International Journal of Service Science, Management, Engineering, and Technology
Volume 15 • Issue 1

Evaluating Cardiac Surgeon Performance:
A Retrospective and Exploratory Study

Jon A. Chilingerian, Heller School for Social Policy and Management, USA
Mitchell P. V. Glavin, Meehan School of Business, Stonehill College, USA*
https://orcid.org/0009-0003-5269-2184

ABSTRACT

This paper uses data envelopment analysis to explore how to evaluate physician and hospital clinical efficiency for coronary artery bypass surgery. First, the DEA models (measuring overall technical and scale efficiency partitioned by severity) identified many inefficient surgeons. Regardless of time period studied, relatively few surgeons were found to be on the best-practicing production frontier (DEA efficiency score = 1). The authors offer some evidence that clinical efficiency may be subject to investing in organizational capabilities that arise from operational strategies such as developing open heart surgery as a specialty, years of experience, training, and ultimately, acquiring a cadre of "efficient" surgeons. At a minimum, these findings support including some measures of "superior" organizational capabilities, strategic focus or product specialization, continuing education, and experience in future work.

KEYWORDS

Date Envelopment Analysis, Measuring Physician Efficiency, Physician Performance, Surgical Performance

INTRODUCTION

Cardiovascular disease (CVD) is a leading cause of death and disability around the globe, as well as a major economic burden contributing to unemployment, disability, lost productivity, and rising health costs (Mensah & Brown, 2007; WHO, 2021). In the U.S., the total direct and indirect costs of CVD and stroke were $555 billion in 2016 and projected to rise to $1.1 trillion by 2035. An estimated 45 percent of the U.S. population will have at least one cardiovascular disease condition by 2035 (American Heart Association, 2017).

The first-line treatment for many people with CVD is coronary artery bypass graft (CABG) surgery, which is among the most common and expensive operative procedures (AHRQ, 2021; Del Rizzo et al., 1998; Wilson et al., 2007). It is also one of the most profitable hospital procedures. These CVD facts raise the question of whether these programs and procedures are effective and efficient. Two CABG trends in the U.S. have been reported in the literature: (1) the number of hospitals performing CABGs has been increasing, and (2) the number of CABG procedures performed annually has been

DOI: 10.4018/IJSSMET.335948 *Corresponding Author

declining (Wilson et al., 2007). According to the Agency for Healthcare Research and Quality's Healthcare Cost and Utilization project, the annual volume of CABG procedures declined nearly 50% between 2001 and 2011, with 207,300 CABG procedures performed in 2011 declining further to 201,600 by 2018 (Agency for Healthcare Research and Quality, 2021; Weiss & Elixhauser, 2014). Although the efficacy of the CABG procedure has been well-established—it prolongs life for some patients and improves the quality of life for most—the efficiency and productivity changes of hospital CABG programs have not been documented.

Hospital managers and physicians are coping with the growing rivalry for CVD patients by developing new marketing approaches. However, this rivalry and threats of alternative and substitute therapies and medical procedures may mean that hospital and clinical service line managers will soon face even more serious accountability and performance challenges. Hospitals are a good example of a multilevel enterprise with centralized management but different levels of clinical units and collaboration among a diverse array of professionals mobilized to provide specific clinical service lines such as coronary artery bypass graft surgeries, the focus in this paper (Sisodia & Agrawal, 2019, Barhoun et al., 2019; Srivastava, 2021; Barros, 2022).

CABG program performance is coming under sharper scrutiny because recent research studies have uncovered striking differences in cost and quality, and interest groups are using these findings to pressure hospitals to change (Anderson et al., 2003; Cowper et al., 2002; Mensah & Brown, 2007; Nagle & Smith, 2004). Future government policy will be aimed toward supporting CABG programs that are more efficient and more effective. Despite much activity and effort, only a few CABG programs have been singled out as superior. In Pennsylvania, for example, only 5 out of 39 hospitals were designated as superior performers in 1999, as measured by lower costs, mortality, morbidity, and short lengths of stay (HCIA, 1999). Studies of physician performance often reveal wide variations in efficiency levels—some are much more successful than others. Why is it that some physicians perform more effectively than their peers? If a CABG surgeon can get better results, what explains her or his relative effectiveness?

In this paper, to illustrate how to evaluate cardiac surgery at the individual physician level, we explore the performance of CABG surgeons operating in Pennsylvania using historical data from 1994-95 and 2003-4. The research objective is to evaluate the performance and analyze the factors that explain performance using a methodology called Data Envelopment Analysis (Banker et al., 1984; Charnes et al., 1978; Charnes et al., 1994; Chilingerian, 1989; Chilingerian, 1995; Chilingerian & Sherman, 1990; Chilingerian & Sherman, 2011; Cooper et al., 2007; Zhu, 2000; Barros & Aguilera, 2022). To study cardiac surgeon performance, the theory-based mathematic called Data Envelopment Analysis (DEA) is used as a tool for health care policy makers and managers to measure and evaluate the relative performance of clinicians, clinics, and health care organizations. DEA is a powerful performance evaluation methodology that identifies top performers in relation to less effective performers. The methodology can handle multiple, non-commensurate[1] clinical inputs and outputs, including qualitative factors such as patient satisfaction. DEA can measure and evaluate the performance of different clinical decision-making units when the care process involves multiple inputs and outputs (Banker et al., 1984). Most importantly, DEA estimates a single summary measure of relative performance without requiring *a priori* weights.

The study demonstrates how to study complex cardiovascular healthcare services at the individual physician level over two time periods from a retrospective data set. We constructed a study file from an amalgam of databases containing the attributes of individual physicians, patients, hospitals, and the markets in which they operate. Although we have more recent data, obtaining all the physician-level information on characteristics such as years out of medical school, where they trained, and how many continuing medical education hours they achieved became difficult. While all these data exist, matching these characteristics to a unique physician identifier is impossible. As we explored several essential variables in our conceptual model, we could test some exploratory hypotheses with that unique historical data set.

Our purpose is two-fold. First, to demonstrate how to measure and evaluate cardiac surgeon performance to understand the variables or mechanisms associated with clinical efficiency and performance change. Second, to show that an informative and more comprehensive data set can uncover new hypotheses for future studies.

LITERATURE AND CONCEPTUAL MODELS

Literature

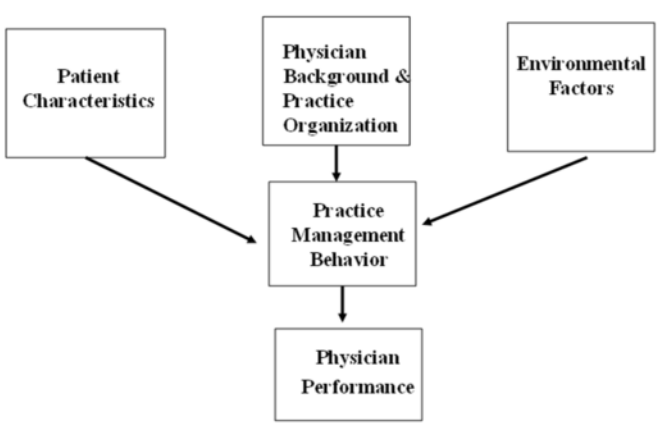
Evaluating a healthcare provider's clinical efficiency requires finding "best practices"—i.e., the minimum set of inputs to produce a successfully treated patient. Clinical inefficiency occurs when a provider uses a relatively excessive quantity of clinical resources (inputs) compared to providers practicing with a similar number and mix of patients. Hence, surgeons will be considered 100% efficient if they care for patients with fewer days of stay and ancillary services and achieve constant quality outcomes.

Previous research on physician utilization has found relationships between the use of resources and factors such as patient case mix, physician specialty training, and practice organization. Eisenberg (1986) suggests that models of physician performance should consider patient characteristics, physician characteristics, practice behavior, and the general context of the practice setting and environment (Figure 1). This model was further developed by Chilingerian and Glavin (1994). These factors can affect the utilization of clinical resources (practice management behavior) and patient management outcomes (quality and cost). This study was organized around the conceptual model displayed in Figure 1.

Multiple patient characteristics have been linked to CABG surgery efficiency and outcomes. Patient-related variables such as gender and age have been known to affect outcomes following CABG procedure. For example, women tend to be more high-risk patients, more often older and have more significant comorbidity (Capdeveille et al., 2001; Rankin, 1990; Regitz-Zagrosek et al., 2004; Vaccarino et al., 2003). Differences in race have also been shown to predict patient outcomes. Black patients more often experience operative mortality and post-operative complications (Bridges et al., 2000; Rumsfeld et al., 2002). Patients residing in the lowest affluent quartile of their area may be twice as likely to experience a complication (Gibson et al., 2009). However, Kim et al. (2007) found that the poorer outcomes experienced by low-SES patients do not exist in low-volume hospitals.

In addition to patient characteristics, other variables can impact patient outcomes following CABG. Several studies have demonstrated that larger-volume hospitals yield better outcomes in

Figure 1. Factors influencing patient management outcomes



CABG procedures, particularly measured in mortality (Birkmeyer et al., 2002; Rathore et al., 2004). Improved outcomes may also be achieved in low-volume hospitals where high-volume surgeons practice (Papadimos et al., 2005). As hospitals develop specialized cardiac surgery programs, this diverts caseload and potentially displaces knowledge away from other hospitals (Wilson et al., 2007). This distribution reduces a hospital's ability to meet the Leapfrog Group for Patient Safety recommended annual CABG volume of 450 procedures. In addition to greater patient mortality, low-volume hospitals experience higher costs (Auerbach et al., 2010; Ho & Petersen, 2007). As the number of CABG procedures being performed is falling and the number of hospitals performing procedures is increasing, patient outcomes are potentially vulnerable.

Physician specialty and practice organization have been correlated with using consultations, hospitalizations, diagnostic tests, and so on (Eisenberg, 1986). Every physician possesses a body of knowledge and a repertoire of skills accumulated through years of education, training, and practice experience (Eisenberg, 1986). Comparative studies of seasoned physicians and medical students have found that with experience, physicians recall more critical cues and can differentiate more relevant from irrelevant information (Coughlin & Patel, 1987). Eisenberg and Nicklin (1981) found that the number of years since medical school graduation was negatively associated with the use of radiologic and laboratory tests.

The experience effect of treating large numbers of patients (volume irrespective of severity) suggests that efficiency should be higher for physicians caring for more patients—i.e., carrying larger caseloads (Lin et al., 2007). Physicians who handle larger hospital caseloads will spend more time managing patients on-site and should be more adept at processing information from nurses and technicians.

