Health Care Manage Sci (2008) 11:67–77
DOI 10.1007/s10729-007-9025-8

Is more better? An analysis of hospital outcomes and efficiency with a DEA model of output congestion

Jan P. Clement · Vivian G. Valdmanis · Gloria J. Bazzoli · Mei Zhao · Askar Chukmaitov

Received: 9 March 2007 / Accepted: 31 August 2007 / Published online: 19 September 2007

Abstract This paper applies a new methodology to the study of hospital efficiency and quality of care. Using a data set of hospitals from several states, we jointly evaluate desirable hospital patient care output (e.g., patient stays) and the simultaneous undesirable output (e.g., risk-adjusted patient mortality) that occurs. With a DEA based approach under two different sets of assumptions, we are able to include multiple quality indicators as outputs. The results show that lower technical efficiency is associated with poorer risk-adjusted quality outcomes in the study hospitals. They are consistent with other studies linking poor quality outcomes to higher cost.

Keywords Hospital efficiency · DEA · Patient outcomes

J. P. Clement (✉) · G. J. Bazzoli
Department of Health Administration,
Virginia Commonwealth University,
P.O. Box 980203, Richmond, VA 23298-0203, USA
e-mail: jclement@vcu.edu

G. J. Bazzoli
e-mail: gbazzoli@vcu.edu

V. G. Valdmanis
Health Policy Program,
University of the Health Sciences in Philadelphia,
Philadelphia, PA 19104, USA
e-mail: v.valdma@usip.edu

M. Zhao
Health Administration Programs, College of Health,
University of North Florida,
4567 St. Johns Bluff Road, South,
Jacksonville, FL 32224-2673, USA
e-mail: mzhao@unf.edu

A. Chukmaitov
Family Medicine and Rural Health, College of Medicine,
Florida State University,
1115 West Call Street, Room 3200-C,
Tallahassee, FL 32306-4300, USA
e-mail: askar.chukmaitov@med.fsu.edu

1 Introduction

Over the last two decades, as third party payers experienced mounting costs, they focused largely on enhancing hospital efficiency through constraining payments for hospital services. The mid to late 1990s were particularly challenging for acute care hospitals in the USA. Private third party payers used discounting, utilization management and risk shifting to control costs. Some states enacted measures to limit their Medicaid costs and Congress passed the Balanced Budget Act (BBA) in 1997. The BBA was intended to cut total Medicare payments, typically the largest revenue source for most hospitals, by $112 million for 1998 through 2002 [1].

Facing the resulting revenue pressure, hospitals had to maintain financial viability by finding ways to cut costs. Indeed, researchers report that the Medicare payment constraints were associated with slower hospital staffing growth among not-for-profit hospitals [1] and small but significant declines in nurse staffing levels at non-safety net hospitals [2]. Another study found that financial pressure during the late 1990s led to less investment in net plant assets [3]. In contrast, from 1995 through 2000, total outpatient visits and hospital admissions increased steadily and hospital days increased in 1999 and 2000 [4]. Although hospitals may have become more efficient, it is also possible that quality of care was harmed as they had to conserve on staffing and other resources and used these resources more intensely to serve growing demand.

This paper applies a new methodology to the study of efficiency and quality of care. Using a data set of hospitals from several states, we jointly evaluate desirable hospital patient care output (e.g., patient stays) and the simultaneous undesirable output (e.g., risk-adjusted patient mortality) that occurs. More specifically, we assess the productive efficiency of hospitals using multiple inputs and outputs with data envelopment analysis (DEA) under two different assumptions. Under the strong disposability of outputs (SDO) assumption, expansion of all patient care outputs, such as admissions and outpatient visits, is desirable. In contrast, the assumption of weak disposability of outputs (WDO) treats expansion of some outputs as undesirable. A measure of the difference in hospital efficiency between the two different assumptions, referred to as congestion, is shown by the ratio of the efficiency scores.

