

Does reporting of coronary artery bypass grafting from administrative databases accurately reflect actual clinical outcomes?

Michael J. Mack, MD, Morley Herbert, PhD, Syma Prince, RN, Todd M. Dewey, MD, Mitchell J. Magee, MD, and James R. Edgerton, MD

Objectives: Quality assessment of coronary artery bypass grafting has traditionally been performed with data from clinical databases. Administrative databases that rely primarily on information collected for billing purposes increasingly have been used as tools for public reporting of outcomes quality. The correlation of administrative data with clinical data for clinical quality assessment has not been confirmed.

Methods: With data from a clinical database, we analyzed the outcomes of all patients who underwent coronary artery bypass grafting surgery in 1 hospital between 1999 and 2001. This information was collected before, during, and after the surgery and hospitalization by designated clinical individuals involved with the patient's care and then entered into an audited clinical database (The Society of Thoracic Surgeons National Cardiac Database). These data were then compared with administrative data collected on the same cohort of patients for the number of procedures performed and mortality rate as reported by the federal government (Medical Provider Analysis and Review), state government (Texas Health Care Information Council), hospital system (HCA, Inc, Casemix Database), and an internet Web site (healthgrades.com). Data were analyzed on the basis of the population reported, definitions used, risk assessment algorithms, and case volumes.

Results: By using the audited The Society of Thoracic Surgeons database as the standard and aggregating the reporting of case volumes by the inclusion criteria of various sources of administrative data, we found variances in the reported procedure volumes and mortality. Case volumes were overreported by as much as 21% in all patients and underreported by up to 16% or more in Medicare patients. Mortality in administrative data exceeded that reported in clinical data by 21%. Reasons for variances included time period reported (calendar vs fiscal year), population reported (all patients, Medicare patients, Medicare patients aged ≥ 65 years), date used for the patient record captured (date of surgery, discharge), and the definition of mortality. Different proprietary risk-adjusting algorithms used magnified variances with risk-adjusted mortality exceeding the Society of Thoracic Surgeons data by as much as 61%.

Conclusions: Substantial variability of reported outcomes is seen in administrative data sets compared with an audited clinical database in the end points of the number of procedures performed and mortality. This variability makes it challenging for the nonclinician unfamiliar with outcomes analysis to make an informed decision.

Reporting the outcomes of surgical procedures, especially coronary artery bypass grafting (CABG), is being increasingly used as a measurement of quality.¹⁻⁴ The early original cardiac surgery databases were clinical in nature and constructed to assess surgical outcomes to perform continuous quality improvement. The Northern New England Cardiovascular Disease Study Group and

From the Cardiopulmonary Research Science and Technology Institute, Dallas, Tex.

Read at the Eighty-fourth Annual Meeting of The American Association for Thoracic Surgery, Toronto, Ontario, Canada, April 25-28, 2004.

Received for publication April 23, 2004; revisions received Oct 21, 2004; accepted for publication Oct 27, 2004.

Address for reprints: Michael J. Mack, MD, 7777 Forest Lane, Suite A323, Dallas, TX 75230 (E-mail: smullins@csant.com).

J Thorac Cardiovasc Surg 2005;129:1309-17
0022-5223/\$30.00

Copyright © 2005 by The American Association for Thoracic Surgery

doi:10.1016/j.jtcvs.2004.10.036

The Society of Thoracic Surgeons (STS) established voluntary clinical databases in the 1980s for these purposes.^{5,6} Beginning in 1986, the then Health Care Financing Administration (HCFA) began the release of unadjusted outcomes data derived from an administrative database used for payment to hospitals for billed charges. Concerns about data quality, adequacy of methods, issues of selection bias and patient population, and inadequate risk adjustment were some of the reasons why HCFA abandoned its decision to publicly release hospital mortality statistics after 1993.⁷ Although published outcomes of these data ceased, administrative data have remained in the public domain and readily available through the successor to HCFA, The Centers for Medicare and Medicaid Services (CMS), as Medical Provider Analysis and Review (MedPAR) data. These data have been repackaged by different sources of quality ratings and released to the public.

Numerous state agencies, including Texas,⁸ and consumer-oriented Web sites such as HealthScope.org⁹ and healthgrades.com¹⁰ also report clinical quality measures using administrative data outcomes on CABG. Lay periodicals such as *US News and World Report*, a widely marketed purveyor of quality, rank hospitals based in part on outcomes data.¹¹ A coalition of private and public purchasers of health insurance, the LeapFrog Group uses outcomes data as a measure of hospital safety.¹² The common denominator for the data for all of these "report cards" is the information extracted from administrative databases collected for hospital billing purposes. The correlation of these data with that gathered from specifically designed clinical sources has not been clearly ascertained.¹³