Research has shown that surgeons performing coronary artery bypass surgery become more efficient with experience (Cromwell et al., 1990; Hannan et al., 2003). Physicians manage patients at every hospital where they admit patients. However, some physicians practice exclusively at one hospital, while others admit patients to several hospitals (Burns et al., 1989). Physicians who admit patients to multiple hospitals inevitably dilute their managerial attention and oversight capability. Therefore, physicians who concentrate their admissions in fewer hospitals should prove more efficient than physicians who operate in a greater number of settings.[2] Finally, training and education can also influence physician behavior. For example, Andriole et al. (1998) found that even an abbreviated continuing medical education course resulted in long-term changes among surgeons attending a one-day workshop on surgical education.

Studies have long documented performance differences between fee-for-service and HMO-affiliated physicians. Some observers argue that prepayment arrangements in HMOs encourage physicians to be more efficient by shifting part of the financial risk to physicians, since HMO viability depends on efficient care (Hillman et al., 1987). One study found that HMO physicians treating inpatients made fewer patient referrals to consultants than fee-for-service physicians (Luke & Thomson, 1980). Similarly, hospitals that receive large shares of revenues from capitated or bundled pricing arrangements can be expected to reduce service utilization, staffing levels, and so forth.

In addition to using the current study to replicate our previous research on CABG surgery in Pennsylvania, we wanted to extend our surgeon-level analysis in 2003-2004 to take advantage of some new potential explanatory variables that were not available. One new variable is MEDINC, which represents the median family income for FY2004 in the Office of Management and Budget's Core-Based Statistical Area (CBSA), where the surgeon's primary hospital is located. We included this environmental control variable to help account for regional price differences not fully accounted for by the hospital charges adjusted by facility-specific cost-to-charge ratios. MEDINC also can help control for regional variations in the intensity of health care resource use where higher family incomes are typically associated with increased health care utilization.

The other two new variables we tested in the Tobit analysis for 2003-2004 were SALMDPCT and HSPTL.[3] The former represents the percentage of all full-time equivalent employees at a hospital that are salaried physicians. A higher value of SALMDPCT should indicate greater alignment between a hospital and the physicians there, potentially generating more efficient care practices. The HSPTL variable is a (0,1) variable indicating whether or not a hospital has adopted the hospitalist physician model for treating inpatients. The hospitalist model became more common after 1994-1995 and has been associated with more efficient care, so we wanted to investigate its association with CABG surgeon efficiency.

Hypotheses

The analysis of physician and hospital services requires adequate measures of case-mix complexity and severity (Brewster et al., 1985; Landon et al., 1996; Steen et al., 1993). In addition to complexity and severity measures, other variables such as age, gender, and race are helpful to account for population and health differences. Based on the conceptual framework presented in the preceding section, the study will focus on testing the following hypotheses:

H1: Controlling for case mix, cardiac surgeons, and cardiac surgery units with higher volumes of CABGs will be more efficient.
H2: Controlling for case mix, cardiac surgeons with more years of experience will be more efficient.
H3: Controlling for case mix, cardiac surgeons who complete more than 50 hours of continuing medical education will be more efficient.
H4: Controlling for case mix, cardiac surgeons, and cardiac surgery units treating a greater proportion of HMO patients will be more efficient.
H5: Controlling for case mix, cardiac surgeons practicing in hospitals that specialize or focus on open heart surgeries in relation to other admissions will be more efficient.
H6: Controlling for case mix, cardiac surgery units that focus on open heart surgeries in relation to other admissions will be more efficient.
H7: Controlling for case mix, cardiac surgeons who concentrate more on the CABG procedures they perform at a single facility will be more efficient.
H8: Controlling for case mix, cardiac surgeons who perform CABG procedures in teaching hospitals will be less efficient.
H9: Controlling for case mix, cardiac surgeons whose primary hospital for CABG procedures has a greater percentage of physicians among total salaried employees will be more efficient.
H10: Controlling for case mix, cardiac surgeons whose primary hospital utilizes hospitalist physicians to treat inpatients will be more efficient.

METHODS

Data Sources

In the study, we compiled our final data set by matching the CABG data supplied by the Pennsylvania Health Care Cost Containment Council with the American Medical Association (AMA), American Hospital Association (AHA), and Pennsylvania Hospital Association (PHA) data available for Pennsylvania physicians. The chief data source was a database of all CABGs performed in Pennsylvania hospitals in 1994-1995 and 2003-2004. These data were obtained from the Pennsylvania Health Care Cost Containment Council (PHC4, 1998, 2005, and 2006), an independent state agency. The second data source employed in the study was the American Medical Association's Directory of Physicians in the United States (AMA, 1996; AMA, 2004). This directory includes both AMA members and non-AMA members. The Pennsylvania Hospital Association's annual hospital survey for 2004-2005 and the American Hospital Association's Annual Survey of Hospitals Database for 1995-2005 represented additional data sources for the study. The PHA and AHA databases contain data items detailing factors such as each hospital's organizational structure, staffing levels, facilities, and services (AHA, 1996).

Sample Construction

For the surgeon-level analysis, we removed all cases handled by surgeons with less than 30 cases (combined) of DRG 106 and DRG 107 in 1994 and 1995.[4] The final 1994-1995 database for the surgeon efficiency analyses thus contained 28,613 discharges for DRGs 106 and 107, representing 139 surgeons. We proceeded analogously with the 2003-2004 data, removing all cases handled by surgeons with less than 30 cases (combined) of DRG 107 and DRG 109 in 2003 and 2004. The final 2003-2004 database for the surgeon efficiency analyses thus contained 19,246 discharges for DRGs 107 and 109, representing 116 surgeons.

For DRG 106 and 107 in 1994-1995 and DRG 107 and 109 in 2003-2004, we classified as high severity approximately 20 percent of all cases with the highest estimated probability of death. For 1994-1995, high-severity cases had an estimated probability of death greater than or equal to 3% ($p \geq 0.03$). In 2003, high-severity cases had an estimated probability of death greater than or equal to 2.5% ($p \geq 0.025$), while high-severity cases in 2004 had an estimated probability greater than or equal to 2.9% ($p \geq 0.029$).

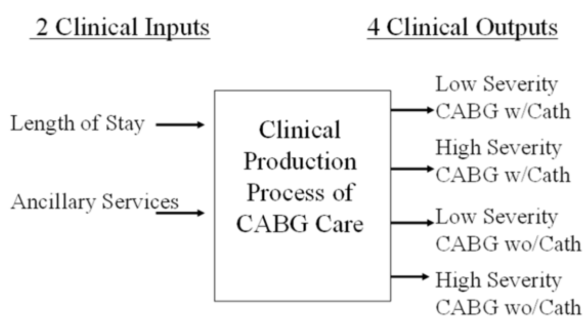
Defining Clinical Inputs and Outputs for Efficiency Estimates

For the surgeon-level analysis, we used a two-input, four-output clinical production model (Figure 2). Surgeons and hospitals are assessed relative to their peers according to how efficiently they convert clinical inputs into the outputs of completed cases. The two inputs are defined as (1) the total length of stay in days (LOS) for the CABG cases handled and (2) the total ancillary and other charges (dollars) for the CABG cases handled. The ancillary and other charges input includes ancillary, drug, equipment, and miscellaneous charges. The first input, length of stay, represents a measure of the duration of CABG admissions and the utilization of clinical inputs such as nursing care and support services. The second input, ancillary and other charges, represents factors such as the intensity of CABG admissions and the utilization of operating rooms, laboratory and radiological testing, and drugs.

The cost-to-charge ratios from the Medicare Cost Reports for each Pennsylvania hospital were applied to the raw totals for ancillary, drug, equipment, and miscellaneous charges to account for facility-specific reimbursement rates for the same DRGs.[5] All charge data for 1994, 1995, and 2003 were converted into 2004 dollars using the all-payer inflation factors from the Office of the Actuary, Centers for Medicare and Medicaid Services (Office of the Actuary, 2007).

The four classes of clinical outputs represent completed CABG surgery cases. Since patients with more severe clinical conditions will likely require more clinical inputs, the efficiency analysis must account for variations in case mix to be fair to surgeons or hospitals treating relatively sicker CABG patient populations. Accordingly, we set up the two CABG DRGs as separate clinical outputs and (as explained above) divided each DRG into low-severity and high-severity classes, as indicated in Figure 2.

Figure 2. Production model for evaluating CABG surgery



Explanatory and Control Variables for Estimates of Surgeon Efficiency

The rationale for the explanatory and control variables was explained above. The variables and definitions are as follows.

Control Variables

1. **AVGAGE:** This variable represents the mean age of the patients treated by each surgeon.
2. **PCTFEMALE:** This variable represents the percentage of male patients treated by each surgeon.
3. **PCTWHT:** This variable represents the percentage of patients treated by each surgeon who are white.
4. **PCTHMO:** This variable represents the percentage of patients treated by each surgeon insured through a health maintenance organization.[6]

The control variables are included in the analysis to account for any case mix differences among surgeons not captured by the admission severity classifications described above. However, given the rigor of the probability of death estimate developed by the Pennsylvania Health Care Cost Containment Council for this specific patient population, we had anticipated in our previous research that one or all of these control variables would be insignificant factors in explaining surgeon efficiency. This was the case, so we proceeded similarly with the 2003-2004 surgeon-level analysis.

Explanatory Variables

1. **OPHTOT:** This variable represents the total number of open-heart surgeries (not only CABGs) performed by each surgeon.
2. **YRSPRA:** This variable represents the number of years since graduation from medical school.
3. **CME:** This variable indicates whether the surgeon received the AMA's Physician Recognition Award for completing at least 50 hours of continuing medical education annually (0/1).
4. **HOSPCT:** This variable represents the percentage of all CABGs handled by a surgeon performed in the surgeon's primary (most commonly used) hospital.
5. **TEACH:** This variable indicates whether the surgeon's primary hospital was a teaching hospital.
6. **HMIX:** This variable represents the percentage of all admissions (all DRGs) at the surgeon's primary hospital for open-heart surgery.[7]
7. **MEDINC:** This variable is the median family income for FY2004 in the Office of Management and Budget's Core-Based Statistical Area (CBSA), where the surgeon's primary hospital is located.
8. **SALMDPCT:** This variable represents the percentage of all full-time equivalent employees at a hospital that are salaried physicians.
9. **HSPTL:** This variable indicates whether or not a hospital has adopted the hospitalist physician model for treating inpatients.

Below is a summary of the values observed for the explanatory and control variables for the study population of 139 Pennsylvania surgeons for 1994-1995 (Table 1). The data show that caseloads, on average, were predominantly male (73%) and white (92%), with few patients (10%) enrolled in health maintenance organizations. The average surgeon has practiced for 20.6 years and performed about 338 open-heart surgeries from 1994 to 1995.

As discussed above, the study population of 116 Pennsylvania CABG surgeons in 2003-2004 were treating remarkably similar patients as in 1994-1995 in terms of average age, percent male, and percent with race white (Table 2). The percentage of patients enrolled in an HMO was much higher (31%) in 2003-2004. The average CABG surgeon was slightly more experienced (22.1 years in practice) in 2003-2004 but performed less than half as many open-heart surgeries (161.5 vs. 338.1 in 1994-1995).