This approach has been used in evaluating the economic bads that detract from the social welfare function of producing goods and services in the economy [5]. A common example is that of pollution. As more of a particular good is produced, such as automobiles, factories produce more pollutants that are emitted into the environment. The same can be said about patient care in hospitals. As more patients are treated with the same production process, holding hospital inputs constant, the likelihood of a poor outcome can increase, detracting from the overall benefits of hospitalization.

We assess the relationship between hospital technical efficiency and quality of patient outcomes using data from 2000. This year is of interest because hospitals had time to adjust inputs in response to the reimbursement constraints of the late 1990s and had also experienced growth in admissions, inpatient days, and outpatient visits [4]. Our findings show that poorer patient outcomes are associated with lower technical efficiency. Most hospitals could improve technical efficiency along with patient care outcomes. The study also examines whether organizational and environmental factors such as a hospital's payer mix, patient characteristics, resources, ownership, teaching status and location are systematically associated with the potential efficiency improvements.

2 Background

Some previous research has found that hospitals utilizing greater amounts of inputs for a given caseload have better patient outcomes. Because nursing care is such a large component of hospital care and there is a lingering nursing shortage in the USA, researchers have been particularly interested in the role of nurse staffing in hospital quality of care. Most of these studies support the notion that higher nurse to patient staffing levels and higher proportions of registered nurses (RNs) in the nurse staffing mix are associated with fewer patient adverse events and lower mortality rates [6–18].

Other studies have found a positive association between cost or service intensity and more desirable patient outcomes. Picone et al. [19] found that higher service intensity lowers mortality rates for selected diagnoses. Burstin et al. [20] found that higher hospital operating cost per discharge was associated with a lower likelihood of negligent medical injury. Indeed, some researchers have assumed that higher service intensity equates with higher quality of care in their study designs [21, 22].

In contrast, other empirical research points to the opposite relationship. That is, more intense or costly care may not be better care. It may be inefficient. Rapoport et al. [23] studied the relationship between resource use and the difference between observed and predicted hospital survival based upon severity of illness at ICU admission. They measured resource use with hospital days with ICU days weighted more highly. Their findings show no tradeoff between quality and efficiency. However, the data requirements limited the number of study hospitals to 25.

Bradbury et al. [24] found positive and significant relationships between hospital expenditures and certain patient morbidity and mortality measures after adjusting for patient risk factors. Although Shultz et al. [15] found that hospitals with greater RN availability had lower mortality rates from acute myocardial infarction (AMI), hospitals with lower operating expenses per patient day also had lower AMI mortality rates. Carey and Burgess [25] estimated a cost function for 137 Veteran's Administration hospitals that included three quality measures, mortality, readmission and follow-up outpatient indices, as regressors. An increase in these adverse events was associated with higher costs implying that poor quality is more costly (less efficient). However, further analyses led the authors to conclude that their quality measures, which they were unable to adjust adequately for severity of illness, may, in part, reflect patient severity as well as quality.

More recently, Deily and McKay [26] hypothesized that higher costs may be positively associated with poor quality because total cost has two components, costs associated with the best use of resources and those associated with waste or inefficiency. They included inefficiency scores in a regression model of the in-hospital mortality rate controlling for predicted risk-adjusted mortality. With data from Florida hospitals over the period 1999–2001, they found that inefficiency was significantly positively associated with the mortality rate. However, as with other studies, these findings may be influenced by omitted risk adjustments. They are also limited by the use of a single outcome measure.

Still other research suggests that the relationship between hospital resource inputs and quality of patient care may be complex. McCloskey [12], Blegen et al. [7], and Blegen and

Vaughn [8] found that as the RN proportion of nursing staff increased beyond some point, the incidence of adverse outcomes began to increase even after controlling for patient acuity. RN staffing proportions that are too high may be inefficient. Alternatively, this result may reflect inadequate controls for patient acuity using standard severity adjustments. Fleming [27] found that cost and quality increased together in low cost hospitals but there was a negative cost-quality relationship for mid to high cost hospitals.