The STS National Cardiac Database (STS-NCD) has grown to include outcomes collected on more than 2 million patients from 60% of all cardiac surgery programs in the United States.¹⁴ It is a voluntary registry but is continuously audited on many levels for completeness and accuracy and is generally accepted in the specialty field of cardiac surgery as the benchmark for clinical outcomes analysis.¹⁵ We recently audited specific outcome data and correlated them with the source data, the patient clinical hospital record, from our own hospital entered into this national clinical database and found this registry to be highly accurate for the reporting of major end points.¹⁶ The report of operative mortality was accurate with complete correlation with source data, and there was an error rate of less than 1% for most major complications. Over the years, we have repeatedly noticed discrepancies between our own outcomes as reported by the clinical STS-NCD and those reported on our program using information extracted from administrative databases. To determine the magnitude and reasons for the variances between the various reports, we compared and analyzed the number of procedures performed and operative

mortality of the same cohort of patients as reported by the STS-NCD and various other reports, all of which used administrative billing data as their source.

Methods

The records of all isolated CABG procedures performed in 1 institution, Medical City Hospital, Dallas, Texas, by 5 surgeons were collected from the STS-NCD clinical database. Patient records from October 1, 1998, to December 31, 2001, were extracted to allow us to define cohorts based on either the calendar or the federal fiscal year as appropriate. We then collected the information regarding CABG procedures from the same time period for the same institution from 4 sources of information using administrative databases and compared them for the specific defined end points of case volume and mortality over the same 3-year period. Specific information regarding data from each source and method of query is as follows.

STS Clinical Database

The Cardiopulmonary Research Science and Technology Institute (CRSTI), a not-for-profit research center, collects clinical data from approximately 2800 cases per year from 22 surgeons in 18 centers and enters the information into a customized STS-approved database (Armus, Burlingame, Calif). Data are harvested to the STS twice annually and used for research and quality-control projects. Our data set is subjected to regular audits of 10% of the forms submitted and has been shown in a separate study to be accurate for reporting of mortality and case volume.¹⁶

For this study, data were exported from the Armus STS database into a spreadsheet format that was analyzed with SAS (version 9.0, SAS Institute, Cary, NC). Data for the 1 study hospital was selected and included only those patients classified as isolated CABG (ie, no concomitant procedures such as valves, carotid endarterectomy, and so forth). We correlated our entries into the STS-NCD database to our surgeons' clinical billing records to ensure completeness. By using the predicted risk of mortality and mortality at discharge and/or 30 days after surgery (the STS-NCD definition of operative mortality), we were able to calculate operative mortality and risk-adjusted operative mortality. Risk adjustment was carried out using the STS national mortality average for CABG or the national mortality average for patients aged more than 65 years as appropriate. Because the values change annually, we weighted the averages to reflect the patient volumes for each year. By using the standard STS operative mortality field (YES, if patient dies during that admission or within 30 days of surgery), we were able to compare mortality similar to the discharge plus 30 days reported by some entities. Because the average length of stay in our data set ranges from 7.3 to 8.6 days, the comparison is not exact because our operative mortality will reflect a slightly shorter period. For analysis of Medicare patients we queried the database for patients whom Medicare was either the primary or secondary payor (Medicare HMO) regardless of age.

Medicare Provider Analysis and Review Data

Effective October 1, 1983, Medicare implemented a prospective payment system for reimbursing in-patient hospital costs. The genesis of a MedPAR file begins with a medical records coder who interprets the physician documentation from the medical record

TABLE 1. Definitions used by reporting sources on patients who underwent coronary artery bypass grafting at Medical City Dallas Hospital 1999 to 2001

Source	Time patient record captured	Patient exclusions	Risk adjusted	Reporting period (y)	Definition of mortality
STS-NCD (CRSTI)	Date of surgery	None	Yes (STS)	Calendar	In-hospital and 30 d postsurgery
HCA Casemix	Date of discharge	None	No	Calendar	In-hospital
MedPAR	Date of discharge	Non-Medicare Medicare HMO	No	Fiscal	In-hospital
THCIC	Date of discharge	None	Yes (APR-DRG)	Calendar	In-hospital
Heathgrades.com	Date of discharge	Non-Medicare Medicare HMO Medicare < 65 y	Yes (proprietary)	Fiscal	In-hospital, discharge date + 1, 3, and 6 mo

CRSTI, Cardiopulmonary Research Science and Technology Institute; STS-NCD, Society of Thoracic Surgeons National Cardiac Database; MedPAR, Medical Provider Analysis and Review; THCIC, Texas Health Care Information Council; HMO, health maintenance organization; APR-DRG, All Patient Refined Diagnosis Related Group.