In the next section, we report on the data sources and analytical results.

Table 1. Summary of input and output measures and factors affecting surgeon efficiency, 1994-1995

Variable	Mean	Standard Deviation	Minimum	Maximum
Outputs:				
Low severity cases (DRG 106)	105.37	51.98	15.00	259.00
High severity cases (DRG 106)	25.24	16.25	1.00	100.00
Low severity cases (DRG 107)	66.11	38.38	7.00	197.00
High severity cases (DRG 107)	9.12	6.32	0.00	38.00
Inputs:				
Total length of stay (days)	2,004.02	875.73	487.00	4,486.00
Total Ancillary, Drugs, Equipment, & Misc. Charge	$3,186,161.58	$1,627,263.96	$810,475.63	$8,878,731.32
Case mix factors:				
Average age of patients (AVGAGE)	64.56	1.23	60.85	67.61
Percentage of patients female (PCTFMALE)	0.27	0.04	0.15	0.38
Percentage of patients white (PCTWHT)	0.92	0.08	0.58	1.00
Percentage of patients in HMOs (PCTHMO)	0.10	0.08	0.00	0.36
Surgeon factors:				
Total open heart surgeries performed (OPHTOT)	338.06	138.76	91.00	771.00
Years in practice (YRSPRA)	20.64	7.56	9.00	44.00
Extra CME award (1=Yes) (CME)	0.50	0.50	0.00	1.00
Percentage of CABGs at primary hospital (HOSPCT)	0.89	0.17	0.24	1.00
Hospital factors for surgeon's primary hospital:				
Teaching hospital (1=Yes) (TEACH)	0.54	0.50	0.00	1.00
Percentage of 1995 admissions for open heart surgery (HMIX)	0.04	0.02	0.01	0.09
Total observations = 139				

Table 2. Summary of input and output measures and factors affecting surgeon efficiency, 2003-2004

Variable	Mean	Standard Deviation	Minimum	Maximum
Outputs:				
Low severity cases (DRG 107)	83.34	38.10	11.00	207.00
High severity cases (DRG 107)	19.86	9.40	4.00	46.00
Low severity cases (DRG 109)	48.66	24.03	7.00	112.00
High severity cases (DRG 109)	7.60	5.54	1.00	29.00
Inputs:				
Total length of stay (days)	1,246.54	433.95	390.00	2,589.00
Total Ancillary, Drugs, Equipment, & Misc. Charge	$1,890,932.06	$736,191.33	$549,256.34	$3,929,059.13
Case mix factors:				
Average age of patients (AVGAGE)	64.54	1.29	61.41	68.43
Percentage of patients female (PCTFMALE)	0.24	0.04	0.09	0.38
Percentage of patients white (PCTWHT)	0.91	0.11	0.38	1.00
Percentage of patients in HMOs (PCTHMO)	0.31	0.14	0.04	0.62
Surgeon factors:				
Total open heart surgeries performed (OPHTOT)	161.48	56.86	52.00	292.00
Years in practice (YRSPRA)	22.09	7.60	9.00	44.00
Extra CME award (1=Yes) (CME)	0.10	0.30	0.00	1.00
Percentage of CABGs at primary hospital (HOSPCT)	0.91	0.11	0.38	1.00
Hospital factors for surgeon's primary hospital				
Teaching hospital (1=Yes) (TEACH)	0.42	0.50	0.00	1.00
Percentage of 2004 admissions for open heart surgery (HMIX)	0.02	0.01	0.01	0.05
Total observations = 116				

FINDINGS AND RESULTS

Findings

The current research examines the performance of CABGs a decade later in the 2003-2004 period. Table 1 shows that the average surgeon handled approximately 206 CABG cases in 1994-1995, with the following distribution across the four clinical output classes: 105 low severity DRG 106 cases, 25 high severity DRG 106 cases, 66 low severity DRG 107 cases, and nine high severity DRG 107 cases. In caring for these 206 total cases, the average surgeon used as clinical inputs 2,004 inpatient days and $3.2 million in ancillary and other charges. The mean length of stay per patient for the average surgeon was 9.7 days and the mean ancillary and other service charges per patient were $15,478.83.[8]

The more recent data in Table 2 shows that the average surgeon in 2003-2004 performed approximately 159 CABG surgeries. This decline in caseload from the 1994-1995 period in Pennsylvania is consistent with the national trend of declining volume of CABGs. The average surgeon in 2003-2004 used as clinical inputs 1,247 inpatient days and $1.9 million in ancillary and other charges.[9] The average length of stay per patient for the average surgeon was 7.8 days, and the mean ancillary and other service charges per patient were $11,858.35. Pennsylvania, therefore, has experienced a reduction in average LOS of about 1.9 days and a reduction in ancillary and other charges (expressed in 2004 dollars) per admission of about $3,620.48 for CABG surgery.

Tables 1 and 2 suggest that the patient population undergoing CABGs in 2003-2004 is largely unchanged from 1994-1995. The average age of patients having the surgery, as well as the percentage of patients who are female and the percentage who are white, are nearly identical in both periods. This result is surprising given that the national trend has been toward performing CABGs on more elderly patients, as discussed above. However, the percentage of patients insured through a health maintenance organization rose from 10% in 1994-1995 to 31% in 2003-2004. This rise reflects the national trend towards greater enrollment in managed care insurance plans and away from traditional indemnity-type health insurance plans.

In the bottom section of Tables 1 and 2, we present data on some surgeon factors we identified in our previous research as potential factors affecting CABG surgeon efficiency. We can see that the average surgeon in 2003-2004 performed fewer open-heart surgeries than in 1994-1995: 161.5 open-heart surgeries vs. the earlier aver-age of 338 open-heart surgeries. There are also fewer CABG surgeons (116) operating in 2003-2004 than in 1994-1995 (139 surgeons). In terms of years in practice, the average surgeon in 2003-2004 had 22.1 years, while the average surgeon in 1994-1995 had 20.6 years. Half of the Pennsylvania CABG surgeons in 1994-1995 had earned the AMA's Physician Recognition Award for completing at least 50 hours of continuing medical education annually. In 2003-2004, this percentage declined to 10%.

For both periods, the average CABG surgeon performs about 90% of surgeries in one hospital, indicating that Pennsylvania CABG surgeons primarily focus their practice within a single facility. In 1994-1995, the primary hospital was a teaching hospital for 54% of the CABG surgeons. In addition, the total number of open-heart surgeries performed (by all surgeons) in a hospital represented 4% of total admissions to that facility (see the bottom of Table 1). Moving to the 2003-2004 period in Table 2, the primary hospital was a teaching hospital for only 42% of CABG surgeons. This split likely reflects the greater maturity and diffusion of the CABG technology to more hospitals across Pennsylvania. As noted above, we have 43 hospitals in our 1994-1995 study database and 60 hospitals in our 2003-2004 database.

Identification of Factors in Clinical Efficiency

This component of the study uses multivariate Tobit analysis of surgeon DEA scores to explore factors that explain differences in efficiency. The specific factors examined included patient, surgeon, and hospital factors. Therefore, the following general model was used:

DEA Score = f (patient, surgeon, and hospital characteristics)

The explanatory and control variables for the surgeon-level multivariate analyses are detailed above. DEA scores for surgeons were computed using discharge data for 1994 and 1995 CABG surgeries together.[10] This approach maximized the per-surgeon CABG caseload quantities and the potential explanatory power of the multivariate analyses.

The physician sample for this study has observations on the DEA efficiency scores and explanatory/control variables for 139 surgeons in 1994-1995 and 116 surgeons in 2003-2004. The DEA scores are a mixture of discrete and continuous parts. Inefficient surgeons have DEA scores that fall into a wide variation of strictly positive values less than 1. Surgeons practicing CABG surgeries at the clinically efficient production frontier all have efficiency equal to 1 because DEA has an upper-value limit of 1. A multivariate regression model other than ordinary least squares (OLS) is required to estimate the coefficients in an equation explaining efficiency. The OLS method assumes a normal and homoscedastic distribution of the disturbance and the dependent variable. When the distribution of the dependent variable is censored at some limiting value, the expected errors will not equal zero. Applying OLS in such circumstances will lead to a biased estimate (Maddala, 1983).

The DEA scores for surgeons performing CABG surgeries thus display a censored distribution where the censoring takes the following form:

Efficiency score = actual score when score < 1
Efficiency score = 1, otherwise

As a result, the distribution of surgeons' DEA scores includes a significant number of observations at the limiting value of 1. Given this condition for the dependent variable, a Tobit-censored regression model is an appropriate method for the multivariate analysis (Chilingerian, 1995).

Greene (1993) has demonstrated that a convenient normalization for computational reasons in Tobit studies assumes a censoring point at zero. To put the surgeon efficiency data into this form, the DEA scores were transformed using the following form:

Inefficiency score = $(1/\theta) - 1$

where:

θ = DEA efficiency score

The result is the dependent variable (DEA inefficiency score), which takes the following form:

For $\theta = 1$
DEA inefficiency score = 0
For all other values of θ, $0 <$ DEA inefficiency score $\leq \infty$

Once DEA scores have been transformed, Tobit regression is readily used to understand variations in efficiency. The slope coefficients for independent variables in Tobit are interpreted as if they were obtained using ordinary least squares regression. A slope coefficient represents the change in the dependent variable (DEA inefficiency score) with respect to a one-unit change in the independent variable, holding all else constant. A log-likelihood ratio statistic is used to test the significance of the Tobit model equation, much like an F-test for OLS (Judge et al., 1985). The log-likelihood ratio test has a chi-square distribution, with degrees of freedom equal to the number of independent variables used in the regression (Chilingerian, 1995).

Results

For each of the 139 surgeons in the study population for 1994-1995, two DEA efficiency scores were obtained: one score for 1994 and one score for 1995. The DEA model for CABG production was composed of two inputs (days of stay and ancillary/other charges) and four outputs (low severity DRG 106 cases, high severity DRG 106 cases, low severity DRG 107 cases, and high severity DRG 107 cases). The 1994 efficiency scores (Theta94) were compared to the 1995 efficiency scores (Theta95). The average Theta94 score was 0.880 (standard deviation = 0.084), while the average Theta95 score was 0.869 (standard deviation = 0.095). As anticipated, Theta94 turned out to be highly correlated with Theta 95. The Spearman's rho correlation was 0.873, significant (2-tailed) at the 0.01 level. For 1994, 21 surgeons out of 139 were located on the best practice frontier and received a DEA efficiency score of 1.000. In 1995, 22 surgeons out of 139 were located on the best practice frontier.