In summary, evidence regarding whether and how hospital efficiency and the quality of patient care are related is mixed. As noted by at least some of the previously cited authors, the results may also be influenced by the ability of patient risk adjusters to reflect patient severity of illness accurately. Another limitation of the studies is that they rely upon the typical assumption that hospitals minimize costs for given output levels, which may or may not be applicable. Picone et al. [19] noted that slack resources and capacity may be inputs to quality. Further, Hoerger [28] suggested that not-for-profit hospitals eschew profits above a certain target level by expanding the quantity and/or quality of services they produce.

Thus, lacking a consensus regarding the relationship between efficiency and quality, we contribute to existing literature with an alternative methodological approach based upon data envelopment analysis under alternative assumptions. The methodological flexibility of DEA allows simultaneous analysis of several quality indicators, which will provide a wider view of the hospital than those studies that just focused on a single quality measure or a specific diagnostic category. In addition, DEA optimizes the performance measure for each hospital in contrast to regression models that optimize a single model across all observations.

3 Data and measures

3.1 Data

Study data are drawn from both patient-level discharge and hospital-level organizational data. Hospital discharge data from ten states in 2000 are used to construct hospital level quality indicators. These data were obtained from the Agency for Healthcare Research and Quality (AHRQ) Healthcare Cost and Utilization Project State Inpatient Database (HCUP SID). The HCUP SID is the largest collection of all-payer, uniform, state-based, inpatient administrative data. Administrative data are routinely collected for each hospital discharge and are not specifically designed for quality assessment. Among the variables included in the data set are patient demographics including age and gender, expected payment source and patient clinical data including principal and secondary diagnoses and procedures, and length of stay. The ten states—Arizona, California, Colorado, Florida, Iowa, Massachusetts, New Jersey, New York, Washington, and Wisconsin—were chosen based upon three criteria: (1) the states had mandatory, rather than voluntary, hospital participation in data collection; (2) hospital specific identifiers are included, which allowed merging with other sources of hospital data; and (3) the state had a history of collecting the discharge database, which implies that difficulties and errors that may arise in early efforts have likely been corrected. The hospitals are geographically dispersed, covering seven of the nine census divisions in the USA. Because we study hospitals in 2000, the legislation in California requiring a minimum nurse staffing ratio at hospitals, which became effective in 2004, does not affect our results [29].

The SID data are merged with data for hospital inputs and outputs from the 2000 American Hospital Association (AHA) Annual Survey. Although the AHA data are frequently used for a variety of cost, quality and efficiency studies, a drawback is that the data are reported only for the hospital level; they are not available for various product lines or hospital departments.

3.2 DEA input measures

For the DEA efficiency estimates, we use the inputs shown in Table 1. The labor inputs are full time equivalent (FTE) registered nurses, FTE licensed practical nurses, and other FTEs. Capital input is captured by licensed staffed beds. The AHA Annual Survey does not provide further information on the "other FTE" category. However, since the outcomes we study can be considered nursing sensitive, the data source does provide information on the most critical staffing levels. Unfortunately, capital input is also not further disaggregated. However, hospital bed size is typically correlated with other capital inputs such as imaging technologies. An alternative source of capital input data, the dollar value of net plant, property and equipment obtained from the Centers for Medicare and Medicaid Services cost report data, proved not to be useful because a of a high number of missing observations.