and adds International Classification of Disease, 9th Revision, Clinical Modification (ICD-9-CM) codes to this extracted information. There are 9 diagnostic ICD-9-CM codes (1 principal and 8 secondary) and 6 procedure codes (1 principal and 5 secondary) that can be assigned to each patient file. The hospital generates a UB-92 billing form, which is submitted to Medicare for reimbursement. CMS then aggregates all facility bills and generates an annual MedPAR file. This file is then analyzed and made available to the public domain. Each admission file is assigned to a specific Diagnostic Related Group (DRG) on the basis of diagnosis and the procedures performed. Patient files are then assigned to 1 of 511 DRGs based on this coding. The ICD-9-CM procedure codes 36.10 to 36.16 or 36.19 are used for the assignment of the DRG for CABG surgery. Most CABG procedures are included in DRG 106, 107, and 109. MedPAR data include all patients who are covered by Medicare regardless of whether they are aged less than or more than 65 years. The reporting base for MedPAR data is based on an October 1 to September 30 year. For this study, we queried MedPAR data for all isolated CABG procedures from 1999 to 2001 by ICD-9-CM procedure codes 36.10 to 36.16 or 36.19 in our hospital, excluding patients who also underwent valve surgery, transmyocardial revascularization, and/or ventricular assist device procedures. For comparison purposes with the other data sources, we queried the data for both calendar and fiscal years for all Medicare patients and for Medicare patients aged 65 years or more.

HCA Casemix

HCA, Inc, is a national hospital system that includes our institution, Medical City Hospital. The hospital completes the UB-92 universal billing form from the patient medical record, and these data are used to generate the HCA Casemix Database for the entire hospital system. This information is filed with CMS for Medicare reimbursement and with other payors for billing purposes. The HCA Casemix Database contains 15 diagnostic codes (1 principal and 14 secondary) and 10 procedure codes (1 principal and 9 secondary) compared with MedPAR files that contain 9 diagnostic and 6 procedure codes. We also queried this database by ICD-9 codes 36.10 to 36.16 or 36.19 excluding patients who also underwent transmyocardial revascularization, valve surgery, and/or ventricular assist device procedures for both calendar and fiscal years and by payor status and age for comparison purposes.

Texas Health Care Information Council

The Texas Health Care Information Council (THCIC) is a state agency created in 1995 that gathers information from hospitals and health maintenance organizations and publishes reports to allow “health consumers to compare and choose their hospitals.” Data are collected from hospital UB-92 billing data that include along with diagnosis and procedure codes information about the patient’s age, gender, accompanying medical conditions, and discharge status. The data collected are based on a calendar year and reflect

TABLE 2. All patients who underwent coronary artery bypass grafting in the calendar year January 1, 1999, to December 31, 2001

Source of data	Risk-adjusting algorithm	Reported volume	In-hospital mortality rate	Predicted mortality	Risk-adjusted rate	
CRSTI	STS database	STS	1121	2.8	3.6	2.3
HCA Casemix	Hospital billing	—	1354	3.3	—	—
THCIC	Administration	3M	1353	3.4	4.0	3.7

CRSTI, Cardiopulmonary Research Science and Technology Institute; STS, Society of Thoracic Surgeons; THCIC, Texas Health Care Information Council.

TABLE 3. All Medicare patients who underwent coronary artery bypass grafting in the fiscal year October 1, 1998, to September 1, 2001

	Source of data	Risk-adjusting algorithm	Reported volume	In-hospital mortality rate	Predicted mortality	Risk-adjusted rate
CRSTI	STS database	STS	505	4.2	5.4	3.1
MedPAR	Medicare	None	423	4.7	N/A	N/A

CRSTI, Cardiopulmonary Research Science and Technology Institute; STS, Society of Thoracic Surgeons; MedPAR, Medical Provider Analysis and Review; N/A, not available.

only patients discharged from a hospital during the calendar year of the data. Mortality is defined by this agency as only deaths during the hospital admission. THIC then uses the 3M Health Information Systems All Patient Refined Diagnosis Related Group (3M APR-DRG) risk algorithm, which provides a risk adjustment based on age, sex, and diagnostic codes reported. Case volumes and risk-adjusted mortality are presented as bar graphs. The risk-adjusted values are given along with the 95% confidence limits and the reference value for all hospitals in the state. The raw information is available, both in summary and in detailed spreadsheet format, on the THIC Web site for our hospital for the period from 1999 to 2001.

Healthgrades.com

Healthgrades.com provides outcomes analysis-based MedPAR data but excludes all Medicare patients who are less than 65 years of age or transferred to another acute care facility. Mortality data are obtained from the hospital discharge data and combined with an analysis of the National Death Index to report mortality as discharge, discharge +1 month, and discharge +6 months. Data are published on the basis of a fiscal year, October 1 to September 30, reporting period. Therefore, 1999 to 2001 data were actually from October 1, 1998, to September 30, 2001, as it is with all reporting that uses MedPAR source data.

A proprietary-based risk algorithm is used by healthgrades.com identified from the ICD-9 codes on Medicare billing. This risk adjustment is compiled from the 9 diagnostic fields in the MedPAR file with proprietary weighting. For this study we queried the healthgrades.com Web site for CABG procedure volumes and outcomes in our hospital from 1999 to 2001. Specific aspects of the reporting and risk adjustment were discussed directly with management of the organization.