For each of the 116 surgeons in the study population for 2003-2004, two DEA efficiency scores were obtained: one for 2003 and one for 2004. Once again, the DEA model for CABG production was composed of two inputs (days of stay and ancillary/other charges) and four outputs (low severity DRG 107 cases, high severity DRG 107 cases, low severity DRG 109 cases, and high severity DRG 109 cases). The 2003 efficiency scores (Theta03) were compared to the 2004 efficiency scores (Theta04). The average Theta03 score was 0.825 (standard deviation = 0.12), while the average Theta04 score was 0.830 (standard deviation = 0.12). As anticipated, Theta03 turned out to be highly correlated with Theta04. The Spearman's rho correlation was 0.680, significant (2-tailed) at the 0.01 level. For 2003, 19 surgeons out of 116 were located on the best practice frontier and received a DEA efficiency score of 1.000. In 2004, 19 out of 116 surgeons were located on the best practice frontier.

Multivariate Analysis of Surgeon Efficiency

A matrix of Pearson correlations among the explanatory and control variables included in the multivariate analysis of surgeon efficiency for 1994-1995 is below (Table 3). The percentage of patients in HMOs (PCTHMO) was significantly correlated with the average age of the patient caseload (AVGAGE) and the percentage of patients who are white (PCTWHT). The negative correlation between PCTHMO and AVGAGE is as expected since HMOs tend to enroll younger populations.

Table 3. Pearson correlations among factors affecting surgeon efficiency 1994-1995

(# of observations = 139)

Variables	1	2	3	4	5	6	7	8	9	10
Case Mix Factors:										
1. AVGAGE	^^^^^	0.135	0.044	-0.190*	0.097	0.170*	0.052	-0.009	0.039	0.197*
2. PCTFMALE		^^^^^	0.067	-0.097	-0.017	-0.183*	0.094	-0.178*	-0.041	0.185*
3. PCTWHITE			^^^^^	-0.413**	0.126	0.099	0.057	-0.029	0.043	0.165
4. PCTHMO				^^^^^	0.156	-0.153	-0.118	0.009	0.269**	0.136
Surgeon Factors:										
5. OPHTOT					^^^^^	-0.013	0.143	0.073	-0.012	0.364**
6. YRSPRA						^^^^^	0.051	0.126	0.101	0.099
7. CME							^^^^^	-0.144	0.116	0.163
8. HOSPCT								^^^^^	-0.062	-0.173*
Hospital Factors for Surgeon's Primary Hospital:										
9. TEACH									^^^^^	0.268**
10. HMIX										^^^^^

** = Correlation significant at $p < 0.01$ (2-tailed) * = Correlation significant at $p < 0.05$ (2-tailed)

The negative correlation between PCTHMO and PCTWHT likely indicates that non-white residents of Pennsylvania tend to be located in larger metropolitan areas where HMOs are more prevalent. PCTHMO also significantly correlates positively with teaching hospital status (TEACH). This relationship is unsurprising as HMOs are more likely to operate in larger metropolitan areas, where teaching hospitals tend to be located.

A matrix of Pearson correlations among the explanatory and control variables included in the multivariate analysis of surgeon efficiency for 2003-2004 is below (Table 4). We again observe a significant correlation (positive) between the surgeon's total open-heart surgeries performed (OPHTOT) and (HMIX), which represents the percentage of all admissions at the surgeon's primary hospital that were open-heart surgery admissions. Again, such a correlation is expected as greater open heart surgery volume at the physician level will likely lead to greater focus on such surgeries at the hospital level. The correlation between AVGAGE and YRSPRA suggests that older surgeons treat relatively older patients.

Tobit Regression of Surgeon Inefficiency Scores

As outlined above, transformed DEA efficiency scores[11] for the 139 CABG surgeons in 1994-1995 were evaluated as the dependent variable in a Tobit regression with the ten control and explanatory variables. These DEA scores were based on the two years of CABG data aggregated into a joint database. The empirical results for the Tobit model are below (Table 5).

Considering the results in Table 5, it is essential to remember that inefficiency scores are regressed in the Tobit estimation. Therefore, the signs of all the coefficients are reversed regarding the hypotheses on factors explaining variations in clinical efficiency. A positive sign on a coefficient indicates an association with inefficiency, while a negative sign indicates an association with efficiency.

First, we observe that the constant is not significant at the 5% level, while the chi-square statistic is significant at the 1% level. Thus, the set of independent variables explains a significant amount of the variance in the dependent variable. Of the four case-mix variables (AVGAGE, PCTMALE, PCTWHT, and PCTHMO), only PCTHMO is statistically significant. This indicates that the severity adjustment built into the 2-input, 4-output clinical production model largely compensates for case mix differences among surgeons' CABG caseloads. Higher efficiency among surgeons with greater

Table 4. Pearson correlations among factors affecting surgeon efficiency 2003-2004

(# of observations = 116)										
Variables	1	2	3	4	5	6	7	8	9	10
Case Mix Factors:										
1. AVGAGE	-----	0.098	-0.007	-0.111	0.037	0.221*	0.055	-0.132	0.018	0.022
2. PCTFMALE		-----	0.022	0.178	-0.086	-0.156	0.058	0.006	0.240*	0.114
3. PCTWHITE			-----	-0.257**	0.085	0.014	0.009	0.211*	-0.207*	0.296**
4. PCTHMO				-----	-0.048	-0.030	0.130	-0.001	0.181	-0.099
Surgeon Factors:										
5. OPHTOT					-----	0.067	-0.118	-0.157	0.011	0.244**
6. YRSPRA						-----	0.154	-0.161	0.132	-0.135
7. CME							-----	-0.047	0.052	0.119
8. HOSPCT								-----	-0.089	-0.047
Hospital Factors for Surgeon's Primary Hospital:										
9. TEACH									-----	0.069
10. HMIX										-----

** = Correlation significant at $p < 0.01$ (2-tailed) * = Correlation significant at $p < 0.05$ (2-tailed)

Table 5. Estimation results for the tobit model of surgeon efficiency 1994-1995

Explanatory & Control Variables	Coefficient	Std. Error	*t*-Statistic
Constant	0.5363	0.6098	0.09
AVGAGE	0.0045	0.0090	0.50
PCTFMALE	0.1820	0.2398	0.76
PCTWHT	-0.1908	0.1507	-1.27
PCTHMO	-0.3895	0.1609	-2.42*
OPHTOT	0.0000	0.0001	0.19
YRSPRA	0.0022	0.0014	1.57
CME	-0.0501	0.0211	-2.38*
HOSPCT	0.0068	0.0608	0.11
TEACH	-0.0044	0.0224	-0.20
HMIX	-1.712	0.6984	-2.45*
*Chi-square^*	*28.34***		

Total observations = 139 * = significance at $p < 0.05$
** = significance at $p < 0.01$

^ The chi-square statistic is based on a likelihood ratio which tests the joint significance of the independent variables. The likelihood ratio is computed as -2log(LO/LI), where LI is the value of the likelihood function for the full model as fitted and LO is the maximum value if all coefficients except the intercept are zero.

proportions of patients enrolled in HMOs may reflect the influence of incentives managed care plans use to promote more efficient service delivery.

Two explanatory variables (CME and HMIX) are statistically significant. The CME variable, indicating the annual completion of 50+ hours of continuing medical education, is associated with higher efficiency (negative coefficient). Surgeons making greater efforts to keep up with development in clinical practice may be using more efficient methods. Lastly, the HMIX variable demonstrates that higher clinical efficiency in producing CABGs is associated with surgeons using hospitals where open heart surgeries represent a greater portion of all admissions.

In the current study, we wanted to apply the same Tobit model to the 2003-2004 data used in the earlier research with the 1994-1995 data. Once again, transformed DEA efficiency scores[12] for the 116 CABG surgeons in 2003-2004 were evaluated as the dependent variable in a Tobit regression with the ten control and explanatory variables. These DEA scores were based on the two years of CABG data for 2003-2004 aggregated into a joint database. Table 6 presents the empirical results for this Tobit model.

The Tobit results for 2003-2004 are presented below (Table 6). In contrast to the earlier 1994-1995 period, the variables CME and PCTHMO are no longer statistically significant. The percentage of Pennsylvania CABG surgeons holding an AMA's Physician Recognition Award for completing at least 50 hours of continuing medical education annually was much lower (10%) in 2003-2004 than in 1994-1995 (50%), and this likely explains why this surgeon characteristic is no longer has a significant association with higher efficiency. In contrast, the average surgeon's patients were more likely to be enrolled in an HMO in 2003-2004 than in 1994-1995, with 31% of patients vs. 10%. It may be that in 1994-1995, the still relatively novel incentives toward efficient care practices fostered by managed care (especially HMOs) had a much more substantial impact on some surgeons and hospitals than others where HMO enrollment was small or non-existent. In the later 2003-2004 period, the entire state had much more exposure and experience with managed care arrangements, which may account for the fact that PCTHMO no longer has a statistically significant association with efficient practice.

Table 6. Estimation results for the tobit model of surgeon efficiency 2003-2004

Explanatory & Control Variables	Coefficient	Std. Error	*t*-Statistic
Constant	1.7344	0.8432	2.06*
AVGAGE	-0.1125	0.0127	-0.89
PCTFMALE	0.3819	0.1577	-3.47**
PCTWHT	-0.5475	0.1507	-1.27
PCTHMO	0.6982	0.1176	0.59
OPHTOT	-0.0002	0.0003	-0.74
YRSPRA	0.0004	0.0022	0.19
CME	0.0616	0.0524	1.18
HOSPCT	-0.2446	0.1349	-1.81
TEACH	-0.0180	0.0341	-0.53
HMIX	-6.597	1.8040	-3.66**
*Chi-square^*	*45.04***		

Total observations = 116   * = significance at $p < 0.05$
** = significance at $p < 0.01$

^ The chi-square statistic is based on a likelihood ratio which tests the joint significance of the independent variables. The likelihood ratio is computed as -2log(LO/LI), where LI is the value of the likelihood function for the full model as fitted and LO is the maximum value if all coefficients except the intercept are zero.

Alternatively, the “managed care backlash” may have blunted the influence of managed care plans on clinical practices in the current century. Surgeons with more female patients are strongly associated with more efficient performance. This result indicates that some aspects of case-mix differences among surgeons is not fully controlled by the severity adjustment built into the DEA model. However, the remaining case mix control variables were not statistically significant.

Most notably, in the Tobit results for 2003-2004, we see the HMIX variable having the strongest association with higher efficiency of surgeons. Surgeons are more efficient in performing CABGs when their primary hospital has open-heart surgeries, representing a more significant product line relative to overall admissions. Presumably, these facilities have assembled specific personnel and protocols to provide more efficient cardiac care over time.

The median family income for the CBSAs in which surgeons perform CABGs ranges from $43,900 to $66,300, with an average of $60,940.52. Most CABG surgeons (66%) had a primary hospital in 2003-04 that used hospitalist physicians. Regarding physicians on salary in CABG surgeons’ primary hospitals, an average of 2% of hospital employees are physicians, ranging from 0% to 19%.

