg 1994; 57:12–19
16. Edwards FH, Graeber GM. The theorem of Bayes as a clinical research tool. Surg Gynecol Obstet 1987; 165:127–129.
17. Edwards FH, Albus PA, Zajtcnuk R, et al. A quality assurance model of operative mortality in coronary artery surgery. Ann Thorac Surg 1989; 47:646–649.
18. Hattler BG, Madia C, Johnson C, et al. Risk stratification using the Society of Thoracic Surgeons program. Ann Thorac Surg 1994; 58: 1348–1352
19. Zinman D. Heart surgeons rated: State reveals patient-mortality records. *Newsweek* 1991, December 18, p 34.
20. Byer MJ. Saint hearts. *New York Times* 1992, March 21, Sect A: 23
21. Daley J, Grover FL, Hammermeister KE. Using outcomes data to improve clinical practice. Ann Thorac Surg 1994; 58:1795–1806.
22. Tencks SF. HCFA's health care quality improvement program and the cooperative cardiovascular project. Ann Thorac Surg 1994; 58: 1858–1862
23. Grover FL, Johnson RR, Shroyer AL, et al. The Veterans Affairs continuous improvement in cardiac surgery study. Ann Thorac Surg 1994; 58:1845–1851.
24. Hannan EL, Kumar D, Racz M, et al. New York State's cardiac surgery reporting system: four years later. Ann Thorac Surg 1994; 58:1852–1857.
25. Deiuban SW, McIlduff JB, Miller SJ, Dal Col RH. How a New York cardiac surgery program uses outcome data. Ann Thorac Surg 1994; 58:1871–1876.

Discussion

DR. ALDEN H. HARKEN (Denver, Colorado): I would like to compliment Dr. Griffith and his colleagues for a superb and uncomfortably important study which serves to emphasize the utility of large national databases and also to communicate the observations of Fred Grover, our chief of cardiothoracic in Colorado. Fred has been involved in the STS Database Committee since its inception under the leadership of Dick Clark and he has also been involved in the development of the VA National Cardiac Surgery Database with Dr. Karl Hammermeister.

The STS Database has now more than 500 centers and, as Dr. Griffith just pointed out, more than 500,000 cardiac surgical procedures. This voluntary database collects group and hospital, not surgeon-specific, data. Because of its huge numbers and large geographic distribution, we believe that it is representative of cardiac surgical practice throughout the nation.

As with all databases, it must be subjected to scrutiny with inter-rater reliability of data and screening reporting. The Pittsburgh group has done a nice job of validating the STS risk model by dividing the patients into six relative risk groups. The observed mortality of the Pittsburgh patients fits very nicely in the STS predicted mortality of each patient risk group.

The VA Cardiac Surgery Database is mandatory and now includes 55,000 cases collected since 1987. Each program director of the 43 VA hospitals performing cardiac surgery is furnished with a graph, like this one, of expected versus observed mortality ratios every 6 months. With a system that works, half the cardiac surgical units should be above this line (Observed: Expected = 1). And half the units should be below this line. But what the lay public does not understand is that there will be groups up here (O:E >1) with a very low raw mortality and groups down here (O:E < 1) with a very high raw mortality.

During this 7-year period, we have seen in the VA system a significant decrease, as Dr. Griffith alluded to, in the observed to expected mortality rates. The Northern New England