A summary of definitions used by the various sources is listed in Table 1.

Results

The initial data obtained from the STS-NCD and 4 administrative data sources were aggregated into those that included all patients, patients reported as Medicare beneficiaries, and Medicare patients aged 65 years or older (Tables 2, 3, and 4). In regard to all patients who underwent CABG in the calendar years 1999 to 2001, the administrative databases report volumes that are 20.8% higher. Patients reported as Medicare beneficiaries were underreported by 16.2%, a figure that may be even higher because the CRSTI-STS database does not record cases as Medicare when the primary payor is a non-Medicare entity. Because some sources only use Medicare data for patients who are 65 years of age or older, we also examined this subset. Again, underreporting of 17% was seen, which is likely a low estimate for the same reasons as stated. Some additional reasons for the observed variances in Medicare patient reporting are listed in Table 5. Differing definitions of event captured (date of surgery or discharge) and the population on whom the data was reported (all patients, Medicare patients, or a subset of Medicare patients) accounted for the majority of procedural volume differences.

Similar variances were observed in mortality reporting with administrative databases reporting observed mortality in all patients to be as much as 21.4% higher (3.4% vs 2.8%) and risk-adjusted mortality as much as 61% higher (3.7% vs 2.3%). Mortality rates in Medicare patients were 11.9% higher (4.7% vs 4.2%) when reported by administrative data.

Discussion

Administrative databases provide an accessible source of information for the evaluation and reporting of health care

TABLE 4. Medicare patients aged 65 years or more who underwent coronary artery bypass grafting in the fiscal year October 1, 1998, to September 31, 2001

	Source of data	Risk-adjusting algorithm	Reported volume	In-hospital mortality rate	Predicted mortality	Risk-adjusted rate
CRSTI	STS Database	STS	458	4.4	5.4	3.3
MedPAR	Medicare	None	389	4.6	N/A	N/A
Healthgrades.com	MedPAR	Proprietary	377	4.5	4.0	N/A

CRSTI, Cardiopulmonary Research Science and Technology Institute; STS, Society of Thoracic Surgeons; MedPAR, Medical Provider Analysis and Review; N/A, not available.

quality. However, the correlation of these data sources for clinical information compared with databases compiled specifically for clinical purposes is not clear.¹⁷ We performed this comparative analysis because of our observed variances in operative mortality in outcomes reporting of our own patients from administrative databases compared with our own clinical STS-NCD.

The STS-NCD is the largest clinical database reporting quality outcomes in a single specialty with clinical information on 2,164,079 cardiac operations performed in 638 hospitals, representing approximately 70% of all cardiac surgery performed in the United States. Data quality improvement efforts and quality assurance mechanisms internally, regionally, and nationally have maximized the completeness and accuracy of this clinical database.¹⁸ We have used this database to collect and analyze clinical outcomes from our own cardiac procedures since 1986. The data are collected by clinical personnel at the time of the procedure and are monitored with internal quality assurance by dedicated personnel with clinical backgrounds in a single specialty with our own customized version of the STS-NCD by a not-for-profit research organization. Our database now tracks outcomes on more than 40,000 cardiac operations. The accuracy of our own STS-NCD data has also been verified compared with the clinical record.¹⁶ We used this data source as the benchmark against which we compared multiple reports on the same group of patients by administrative sources. Although one can debate the definitions used by the STS database, this is the clinical data set most commonly used nationally and the only one available for comparison. It is not our intent to laud it as the gold standard but to use it as a widely accepted and locally available audited clinical data set against which we could benchmark the various administrative sources. Although we had previously validated the accuracy of our own STS-NCD data set by audit compared with the clinical record, for this study we further confirmed the reported volume by comparison with our own practice billing information for date and type of procedure performed. Because of this, we confirmed, for example, that the reported procedure volume by the STS database in Table 3 (n = 505) is an accurate value. The discrepancies between this value and the MedPAR reported volume of 428 procedures were attributable to the factors listed in Table 5. We therefore concluded that the STS-NCD clinical data set is the most accurate for answering the simple question “How many isolated CABG procedures were performed in this hospital in a particular time period?”

Early studies of the accuracy of discharge data found that the coding of nonclinical data such as age, gender, and dates of admission and discharge was highly accurate, but that diagnosis and procedures were less reliably coded.¹⁹ The common source of all administrative data is the UB-92 claim form used for billing purposes. This claim form is

TABLE 5. Reasons for variations in MedPAR population data

- Medicare not primary payor, therefore not included (9%)
 - Medicare HMO member
 - Working > 65 y
 - Covered by working spouse
- Medicare is primary payor, but patient is aged < 65 y (dialysis)
- Transferred to another acute care facility

MedPAR, Medical Provider Analysis and Review; *HMO*, health maintenance organization.

generated using data gathered from documentation in the patient’s medical record by administrative coding personnel. The chart information is then extracted, and appropriate diagnostic and procedure codes called ICD-9-CM codes are selected and entered. There are 9 diagnostic code fields (1 primary and 8 secondary) and 6 procedural code fields (1 primary and 5 secondary). Billing software systems then trigger the order of coding leading to the selection of a DRG for the record. The HCA Casemix data and the THIC information are compiled from extracting this information.