We can see that the additional variables SALMDPCT and HSPTL have no significant correlations with the other variables (Table 7). The median family income variable (MEDINC) negatively correlates with the HMIX and HOSPCT variables. A given surgeon may be likelier to perform CABGs at more hospitals in more resource-rich (higher income) areas. In addition, higher-income areas tend to have higher healthcare utilization, where CABGs would likely be a smaller portion of hospital admissions (i.e., lower values of HMIX). The results of the expanded Tobit model of CABG surgeon efficiency are below (Table 8).

Two of the control variables, PCTWHT and MEDINC, are statistically significant. Surgeons with a higher percentage of white patients were associated with more efficient performance in CABG surgeries. This result may reflect some residual case-mix variation or perhaps some underlying social determinants. Surgeons practicing in areas with higher median family incomes were likelier to be less efficient. This result may reflect some regional cost variations not fully adjusted for by applying hospital-specific cost-to-charge ratios or be evidence of more resource-intensive practice styles in

Table 7. Pearson correlations among factors affecting surgeon efficiency 2003-2004

Variables	1	2	3	4	5	6	7	8	9
Case Mix Factors:									
1. AVGAGE	-----	0.098	-0.007	0.210*	0.037	-0.132	0.022	0.087	-0.098
2. PCTFMALE		-----	0.022	-0.022	-0.086	0.006	0.114	-0.129	0.009
3. PCTWHITE			-----	-0.443**	0.085	0.211*	0.296**	-0.117	-0.134
Environment Factor:									
4. MEDINC				-----	0.045	-0.215*	-0.318**	-0.097	0.040
Surgeon Factors:									
5. OPHTOT					-----	-0.157	0.244**	0.068	0.072
6. HOSPCT						-----	-0.047	-0.039	0.121
Hospital Factors for Surgeon's Primary Hospital:									
7. HMIX							-----	0.070	0.058
8. SALMDPCT								-----	0.180
9. HSPTL									-----

** = Correlation significant at $p < 0.01$ (2-tailed) * = Correlation significant at $p < 0.05$ (2-tailed)

Table 8. Estimation results for the expanded tobit model of surgeon efficiency 2003-2004

Explanatory & Control Variables	Coefficient	Std. Error	*t*-Statistic
Constant	1.6618	0.8623	1.93
AVGAGE	-0.0177	0.0130	-1.36
PCTFMALE	0.3989	0.3491	1.14
PCTWHT	-0.5418	0.1581	-3.43**
MEDINC	7.61e-06	3.16e-06	2.41*
OPHTOT	-0.0005	0.0003	-1.68
HOSPCT	-0.1225	0.1367	-0.90
HMIX	-4.7605	1.7203	-2.77**
SALMDPCT	-1.2200	0.4772	-2.56*
HSPTL	-0.0545	0.0338	-1.61
*Chi-square^*	*60.86***		

Total observations = 116 * = significance at $p < 0.05$
** = significance at $p < 0.01$

^ The chi-square statistic is based on a likelihood ratio which tests the joint significance of the independent variables. The likelihood ratio is computed as -2log(LO/LI), where LI is the value of the likelihood function for the full model as fitted and LO is the maximum value if all coefficients except the intercept are zero.

relatively wealthier areas of Pennsylvania. None of the remaining factors (use of hospitalists, total volume of open-heart surgeries performed by the surgeon, percentage of CABG surgeries performed at the surgeon's primary hospital) had a statistically significant relationship to higher efficiency among CABG surgeons.

HMIX is once again a statistically significant factor, as it was in the 1994-1995 surgeon-level analysis and the previous 2003-2004 Tobit model. More efficient CABG surgeons are associated with having a primary hospital where open-heart surgeries are a higher percentage of overall admissions. The new variable SALMDPCT is also a statistically significant factor, with the expected association with more efficient surgery. CABG surgeons tend to be more efficient when associated with a primary hospital where a more significant proportion of salaried employees are physicians.

Summary of Findings for Study Hypotheses

H1: Controlling for case mix, cardiac surgeons, and cardiac surgery units with higher volumes of CABGs will be more efficient.

This hypothesis is not supported. For the Tobit regressions of DEA efficiency scores, the surgeon's total open-heart surgeries performed (OPHTOT) was not a statistically significant factor associated with efficiency for either the 1994-95 period or the 2003-04 period (see Table 5 and Table 6).

H2: Controlling for case mix, cardiac surgeons with more years of experience will be more efficient.

This hypothesis is not supported. For the Tobit regressions of DEA efficiency scores, the surgeon's years since graduation from medical school (YRSPRA) was not a statistically significant factor associated with efficiency for either the 1994-95 period or the 2003-04 period (see Table 5 and Table 6).

H3: Controlling for case mix, cardiac surgeons who complete more than 50 hours of continuing medical education will be more efficient.

This hypothesis is supported. For the Tobit regressions of DEA efficiency scores, completing more than 50 hours of CME annually (CME) was a statistically significant factor associated with greater efficiency for the 1994-95 period but not for the 2003-04 period (see Table 5 and Table 6).

H4: Controlling for case mix, cardiac surgeons, and cardiac surgery units treating a greater proportion of HMO patients will be more efficient.

This hypothesis is partially supported. For the Tobit regressions of DEA efficiency scores, the percentage of patients treated by each surgeon that are insured through a health maintenance organization (PCTHMO) was a statistically significant factor associated with greater efficiency for the 1994-95 period but not for the 2003-04 period (see Table 5 and Table 6).

H5: Controlling for case mix, cardiac surgeons practicing in hospitals that specialize or focus on open heart surgeries in relation to other admissions will be more efficient.

This hypothesis is supported. For the Tobit regressions of DEA efficiency scores, practicing in hospitals that specialize or focus on open heart surgeries concerning other admissions (HMIX) was a statistically significant factor associated with greater efficiency of surgeons for both the 1994-95 period and the 2003-04 period (see Table 5, Table 6, and Table 8).

H6: Controlling for case mix, cardiac surgery units that focus on open heart surgeries in relation to other admissions will be more efficient.

This hypothesis was not tested in the surgeon-level analyses.

H7: Controlling for case mix, cardiac surgeons who concentrate more on the CABG procedures they perform at a single facility will be more efficient.

This hypothesis is not supported. Measured via the HOSPCT variable, concentrating more on the CABG procedures performed at a single facility was not a statistically significant factor associated with the efficiency of surgeons for both the 1994-95 period and the 2003-04 period (see Table 5, Table 6, and Table 8).

H8: Controlling for case mix, cardiac surgeons who perform CABG procedures in teaching hospitals will be less efficient.

This hypothesis is not supported. Measured via the TEACH variable, the primary hospital for a surgeon's CABG procedures being a teaching hospital was not a statistically significant factor associated with the efficiency of surgeons for both the 1994-95 period and the 2003-04 period (see Table 5, Table 6, and Table 8).

H9: Controlling for case mix, cardiac surgeons whose primary hospital for CABG procedures has a greater percentage of physicians among total salaried employees will be more efficient.

This hypothesis is supported. Measured via the SALMDPCT variable, the primary hospital for a surgeon's CABG procedures with physicians representing a greater percentage of all salaried employees was a statistically significant factor associated with greater efficiency of surgeons for the 2003-04 period (see Table 8).

H10: Controlling for case mix, cardiac surgeons whose primary hospital utilizes hospitalist physicians to treat inpatients will be more efficient.

This hypothesis is not supported. Measured via the HSPTL variable, the primary hospital for a surgeon's CABG procedures being a hospital utilizing hospitalist physicians for the care of inpatients, was not a statistically significant factor associated with the efficiency of surgeons for the 2003-04 period (see Table 8).

Identification of High Efficiency, High Effectiveness (Low Mortality, Low Readmission Rates) Surgeons

Difficulty in judging the service quality of healthcare providers delivering complex clinical services is widely acknowledged (Azam et al., 2017; Mandal, 2022), but information such as the identification of high-efficiency, high-effectiveness providers as outlined here can yield valuable information for purchasers of clinical services such as private or public health insurance schemes and facilitate the providers in distinguishing themselves from competitors.

As described above, DEA efficiency scores were obtained for the 116 surgeons using all CABG surgeries performed in 2003 and 2004, where the patient was discharged home. The average efficiency score obtained was 0.83, with a standard deviation of 0.11, and the median score was 0.82. The minimum score was 0.53, and the maximum score was 1.0, which indicates that the surgeon was located on the production frontier. Sixteen surgeons had scores equal to 1.0. Surgeons with clinical efficiency scores above the 75th percentile were classified as high-efficiency surgeons, while those below the median were classified as low-efficiency surgeons.

This study developed DEA clinical effectiveness scores to complement the efficiency estimates by finding a quality-outcome frontier. For this application, a one-input, two-output model was used.

The single input was the total quantity of CABG procedures, and the two outputs were (1) the total quantity of CABG patients minus the inpatient CABG mortality rates plus the 30-day post-discharge mortality rate and (2) the total number of patients who were not readmitted 30-days later (i.e., 30-day readmission rates).[13]

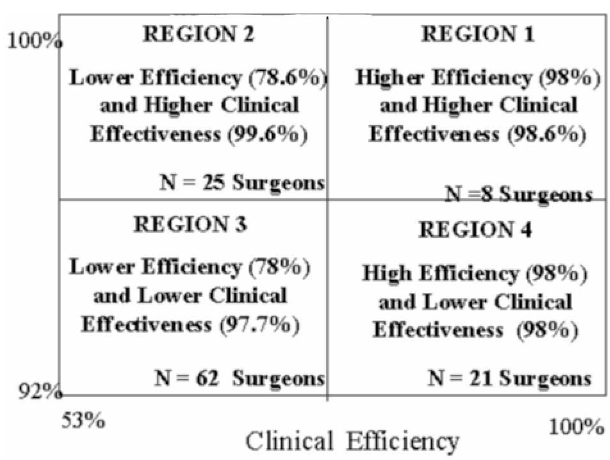
The average effectiveness score for each surgeon was then the average of the three rates (inpatient mortality, 30-day mortality, and 30-day readmission), subtracted from 1. The average effectiveness score per surgeon was 0.98 for all CABG procedures, and the standard deviation was 0.01. The minimum clinical effectiveness score was 0.92, and the maximum effectiveness score was 1. Twelve surgeons had scores equal to 1. The median effectiveness rate was .99. Surgeons with mortality rates above the 75th percentile were classified as highly effective surgeons, while those below the 75th percentile were classified as less effective surgeons.

The two-dimensional array of surgeon performance, efficiency, and effectiveness are presented below (Figure 3). Of particular interest are the eight surgeons in Region I, who display high efficiency and higher effectiveness (low mortality/readmission). This region had an average efficiency and effectiveness score of 0.98 and 0.98, respectively. By comparison, the 62 surgeons located in Region III had an average DEA score of 0.782 and an average effectiveness score of 0.977. The surgeons located in Region I could serve as potential role models for clinical improvement efforts.