Although private payors generally ignore the DRG code, Medicare uses it for billing purposes. The software that generates the appropriate DRG code takes the ICD-9-CM diagnosis and procedure codes, and prioritizes. There are 511 DRG codes, of which CABG is most commonly coded as DRGs 106, 107, and 109.

There are a host of reasons to explain the variances we observed between our outcomes as reported by our clinical database and those reported by various administrative sources. Reasons for variances in procedure volume and operative mortality reported by sources using outcomes from administrative databases include the use of different “captured events,” discharge date versus surgery date, population reported (all patients, Medicare patients, or Medicare patients aged 65 years or more), different reporting time periods (calendar vs fiscal year), which DRG codes were used to capture the study population, different definitions of operative death, and whether the reporting was risk adjusted and the adequacy of the risk-adjusted algorithm.²⁰⁻²³

Discrepancies are attributable to the designation of the population about whom outcomes are being reported. THIC and HCA Casemix report data on *all* patients undergoing CABG, whereas MedPAR and healthgrades.com report outcomes only on *Medicare* beneficiaries. Medicare patients comprised only 45% (505/1121) of the patients undergoing CABG in our analysis; thus, MedPAR and healthgrades.com are reporting outcomes on less than half of the patients actually undergoing CABG. This proportion changes according to local demographics but ranges from 40% to 60% of patients undergoing CABG in most

programs. This difference in population being reported on is also the reason for the largest discrepancy in operative mortality. The observed mortality in our Medicare population was 1.5 times higher (4.2% vs 2.8%) than in our total patient cohort in our clinical database.

Other reasons for reporting variances included whether the data was reported on a calendar year or fiscal year basis and the "event" being captured. The Federal Fiscal Year (October 1 to September 31) is used for reporting by MedPAR and usually all sources that rely on MedPAR data. The captured event in MedPAR is actually the date of discharge, not the date of surgery, and reports on discharges occurring between October 1 and September 30 of any year. Therefore, 1999 to 2001 data from this source include patients who were discharged between October 1, 1998, and September 30, 2001, and therefore underwent surgery before these dates.

The STS-NCD captures data based on the date the CABG procedure was performed and reports the data on a calendar year basis, that is, all patients actually undergoing a CABG in the year queried are captured. Therefore, the clinical database registered patients actually undergoing surgery between January 1, 1999, and December 31, 2001. Some administrative databases report the patient encounters as the date of discharge or the date the billing form is generated. This accounts for some of the variances between the different databases in reporting the number of patients undergoing surgery.

There are also significant differences in the definition of a Medicare beneficiary. Patients who are members of a Medicare HMO and those patients who are aged more than 65 years and working or are covered by a spouse's insurance are not reported as "Medicare" patients. Approximately 9% of patients captured as "Medicare" by our clinical database were actually in 1 of these categories and therefore are not included in MedPAR reporting. In addition, some sources, such as healthgrades.com, reported on only Medicare patients aged more than 65 years. Therefore, this would exclude a number of patients presented as high risk for CABG, including all patients with end-stage renal disease, accounting for both variances in procedure volume and operative mortality.

Another important reason for variance includes the method of query of the study population. Whereas the STS-NCD reliably reports all patients undergoing surgery during a time period, administrative databases may fail to report patients undergoing surgery according to how the database is queried. If only DRGs 106, 107, and 109 are used, a significant number of patients who underwent CABG may not be captured. We queried the various databases by ICD-9-CM codes 36.10 through 36.16 plus 36.19 to capture all patients undergoing CABG. However, if the method of query is by DRG assignment, then patients un-

dergoing CABG may not be captured because of assignment to a different DRG (because of additional procedures performed) by software analysis systems. A patient who underwent a tracheostomy is one such example (DRG 483). In our institution we also discovered CABG coded into DRGs 148 (bowel surgery), 149 (bowel surgery), 357 (uterine surgery), 415 (infectious surgery), and 468 (extreme or procedure unrelated to principal diagnosis), as well as DRG 483.

The definition of mortality varies between the different data information sources. The definition in the STS clinical database is any death occurring in-hospital or within 30 days of surgery date if discharged from the hospital. MedPAR data include only deaths occurring in the hospital. healthgrades.com data include in-hospital deaths and those occurring at 1, 3, and 6 months after discharge. In this data set it is actually possible to have mortality counted twice. If a patient was discharged from 1 hospital and admitted to another hospital and subsequently died, the death would be counted for both hospitals.