DISCUSSION

The results in this report provide an exploration of physician clinical efficiency for CABG surgeries and no other procedures. Subject to the focus on CABGs, the following findings and insights are summarized from the analysis. First, the DEA models (measuring overall technical and scale efficiency partitioned by severity) identified many inefficient surgeons and hospitals. Regardless of the period studied, relatively few surgeons were found to be on the best-practicing production frontier (DEA efficiency score = 1). The amount of slack associated with these less efficient CABG surgeons is consistent with a hypothesis that a substantial amount of money could be saved if the majority of physicians could practice closer to every physician were as good as the most efficient or best-practicing physician (see Chilingerian 1989; Chilingerian, 1995; Thurow,1985). If physicians (and hospitals) could use this information to adjust their "style of practice," real clinical efficiency improvements could result.

Figure 3. Surgeon's performance reflected on clinical effectiveness and clinical efficiency



Our examination of the consistency of DEA efficiency estimates for surgeons and hospitals yielded encouraging results. Since the actual clinical efficiency of surgeons or hospitals is unlikely to vary significantly over two consecutive years, we expected DEA efficiency scores in Year 2 to be closely linked to DEA efficiency scores in Year 1, as long as "noisy" data does not seriously threaten the consistency of DEA estimates. For example, it turned out that surgeons' DEA scores in 1994 were highly correlated with these surgeons' DEA scores in 1995. In addition, receiving a DEA efficiency score of 1.0 in 1994 was found via chi-square testing to be a significant predictor of receiving a DEA efficiency score of 1.0 in the following year (1995). Eighteen of the 21 surgeons with DEA scores of 1.0 in 1994 also had a DEA score of 1.0 in 1995. The surgeons' DEA scores for 2003 and 2004 yielded similar results that indicate the consistency of DEA scores in consecutive years. Although there are alternative hypotheses, it is reasonable to propose that DEA scores are stable yearly.

To the best of our knowledge, our research represents the only healthcare analyses to specifically examine and demonstrate the consistency and stability of DEA efficiency score across consecutive years at the physician level. Examining the surgeons' performance, our results demonstrate a strong relationship between performance in one year and the following year.

One question that must be addressed in all performance studies is whether the models are muddling output differences with differences in efficiency. Using DEA in a multi-factor Tobit helps clarify what is going on. From the Tobit regression results for the 139 surgeons over the 24 months 1994-1995, among the four case-mix variables (AVGAGE, PCTFMALE, PCTWHT, and PCTHMO), only the percentage of a surgeon's patients insured through HMOs showed a statistically significant (positive) relationship to surgeon efficiency. This indicates that the severity adjustment built into the 2-input, 4-output clinical production model largely compensated for case mix differences among surgeons' CABG caseloads.

Looking at the Tobit regression results for the 116 surgeons in 2003-2004, the percentage of a surgeon's female patients had a statistically significant (positive) relationship to surgeon efficiency. AVGAGE, PCTWHT, and PCTHMO were not statistically significant. However, as shown in Table 8, for the Tobit regression that included the median family income (MEDINC) in a surgeon's practice locale as an environmental control variable, AVGAGE, PCTFMALE, and PCTHMO did not have a statistically significant association with surgeon efficiency. The percentage of surgeon's patients who were white did show a statistically significant (positive) relationship to surgeon efficiency.

Given, then, that patient case-mix differences are not the primary drivers of differences in clinical efficiency among surgeons; we considered surgeon and hospital characteristics in our Tobit regression results that have significant associations with surgeon efficiency. For the 1994-1995 period, surgeons who completed 50+ hours of continuing medical education annually (CME) were more efficient, as were surgeons with a primary hospital where open heart surgeries represented a greater proportion of total admissions (HMIX). Repeating the surgeon-level analysis for the 2003-2004 period, the CME variable no longer had a statistically significant association with surgeon efficiency, but HMIX still had a strongly positive association with efficiency (see Tables 6 and 8).

The sustained significance of the HMIX variable demonstrates that higher clinical efficiency in producing CABGs is associated with surgeons using hospitals where open heart surgeries represent a greater portion of all admissions. Hospitals have unique combinations of clinical professionals' skills and delivery capabilities. A hospital's cardiac surgery capability depends on the co-alignment of knowledge, service activities, equipment, and professionals. While each resource, such as the operating rooms, intensive care units, scrub nurses, cardiac surgeons, surgical tools, and equipment, is costly, their value depends on organizing and integrating tangible and intangible resources to create a highly productive activity. Physicians practicing in hospitals that are more heavily focused on cardiac surgeries increased their likelihood of being efficient under all conditions tested, irrespective of patient illness characteristics. While the thrust of this finding emphasizes the potentially important role that a more focused strategy could have in inefficient care and cost containment, further research is needed to learn about the difference in operating strategies and capabilities hospitals with a higher proportion of cardiac admissions develop.

Lastly, our expanded Tobit regression analysis for surgeons in 2003-2004 identified an additional factor in clinical efficiency – the percentage of hospital FTEs that were salaried physicians (SALMDPCT). Greater clinical efficiency was significantly associated with higher values of SALMDPCT (see Table 8). In other words, CABG surgeons tend to be more efficient when associated with a primary hospital where a greater proportion of salaried employees are physicians. Such hospitals presumably enjoy a greater alignment between the institution and physicians.

These findings suggest surgeons and surgical units can develop internal capabilities involving key choices. The time-honored "managerial" strategy of focusing on a more limited set of procedures (in this case, CABGs) may improve efficiency. Alternatively, since it may not be easy to turn away some hospital cases, hospitals with a higher than average number of CABGs in relation to other admissions specialize and become more efficient. Alignment between a hospital and the physicians practicing within it may also be a crucial capability, and having more physicians as salaried employees indicates stronger alignment. Hospitals may well have an interest in attracting and retaining more efficient surgeons. These factors associated with clinical performance are under a clinical department's control.

According to the strategy literature (for the earliest references, see Hamel and Prahalad, 1994), an organization's advantage requires mobilizing and "leveraging" resources and capabilities in several ways. Organizations can concentrate (or focus) resources, accumulate resources, conserve resources, and recover resources.

The resources employed by hospitals comprise the strengths and weaknesses of the organization. For example, the performance of cardiac surgeons depends on the surgeon's ability (knowledge, skills, and experience) and the capability of the hospital organization (amount invested in cardiac surgery, team skills, nursing skills, etc.) in which they perform the cardiac surgery. Though hospitals have unique combinations of clinician skills, team skills, and delivery capabilities, some perform better than others. There is some evidence that clinical efficiency may be subject to investing in organizational capabilities that arise from operational strategies such as developing open heart surgery as a specialty, years of experience, training, and, ultimately, acquiring a cadre of "efficient" surgeons. At a minimum, these findings support including some measures of "superior" organizational capabilities, strategic focus or product specialization, continuing education, and experience in future work.

Limitations

This retrospective study was intended to illustrate how to understand variations in performance and the importance of studying performance at the level of the individual physician. We have observed cardiovascular outcomes in one state over a decade to demonstrate how to measure and evaluate cardiac surgeon performance. The age of the data would be a severe limitation if our purpose were not a demonstration of a methodology. This retrospective study was intended to illustrate how to understand variations in performance and the importance of studying performance at the level of the individual physician.

Several limitations of this analysis are noted here. First, the study was conducted in only one U.S. state, Pennsylvania, employing a sampling rule that only selected physicians with at least 30 CABG patients each year. It remains to be seen whether adding physicians with smaller caseloads (less than 30) to the study would substantially change the results. The research results for CABGs may not apply to other surgical or non-surgical admissions.

Another weakness of this study is "site-specific" biases—some factors present in this study may not exist in identical intensity or magnitude elsewhere. The technical efficiency measures derived in this study are relative to the best practices observed among the physicians and hospitals for cardiac surgery in Pennsylvania, and not to the best clinical practices that might have been observed if the sample included physicians from hospitals throughout the U.S.

Finally, this study's results depend on the choice of inputs and outputs, and how DEA measures efficiency. A thorny issue is the comparability of DEA scores obtained from different DEA's comparison groups. Is a physician from comparison group 1 with a score of 90% more efficient than a physician from group 2 with a score of 80%? The DEA scores are based on actual achieved performance by members within a group studied, but the scores are relative to the sample under evaluation. There are also issues around estimating savings using DEA. Only further research with current data will determine whether or not DEA is producing clinically meaningful information.

Challenges

Health policymakers in countries around the globe are searching for ways to improve the performance of health systems. The drive for performance has spawned policies and innovations such as high-value health care, accountable care organizations, world-class commissioning, value-based purchasing, and pay-for-performance. The idea of holding physicians, nurses, allied health professionals, and clinical managers accountable for clinical outcomes, patient safety, patient experience, patient satisfaction, and clinical efficiency is what underlies all of these programs.

Why do policy makers want to create more accountable health care organizations? First, there is growing evidence that quality and efficiency vary widely among clinicians and allied health professionals. Second, while demands for safe and acceptable quality and better access to health care once justified investment and cost escalation, current budget restrictions refocus attention on clinical efficiency and cost-effective care goals.

Improving the quality and efficiency of care delivered requires measuring and comparing clinical performance and operational improvement of the clinics delivering services. Much attention centers on managerial control systems that provide information to help clinical departments monitor, profile, and compare clinical care and treatment, relying on audits and reviews to ensure that health care is neither over-utilized nor under-utilized. The question is—are there better ways to measure and evaluate clinical practice patterns?

As a tool to establish accountable care, DEA finds the top performers and estimates a relative performance score for the rest of the units. More than measuring and evaluating relative performance, the methodology also identifies which top performers offer the most realistic benchmarks for a unit. The method estimates what needs to change to improve performance by pinpointing the clinical inputs being used excessively or projecting how many more patients could be treated if the clinicians operated as well as the top performers. Consequently, the information provides benchmarks to guide performance conversations.

Understanding physician decision-making and the resulting variations in the production of hospital clinical services is important for improving hospital performance and combating rising hospital costs. This analysis offers one management science technique that might illuminate the healthcare crisis. While this early work has been promising, physicians at most hospitals will probably be reluctant to be studied at such a close level because it threatens their autonomy. Researchers interested in physician behavior will find many physicians willing to endorse research efforts on clinical outcomes, patient compliance, and medical efficacy, but few will tolerate studies of their decision-making efficiency. Therefore, researchers and management scientists must help managers use efficiency studies to learn about these very complex clinical production characteristics and not to reward or punish physicians, until we are certain what we are really finding.

This paper has identified a novel way to analyze clinical performance and offers several significant findings about potential connections between resource utilization and organizational capabilities. The resources employed by hospitals comprise the strengths and weaknesses of the organization. For example, cardiac surgeons's performance depends on their ability (knowledge, skills, and experience) and the capability of the hospital organization (such as the amount invested in cardiac surgery, team skills, and nursing skills) in which they perform cardiac surgery. Though hospitals have unique combinations of clinician skills, team skills, and delivery capabilities, some perform better than others.

We offer some evidence that clinical efficiency may be subject to investing in organizational capabilities that arise from operational strategies such as developing open heart surgery as a specialty, years of experience, training, and, ultimately, acquiring a cadre of "efficient" surgeons. At a minimum, these findings support including measures of "superior" organizational capabilities, strategic focus or product specialization, continuing education, and experience in future work.

ACKNOWLEDGMENT

The authors would like to acknowledge the assistance of Seyed Moaven Razavi, Ph.D., and Meghan Connor in completing this research work.

COMPETING INTERESTS

The authors have no competing interests to declare for this manuscript.

FUNDING

This research work was supported by a grant from the Commonwealth Fund (Grant # 20070617).

REFERENCES

Agency for Healthcare Research and Quality. (2021). *Overview of operating room procedures during inpatient stays in U.S. hospitals, 2018*. https://www.hcup-us.ahrq.gov/reports/statbriefs/sb281-Operating-Room-Procedures-During-Hospitalization-2018.jsp

American Heart Association. (2017). *Cardiovascular disease: A costly burden for America – Projections through 2035*. https://www.heart.org/-/media/Files/About-Us/Policy-Research/Fact-Sheets/Public-Health-Advocacy-and-Research/CVD-A-Costly-Burden-for-America-Projections-Through-2035.pdf

American Hospital Association. (1996). *AHA Annual survey database (1995)*. https://www.aha.org/

American Medical Association. (1996). *Directory of physicians in the United States* (35th ed.).

American Medical Association. (2004). *Directory of physicians in the United States* (39th ed.).

Anderson, G. F., Reinhardt, U. E., Hussey, P. S., & Petrosyan, V. (2003). It's the prices, stupid: Why the United States is so different from other countries. *Health Affairs (Project Hope)*, *22*(3), 89–105. doi:10.1377/hlthaff.22.3.89 PMID:12757275

Andriole, D. A., Evans, S. R., Foy, H. M., Atnip, R. G., & Mancino, A. T. (1998). Can a one-day workshop alter surgeons' teaching practices? *American Journal of Surgery*, *175*(6), 518–520. doi:10.1016/S0002-9610(98)00077-4 PMID:9645786

Auerbach, A. D., Hilton, J. F., Maselli, J., Pekow, P. S., Rothberg, M. B., & Lindenauer, P. K. (2010). Case volume, quality of care, and care efficiency in coronary artery bypass surgery. *Archives of Internal Medicine*, *170*(14), 1202–1208. doi:10.1001/archinternmed.2010.237 PMID:20660837

Azam, M., Qureshi, M. R., & Talib, F. (2017). Quality evaluation of health care establishment utilizing Fuzzy AHP. *International Journal of Service Science, Management, Engineering, and Technology*, *8*(4), 83–120. doi:10.4018/IJSSMET.2017100105

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, *30*(9), 1078–1092. doi:10.1287/mnsc.30.9.1078

Barhoun, R., Ed-daibouni, M., & Namir, A. (2019). An extended attribute-based access control (ABAC) model for distributed collaborative healthcare system. *International Journal of Service Science, Management, Engineering, and Technology*, *10*(4), 81–94. doi:10.4018/IJSSMET.2019100105

Barros, O. (2022). A management and enterprise architecture framework for comprehensive structure design of complex services. *International Journal of Service Science, Management, Engineering, and Technology*, *13*(1), 1–24. doi:10.4018/IJSSMET.315583

Barros, O., & Aguilera, I. (2022). Improving the efficiency of a public health service by designing its innovation management structure based on data envelopment analysis. *International Journal of Service Science, Management, Engineering, and Technology*, *13*(1), 1–17. doi:10.4018/IJSSMET.315583

Birkmeyer, J. D., Siewers, A. E., Finlayson, E. V., Stukel, T. A., Lucas, F. L., Batista, I., Welch, H. G., & Wennberg, D. E. (2002). Hospital volume and surgical mortality in the United States. *The New England Journal of Medicine*, *346*(15), 1128–1137. doi:10.1056/NEJMsa012337 PMID:11948273

Brewster, A. C., Karlin, B. G., Hyde, L. A., Jacobs, C. M., Bradbury, R. C., & Chae, Y. M. (1985). MEDISGRPS: a clinically based approach to classifying hospital patients at admission. *Inquiry: A Journal of Medical Care Organization. Provision and Financing*, *22*(4), 377–387.

Bridges, C. R., Edwards, F. H., Peterson, E. D., & Coombs, L. P. (2000). The effect of race on coronary bypass operative mortality. *Journal of the American College of Cardiology*, *36*(6), 1870–1876. doi:10.1016/S0735-1097(00)00956-6 PMID:11092658

Burns, L. R., Wholey, D. R., & Huonker, J. (1989). Physician use of hospitals: Effects of physician, patient, and hospital characteristics. *Health Services Management Research*, *2*(3), 191–203. doi:10.1177/095148488900200303 PMID:10296915

Capdeville, M., Chamogeogarkis, T., & Lee, J. H. (2001). Effect of gender on outcomes of beating heart operations. *The Annals of Thoracic Surgery*, *72*(3), S1022–S1025. doi:10.1016/S0003-4975(01)02954-X PMID:11565719

Charnes, A., Cooper, W. W., Lewin, A. Y., & Seiford, L. M. (1994). Basic DEA models. In A. Charnes, W. W. Cooper, A. Y. Lewin, & L. M. Seiford (Eds.), *Data envelopment analysis: Theory, methodology, and application*. Kluwer Academic Publishers.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, *2*(6), 429–444. doi:10.1016/0377-2217(78)90138-8

Chilingerian, J. A. (1989). Investigating non-medical factors associated with the technical efficiency of physicians in the provisions of hospital services: A pilot study. *Proceedings - Academy of Management*, *1989*(1), 85–89. doi:10.5465/ambpp.1989.4977968

Chilingerian, J. A. (1995). Evaluating physician efficiency in hospitals: A multivariate analysis of best practices. *European Journal of Operational Research*, *80*(3), 548–574. doi:10.1016/0377-2217(94)00137-2

Chilingerian, J. A., & Glavin, M. P. V. (1994). Temporary firms in community hospitals: Elements of a managerial theory of clinical efficiency. *Medical Care Review*, *51*(3), 289–335. doi:10.1177/107755879405100303 PMID:10138050

Chilingerian, J. A., & Sherman, D. (2011). Health-care applications: From hospitals to physicians, from productive efficiency to quality frontiers. In W. Cooper, L. Seiford, & J. Zhu (Eds.), *Handbook on data envelopment analysis*. Springer. doi:10.1007/978-1-4419-6151-8_16

Chilingerian, J. A., & Sherman, H. D. (1990). Managing physician efficiency and effectiveness in providing hospital services. *Health Services Management Research*, *3*(1), 3–15. doi:10.1177/095148489000300101 PMID:10104280

Cooper, W. W., Seiford, L. M., & Tone, K. (2007). *Data envelopment analysis: A comprehensive text with models, applications, references and DEA-Solver software* (2nd ed.). Springer. doi:10.1007/978-0-387-45283-8

Coughlin, L. D., & Patel, V. L. (1987). Processing of critical information by physicians and medical students. *Journal of Medical Education*, *62*(10), 818–828. doi:10.1097/00001888-198710000-00005 PMID:3656383

Cowper, P. A., DeLong, E. R., Peterson, E. D., Hannan, E. L., Ray, K. T., Racz, M., & Mark, D. B. (2002). Variability in cost of coronary bypass surgery in New York State: Potential for cost savings. *American Heart Journal*, *143*(1), 130–139. doi:10.1067/mhj.2002.119617 PMID:11773923

Cromwell, J., Mitchell, J. B., & Stason, W. B. (1990). Learning by doing in CABG surgery. *Medical Care*, *28*(1), 6–18. doi:10.1097/00005650-199001000-00003 PMID:2404169

Del Rizzo, D. F., Boyd, W. D., Novick, R. J., McKenzie, F. N., Desai, N. D., & Menkis, A. H. (1998). Safety and cost-effectiveness of MIDCABG in high-risk CABG patients. *The Annals of Thoracic Surgery*, *66*(3), 1002–1007. doi:10.1016/S0003-4975(98)00660-2 PMID:9768990

Eisenberg, J. M. (1986). *Doctors' decisions and the cost of medical care*. Health Administration Press.

Eisenberg, J. M., & Nicklin, D. (1981). Use of diagnostic services by physicians in community practice. *Medical Care*, *19*(3), 297–309. doi:10.1097/00005650-198103000-00005 PMID:7218895

Gibson, P. H., Croal, B. L., Cuthbertson, B. H., Gibson, G., Jeffrey, R. R., Buchan, K. G., El-Shafei, H., & Hillis, G. S. (2009). Socio-economic status and early outcome from coronary artery bypass grafting. *Heart (British Cardiac Society)*, *95*(10), 793–798. doi:10.1136/hrt.2008.149849 PMID:19304571

Greene, W. H. (1993). *Econometric analysis* (2nd ed.). Macmillan.

Hamel, G., & Prahalad, C. K. (1994). *Competing for the future*. Harvard Business School Press.

Hannan, E. L., Wu, C., Ryan, T. J., Bennett, E., Culliford, A. T., Gold, J. P., Hartman, A., Isom, O. W., Jones, R. H., McNeil, B., Rose, E. A., & Subramanian, V. A. (2003). Do hospitals and surgeons with higher coronary artery bypass graft surgery volumes still have lower risk-adjusted mortality rates? *Circulation*, *108*(7), 795–801. doi:10.1161/01.CIR.0000084551.52010.3B PMID:12885743

HCIA (1999). *The hundred top cardiovascular hospitals*.

Hillman, A. L., Pauly, M. V., & Kerstein, J. J. (1989). How do financial incentives affect physicians' clinical decisions and the financial performance of health maintenance organizations? *The New England Journal of Medicine*, *321*(2), 86–92. doi:10.1056/NEJM198907133210205 PMID:2733758

Ho, V., & Petersen, L. A. (2007). Estimating cost savings from regionalizing cardiac procedures using hospital discharge data. *Cost Effectiveness and Resource Allocation*, *5*(7), 7. Advance online publication. doi:10.1186/1478-7547-5-7 PMID:17603890

Judge, G., Hill, C., Griffiths, W., Lutkepohl, H., & Lee, T. (1985). *The theory and practice of econometrics* (2nd ed.). Wiley.

Kim, C., Diez Roux, A. V., Hofer, T. P., Nallamothu, B. K., Bernstein, S. J., & Rogers, M. A. (2007). Area socioeconomic status and mortality after coronary artery bypass graft surgery: The role of hospital volume. *American Heart Journal*, *154*(2), 385–390. doi:10.1016/j.ahj.2007.04.052 PMID:17643593

Landon, B., Iezzoni, L. I., Ash, A. S., Shwartz, M., Daley, J., Hughes, J. S., & Mackiernan, Y. D. (1996). Judging hospitals by severity-adjusted mortality rates: The case of CABG surgery. *Inquiry: A Journal of Medical Care Organization. Provision and Financing*, *33*(2), 155–166.