Similar variances exist from the use of different risk-adjustment algorithms. The STS risk algorithm includes 30 variables known to impact operative mortality. Although the factors included in this risk algorithm are known, the "weighting" of these factors is proprietary. Administrative risk-adjustment algorithms are subject to the limitations of administrative data, including insufficient clinical information, the use of administrative diagnoses and procedures, and possible confusion regarding the difference between complications and pre-existing conditions.

Inconsistencies in coding and limited numbers of variables coded are also problematic and explain some of the differences in risk-adjusted mortality. Risk models based on administrative data can lead to substantial misclassification of hospitals compared with models based on higher quality data.⁷ The risk algorithms used for risk adjustment by 3M APR-DRG and healthgrades.com are both proprietary and therefore cannot be assessed. Neither risk algorithm has been validated in a published peer review publication. Yet these risk-adjustment formulas relying on the limitations of administrative data are used to discriminate between factors in different hospitals on the assessment of quality.

Although risk adjustment is a desirable tool to provide a level playing field for outcomes analyses, the limited number of diagnostic and procedure codes allowed (9 and 6, respectively) for UB-92 form-generated data, and the reliance on nonclinical personnel to code and prioritize these factors, make the accuracy of this risk adjustment questionable. The STS-NCD risk algorithm uses 30 preoperative risk variables known to be directly associated with clinical outcomes as the method of risk adjustment. The proprietary nature of all risk algorithms (STS, APR-DRG, healthgrades.com)

com), however, does not allow for comparison of the “weighting” of different factors into the risk equation.

Conclusion

Whereas variances in outcomes reporting of procedural volume and operative mortality are explained by differences in factors including definitions of reporting period and mortality, specific population studied, method of query, and whether risk adjusting of mortality is used, this was not clear to us as clinicians before this analysis. Although this may be readily apparent to statisticians and individuals conversant in outcomes analysis, it is beyond the capacity of most clinicians and therefore the lay public to understand the subtleties of query and analysis that cause the variances. One could argue that it does not matter as long as the same “measuring stick” is used to assess all programs. However, significant local differences exist in the percentage of Medicare patients undergoing CABG and the inclusion or exclusion of many patients because of coding differences or analysis criteria.

A significant shortcoming of all administrative quality assessment sources is their reliance on administrative data, which are by necessity collected by individuals not actively involved in clinical care and fit into a coding system with inherent limitations. Inconsistencies in coding of diagnoses and procedures may lead to lower risk-adjusted mortality or higher risk-adjusted mortality, which can affect a hospital's rating.²³ Administrative data are more accessible and less expensive to collect and analyze than detailed clinical data and as a result are commonly used in public reporting. Administrative data may vary from clinical data in areas such as the quality of the data collected, methods used to analyze the data, and uncertainty as to the risk-adjustment models used. These issues all led the HCFA to stop releasing hospital mortality statistics in 1993. Subsequent analysis of this same administrative data by other groups does not solve the inherent issues of administrative data. A health care consumer, no matter how astute, cannot be expected to have the degree of sophistication necessary to properly analyze and assess public information, which is all based on administrative and not clinical databases, and appreciate the variances compared with outcomes reported by clinical databases.

We are indebted to April Simon, RN, for data and her patient explanation of the intricacies of administrative databases to clinicians and to Sherry Hill for manuscript preparation.

References

- Hannan EL, Racz MJ, Jollis JG, Peterson ED. Using Medicare claims data to access provider quality for CABG surgery: does it work well enough? *Health Serv Res.* 1997;31:659-78.
- Landon B, Iezzoni LI, Ash AS, et al. Judging hospitals by severity-adjusted mortality rates: the case of CABG surgery. *Inquiry.* 1996;33:155-66.
- Griffith BP, Hattler BG, Hardesty RL, Kormos RL, Pham SM, Bahnson HT. The need for accurate risk-adjusted measures of outcome in surgery. lessons learned through coronary artery bypass. *Ann Surg.* 1995;222:593-8.
- 3M Health Information Systems. Available at: www.3Mhis.com. Accessed: April 5, 2004.
- O'Connor GT, Plume SK, Olmstead EM, et al. A regional prospective study of in-hospital mortality associated with coronary artery bypass grafting. The Northern New England Cardiovascular Disease Study Group. *JAMA.* 1991;266:803-9.
- Clark RE. The STS National Database: alive, well, and growing. *Ann Thorac Surg.* 1991;52:5.
- Krumholz HM, Rathore SS, Chen J, Wang Y, Radford MJ. Evaluation of a consumer-oriented internet health care report card. *JAMA.* 2002; 287:1277-87.
- Texas Health Care Information Council. Available at: www.thcic.com. Accessed: April 5, 2004.
- HealthScope. Available at: www.healthscope.org. Accessed: April 5, 2004.
- Healthgrades.com: The Healthcare Quality Experts. Available at: www.healthgrades.com. Accessed: April 5, 2004.
- America's Best Hospitals: 2001 Hospital Guide. U.S. News and World Report; 2001. Available at: www.htm.usnews.com/usnews/nycu/health/hosptl/tophosp.htm. Accessed: April 6, 2004.
- A coalition of private and public purchasers of health insurance. Leapfrog Group. Available at: www.leapfroggroup.org. Accessed: April 5, 2004.
- Hannan EL, Kilburn H Jr, Lindsey ML, Lewis R. Clinical versus administrative databases for CABG surgery: does it matter? *Med Care.* 1992;30:892-907.
- Ferguson TB Jr, Dziuban SW, Edwards FH, et al. The STS National Database: current changes and challenges for the new millennium. *Ann Thorac Surg.* 2000;69:680-91.
- Shahian DM, Blackstone EH, Edwards FH, et al. Cardiac surgery risk models. a position paper from The Society of Thoracic Surgeons Workforce on Evidence-Based Surgery. *Ann Thorac Surg.* In press.
- Herbert MA, Prince SL, Williams JL, Magee MJ, Mack MJ. Are unaudited records from an outcomes registry database accurate? *Ann Thorac Surg.* 2004;77:1960-5.
- Shahian DM, Normand SL, Torchiana DF, et al. Cardiac surgery report cards: comprehensive review and statistical critique. *Ann Thorac Surg.* 2001;72:2155-68.
- Welke KF, Ferguson TB Jr, Coombs LP, et al. Validity of the Society of Thoracic Surgeons National Adult Cardiac Database. *Ann Thorac Surg.* 2004;74:1137-9.
- Fisher ES, Whaley FS, Krushat WM, et al. The accuracy of medicare's hospital claims data: progress has been made, but problems remain. *Am J Public Health.* 1992;82:243-8.
- Finlayson EV, Birkmeyer JD, Stukel TA, Siewers AE, Lucas FL, Wennberg DE. Adjusting surgical mortality rates for patient comorbidities: more harm than good? *Surgery.* 2002;132:787-94.
- Izezzoni LI. The risks of risk adjustment. *JAMA.* 1997;278:1600-7.
- Shroyer AL, Coombs LP, Peterson ED, et al. The Society of Thoracic Surgeons: 30-day operative mortality and morbidity risk models. *Ann Thorac Surg.* 2003;75:1856-65.
- Krumholz HM Chen J, Wang Y, et al. Comparing AMI mortality among hospitals in patients 65 years of age and older: evaluating methods of risk adjustment. *Circulation.* 1999;99:2986-92.

Discussion

Dr T. Bruce Ferguson, Jr (New Orleans, La). I thank Dr Mack and his colleagues for this important contribution. As CMS and all major insurance stakeholders embark on pay-for-performance programs targeted at the hospital and the individual provider level, the mechanism used to evaluate both hospital and providers is critical. Dr Mack, I have 3 questions for you and then a comment.

One of the most disturbing facts in this analysis is that the HCA Casemix and the THCIC data overestimated the actual number of bypass procedures performed, whereas the MedPAR and Health-

grades underreported candidate patients by 65 and 77 cases, respectively, in the fiscal year analysis.

Could you identify whether this was because of the inability to select out isolated coronary bypass procedures from concomitant procedures in the administrative data sets, whether it was, as you pointed out, just assignment to incorrect DRGs, or even additional factors that we as yet do not understand?

Dr Mack. There are a number of reasons for that. First of all, all administrative data come back to the common point that you are depending on a nonclinician to extract data from the medical records. It all depends how that data are extracted, whether actually concomitant procedures are included or not, and then how the procedure is coded and what code order it goes in. For instance, if there are only 6 procedure codes for CABG, if the coder happens to put CABG as the seventh code on the sheet, it never gets captured by the data. So it is very dependent on the coder.

Second, the triggering software for DRG payment automatically aggregates patients to the highest paying DRG, and in many instances it is DRG 483, which is with tracheostomy. Therefore, some of the highest risk patients actually never get captured.

Third, it all depends on how you query the data. Do you query it by ICD-9 codes, by all the ICD-9 codes that CABG can be caught under, or by DRG, and which DRG do you use?

Dr Ferguson. Second, the STS analysis published in 2004 by Carl Welke documented that there was no discernible underreporting of cases compared with Medicare data on a national level in the STS database. Your data document that, at least in this analysis, these case volume issues extend to state-mandated and hospital system data as well.

Could you speculate on how these data could be used to dissuade payor and public stakeholders from seizing on volume as a quality metric for programs?

Dr Mack. I think that you are clearly correct that the volume metric and the number of procedures that you perform are being used as a surrogate for quality surgery. There are a couple of problems with that. One is that there are many excellent surgeons today who perform isolated CABG in only half of their cases. So they do not meet volume quotas.

Second, it all depends on which data source is used. By using the administrative data, we find that there is a significant variance with the actual number of patients who undergo surgery compared with a clinical database.