Lin, H. C., Xirasagar, S., Chen, C. H., Lin, C. C., & Lee, H. C. (2007). Association between physician volume and hospitalization costs for patients with stroke in Taiwan: A nationwide population-based study. *Stroke*, *38*(5), 1565–1569. doi:10.1161/STROKEAHA.106.474841 PMID:17395867

Luke, R. D., & Thomson, M. A. (1980). Utilization of within-hospital services: A study of the effects of two forms of group practice. *Medical Care*, *18*(2), 219–227. doi:10.1097/00005650-198002000-00008 PMID:7206842

Maddala, G. S. (1983). *Limited dependent and qualitative variables in econometrics*. Cambridge University Press. doi:10.1017/CBO9780511810176

Mandal, P. C. (2022). Service excellence in marketing for firms: Strategies and initiatives. *International Journal of Service Science, Management, Engineering, and Technology*, *13*(1), 1–22. doi:10.4018/IJSSMET.290331

Mensah, G. A., & Brown, D. W. (2007). An overview of cardiovascular disease burden in the United States. *Health Affairs (Project Hope)*, *26*(1), 38–48. doi:10.1377/hlthaff.26.1.38 PMID:17211012

Nagle, P. C., & Smith, A. W. (2004). Review of recent U.S. cost estimates of revascularization. *The American Journal of Managed Care*, *10*(11, Suppl), S370–S376. PMID:15603246

Office of the Actuary, Centers for Medicare and Medicaid Services. (2007). *Health expenditures by state of provider: State-specific tables, 1980-2004*. https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/nationalhealthexpenddata/nationalhealthaccountsstatehealthaccountsprovider

Papadimos, T. J., Habib, R. H., Zacharias, A., Schwann, T. A., Riordan, C. J., Durham, S. J., & Shah, A. (2005). Early efficacy of CABG care delivery in a low procedure-volume community hospital: Operative and midterm results. *BMC Surgery*, *5*(1), 10. doi:10.1186/1471-2482-5-10 PMID:15865623

Pennsylvania Health Care Cost Containment Council (1998). *Coronary artery bypass graft surgery —1994-95 data: Research methods and results*. https://www.phc4.org/wp-content/uploads/cabg1994_95researchmethods.pdf

Pennsylvania Health Care Cost Containment Council. (1998). *About the Pennsylvania Health Care Cost Containment Council*. www.phc4.org/aboutthe.htm

Pennsylvania Health Care Cost Containment Council. (2005). *Technical notes: Pennsylvania's guide to coronary artery bypass graft surgery, calendar year 2003*. https://www.phc4.org/wp-content/uploads/cabg2003technotes.pdf

Pennsylvania Health Care Cost Containment Council. (2006). *Technical notes: Pennsylvania's guide to coronary artery bypass graft surgery, calendar year 2004*. https://www.phc4.org/wp-content/uploads/cabg2004technotes.pdf

Rankin, S. H. (1990). Differences in recovery from cardiac surgery: A profile of male and female patients. *Heart & Lung: The Journal of Critical Care*, *19*(5 Pt 1), 481–485. PMID:2211155

Rathore, S. S., Epstein, A. J., Volpp, K. G., & Krumholz, H. M. (2004). Hospital coronary artery bypass graft surgery volume and patient mortality, 1998-2000. *Annals of Surgery*, *239*(1), 110–117. doi:10.1097/01.sla.0000103066.22732.b8 PMID:14685108

Regitz-Zagrosek, V., Lehmkuhl, E., Hocher, B., Goesmann, D., Lehmkuhl, H. B., Hausmann, H., & Hetzer, R. (2004). Gender as a risk factor in young, not in old, women undergoing coronary artery bypass grafting. *Journal of the American College of Cardiology*, *44*(12), 2413–2414. doi:10.1016/j.jacc.2004.09.031 PMID:15607409

Rumsfeld, J. S., Plomondon, M. E., Peterson, E. D., Shlipak, M. G., Maynard, C., Grunwald, G. K., Grover, F. L., & Shroyer, A. L. (2002). The impact of ethnicity on outcomes following coronary artery bypass graft surgery in the Veterans Health Administration. *Journal of the American College of Cardiology*, *40*(10), 1786–1793. doi:10.1016/S0735-1097(02)02485-3 PMID:12446062

Sisodia, S., & Agrawal, N. (2019). Examining employability skills for healthcare services in India: A descriptive literature review. *International Journal of Service Science, Management, Engineering, and Technology*, *10*(3), 63–79. doi:10.4018/IJSSMET.2019070105

Srivastava, S. (2021). Exploring enablers of value co-creation in the healthcare service delivery: An internal customer perspective. *International Journal of Service Science, Management, Engineering, and Technology*, *12*(1), 111–126. doi:10.4018/IJSSMET.2021010107

Steen, P. M., Brewster, A. C., Bradbury, R. C., Estabrook, E., & Young, J. A. (1993). Predicted probabilities of hospital death as a measure of admission severity of illness. *Inquiry: A Journal of Medical Care Organization. Provision and Financing*, *30*(2), 128–141.

Thurow, L. C. (1985). Medicine versus economics. *The New England Journal of Medicine*, *313*(10), 611–614. doi:10.1056/NEJM198509053131005 PMID:3927162

Vaccarino, V., Lin, Z. Q., Kasl, S. V., Mattera, J. A., Roumanis, S. A., Abramson, J. L., & Krumholz, H. M. (2003). Gender differences in recovery after coronary artery bypass surgery. *Journal of the American College of Cardiology*, *41*(2), 307–314. doi:10.1016/S0735-1097(02)02698-0 PMID:12535827

Weiss, A. J., & Elixhauser, A. (2014). Trends in Operating Room Procedures in U.S. Hospitals, 2001–2011. In *Healthcare Cost and Utilization Project (HCUP) Statistical Briefs*. Agency for Healthcare Research and Quality.

Wilson, C. T., Fisher, E. S., Welch, H. G., Siewers, A. E., & Lucas, F. L. (2007). U.S. trends in CABG hospital volume: The effect of adding cardiac surgery programs. *Health Affairs (Project Hope)*, *26*(1), 162–168. doi:10.1377/hlthaff.26.1.162 PMID:17211025

World Health Organization. (2021). *Cardiovascular diseases fact sheet*. https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

Zhu, J. (2000). Further discussion on linear production functions and DEA. *European Journal of Operational Research*, *127*(3), 611–618. doi:10.1016/S0377-2217(99)00344-6

ENDNOTES

1. Healthcare performance often involves mixing non-commensurate information, or physical quantities with different units. For example, clinical services can be measured in monetary units, full time equivalents staff, lengths-of-stay, nursing units, minutes in the operating room, number of lab tests or junior registrar visits. Although these are physical quantities with different units, DEA can combine them in a consistent way.
2. It is also possible that physicians who admit too many patients may spend less time with each patient and physicians could become more inefficient with very high volumes.
3. SALMDPCT and HSPTL are from the 2004 AHA Annual Survey of Hospitals.
4. DRG 106 is for CABG with catheterization and DRG 107 is for CABG without catheterization. By 2003, however, DRG 107 was used for CABG with catheterization and DRG 109 for CABG without catheterization.
5. Since 1994 and 1995 Medicare Cost Report data were unavailable, the 1996 cost-to-charge ratios for each Pennsylvania hospital were used for these years.
6. The PAHC4 data base was the source of all control variables.
7. The sources for the explanatory variables were the PAHC4 data base (OPHTOT and HOSPCT), the AMA data base (YRSPRA and CME), and the AHA data base (TEACH). The HMIX variable was constructed from PAHC4 and AHA data.

8 The raw charges were adjusted by applying each hospital's cost-to-charge ratio per CMS cost reports, and then inflated to 2004 dollars.

9 The raw charges for 2003-2004 were adjusted by applying each hospital's cost-to-charge ratio per CMS cost reports, and then the 2003 charges were inflated to 2004 dollars.

10 Adjusted charge data for 1994, 1995, and 2003 were inflated to 2004 dollars before computing totals per surgeon.

11 Transformation: Inefficiency score = $(1/\theta) - 1$, where θ = DEA efficiency score.

12 Transformation: Inefficiency score = $(1/\theta) - 1$, where θ = DEA efficiency score.

13 Mortality rates were not adjusted for patient severity. CABG inpatient mortality rates for each surgeon were obtained by dividing the number of CABG patients with discharge status of expired by the total number of CABG patients (any discharge status). The 30-day mortality rate for each surgeon was determined by dividing the number of patients who died within 30 days of discharge by the number of patients discharged home alive. The 30-day readmission rate for each surgeon was determined by dividing the number of patients who were readmitted within 30 days of discharge by the number of patients discharged home alive.

Jon A. Chilingerian, Ph.D., is a tenured Professor of Management at Brandeis University and an Adjunct Professor of Public Health and Community Medicine at Tufts School of Medicine. He is the founding director of the Heller School's EMBA physician Program and the Brandeis-Tufts School of Medicine MD-MBA. He received his Ph.D. in Management from MIT's Sloan School of Management. His research focuses on managing healthcare organizations, ranging from studies of executive leadership and management of professionals to the measurement of performance (i.e., productive efficiency, quality, etc.), identification of physician best practices, and the analysis of effective operating strategies. An expert on healthcare performance studies using mathematical programming and frontier analysis, Dr. Chilingerian is the lead editor of a forthcoming Springer Nature Book on "The New Science of Medicine and Management." In his former life, he held various managerial and leadership positions, including Assistant Health Commissioner for the City of Boston. Editor's note: Unfortunately, Jon Chilingerian passed away a few months after submitting this manuscript. Further information on Jon and his work can be accessed at https://heller.brandeis.edu/news/items/releases/2023/remembering-jon-chilingerian.html and at https://mba.insead.edu/centres/healthcare-management.

Mitchell P.V. Glavin is Chair of the Department of Healthcare Management in the Meehan School of Business, Stonehill College. He also serves as Lead for the Department's Healthcare Supply Chain Management Initiative. He received his Ph.D. in health policy and management from the Heller Graduate School, Brandeis University with a Health Care Financing Administration Dissertation Fellowship for his doctoral research. Dr. Glavin earned master's degrees from the University of Chicago and the London School of Economics and Political Science. He has published on a broad range of topics including the care production process in hospitals, organizational strategies to reduce medical errors, the withdrawal of HMOs from participation in Medicare contracting, and employee engagement and productivity. His current research focuses on assessing performance and identifying the best practices and adaptations of top-performing organizations. He has also published on incorporating problem-based learning into undergraduate course design.