Third, they report selected groups of patients, namely, the high-risk patients, the Medicare patients, and the Medicare patients aged 65 years or more, who are going to have, as demonstrated, a higher mortality.

I think that using these administrative data as the barometer for determining quality solely on the basis of volume is a dangerous one.

Dr Ferguson. Third, it is ironic that after all the years of being criticized for having a proprietary risk model, we find ourselves being evaluated by APR-DRG and Healthgrades.com proprietary models. CMS has developed a risk model for CABG that is being used on administrative data without hierarchic modeling, and it is proposing to use this model in the Premier Hospital Demonstration Project. In addition to the unadjusted and risk-adjusted differences you have defined, one other important aspect of the STS database is the observed-to-expected (O/E) ratio data.

Do you have information on whether this metric was available with these other models, and can you comment on the usefulness of the O/E ratio to evaluate the quality of care?

Dr Mack. When Medicare initially released data to begin with, they stopped in 1993 because it was not risk-adjusted data, and they viewed that this was important and they stopped public release of Medicare data, or publishing this; however, it is in the public domain and available, and what these other sources do is repackaging these data and place their own proprietary risk algorithms on it.

The 3M risk algorithm, which is the APR-DRG, is totally proprietary, and they would not tell us what variables go into it or the weighting of those variables. Healthgrades.com did share with us what variables were included. There are 9 procedural codes that are included in the Healthgrades.com risk algorithm. They would not share the weighting of these, however.

Third, the STS database contains 30 clinical variables. The ones that are higher in the hierarchy, if you will, are known; however, the exact weighting, as you have alluded to, is proprietary and not known. I realize there are a lot of issues with making this nonproprietary, but I think that that is important because it is only the real measure of risk adjustment. In addition to being dependent on the coder to take this data, there is an inability to distinguish between administrative data on a preoperative morbidity and a postoperative complication (eg, renal failure). Is that a risk factor or a complication? Administrative databases cannot sort that out, and that significantly affects the accuracy of the risk algorithm and therefore the O/E ratio.

Dr Ferguson. My comment is to again congratulate Dr Mack and his colleagues on an important contribution that in terms of the National Quality Forum Cardiac Surgery Performance Measures Project, which is ongoing, could not be more timely.

The last and perhaps most important difference between clinically driven data sets and administrative data sets is that although both are purported to evaluate metrics of quality, the former is the only one that can be actually used for continuous quality improvement.

Dr Stephen J. Lahey (Worcester, Mass). As a long-time representative of the Northern New England Consortium, I am pleased to see this article, and obviously you have done an excellent job. I wonder if this is a call to action for the various committees of the STS and The American Association for Thoracic Surgery? We began speaking of this issue of maybe educating the public, politicians, and businesses about interpretation of data.

In the Northern New England Consortium, our latest CABG mortality is now less than 1%. This rate is in institutions that would never pass Leapfrog muster as far as volume. So it is clear to all of us in the Northern New England Cardiovascular Disease Study Group that volume is not a surrogate for quality. So your article shows that we have a lot of work to do here but that we have to do it very, very quickly. Is this what we should be doing to educate the public?

Dr Mack. I think there are a couple of answers to that question. One is there are multiple stakeholders who want to know outcomes data, and not all of the data are clinical. So I think there should be some integration between financial outcomes and clinical outcomes.

Second, I think that you have to be careful what you wish for, it may come true, and if we want to replace our own clinical data in the public domain, then that means (I realize in New York State that individual data are actually being published on individuals) we need to be prepared that our clinical data are out there and can be accessed and analyzed by anybody. Although the STS does a wonderful job of it right now, I am not so sure that everybody else would with those data.

The advantage of clinical data is that they are relevant and timely. The problem is that the data are expensive to collect. The advantage of administrative data is that they are already out there, they are out there in large volumes, and they are inexpensive because they are readily available, but the data are dated by the time they are made available. That is why we picked the study year of 2001, because that is the last year that data were available from all of those sources on our patients.

ON THE MOVE?

Send us your new address at least six weeks ahead

Don't miss a single issue of the journal! To ensure prompt service when you change your address, please photocopy and complete the form below.

Please send your change of address notification at least six weeks before your move to ensure continued service. We regret we cannot guarantee replacement of issues missed due to late notification.

JOURNAL TITLE:

Fill in the title of the journal here. _____

OLD ADDRESS:

Affix the address label from a recent issue of the journal here.

NEW ADDRESS:

Clearly print your new address here.

Name _____

Address _____

City/State/ZIP _____

COPY AND MAIL THIS FORM TO:

Elsevier Inc.
Subscription Customer Service
6277 Sea Harbor Dr
Orlando, FL 32887

OR FAX TO:

407-363-9661

OR E-mail:

elspcs@elsevier.com

OR PHONE:

800-654-2452

Outside the U.S., call

407-345-4000