3.3 DEA output measures

The output measures are also shown in Table 1. They include the outputs commonly used in frontier estimation: number of births, outpatient surgeries, emergency room visits, outpatient visits, and case mix adjusted admissions. Case mix adjusted admissions are calculated as the number of admissions multiplied by the hospital's average case mix. For example, if a hospital has 1,000 admissions and its average case mix index is 1.2, we credit this hospital as

Table 1 Descriptive statistics for data envelopment efficiency model variables (n=667)

Variable	Mean	SD	Minimum	Maximum
Inputs				
FTE registered nurses	313.34	254.68	35	1,683
FTE licensed practical nurses	30.89	28.44	0	262
FTE other	899.10	779.13	94	7,621
Staffed beds	246.7	153.8	24	1,049
Outputs				
Births	1,464.5	1,359.2	0	9,892
Outpatient surgeries	8,407.7	6,075.9	0	77,716
Emergency room visits	34,103.2	22,913.0	0	227,457
Outpatient visits	135,558.3	153,434.9	4,892	1,548,092
Case mix adjusted admissions	13,097.1	9,865.6	924.93	84,644.79
Risk-Adjusted Mortality Rate (%)				
Acute myocardial infarction	11.80	5.63	0	34.89
Congestive heart failure	5.01	2.11	0	12.50
Stroke	13.14	4.69	0	28.87
Gastrointestinal hemorrhage	3.67	1.89	0	10.06
Pneumonia	9.45	3.44	0	18.85
Number of Patients at Risk at Hospital for Mortality				
Acute myocardial Infarction	176.7	160.7	30	987
Congestive heart failure	353.5	221.2	39	1,864
Stroke	184.2	116.5	30	770
Gastrointestinal hemorrhage	172.6	97.5	36	780
Pneumonia	403.9	231.9	60	2,239

treating 1,200 patients (1,000×1.2). If another hospital also admits 1,000 patients but only has an average case mix index of 1.0, this hospital is credited with admitting 1,000 patients (1,000×1.0). Thus, hospitals treating a more serious case mix of patients are not penalized.

It should be noted that some hospitals do not hire some inputs or provide some outputs. Since these decisions are made by hospital managers, they are considered pertinent in determining the product mix of our sample. However, since variable returns to scale is used, they will be compared with their peer groupings which are similar in output mixes. Not providing ER services will not penalize a hospital. It only means the hospital will be compared to other hospitals not providing ER services or some other similar mix.

In addition, we include as outputs a series of indicators for the quality of patient outcomes, which were risk-adjusted for patient characteristics, in both the SDO and WDO models. Donabedian [30] recognized that quality is a multidimensional concept and identified structural, process and outcome dimensions for health care quality. Certain structures, such as up-to-date facilities and highly trained staffing, can provide a foundation for good quality patient care. Ideally, these structural inputs should be coupled with high quality care processes to achieve good patient outcomes. In reality, though, patient outcome measures like the ones we study are affected not only by the quality of relevant structures and care processes, but also by patient severity of illness. For example, it is possible to achieve good patient outcomes, such as low mortality rates or adverse events, even if good structures or processes are not in place because of differences in patient severity of illness. Thus, it is necessary to risk-adjust our measures of patient outcome quality.

More specifically, we use the SID data to construct risk-adjusted inpatient quality indicators (IQIs) for each hospital using methods and software developed by the Agency for Health Care Research and Quality (AHRQ) [31]. There are a total of fifteen risk-adjusted mortality IQIs. However, many hospitals did not report data for a number of the IQIs or, as is explained later, reported too few patients at risk to construct an IQI. To choose IQIs relevant to a large number of hospitals and patients, we included only IQIs for which 75% or more of the hospitals in our data set reported having patients at risk. These IQIs are for acute myocardial infarction (AMI), congestive heart failure (CHF), stroke, gastrointestinal hemorrhage, and pneumonia.

The IQIs are risk-adjusted to account for differences in patient mix across hospitals using a linear multivariate regression model developed by AHRQ and included within their downloadable software. The AHRQ risk adjustment procedure involves estimating hospital fixed effects models with controls for patient age category, gender, the interaction of age category and gender, and all patient diagnosis related group (APR-DRG) classifications. The APR-DRGs expand the basic classification the Centers for Medicare and Medicaid Services uses to classify Medicare patient discharges according to diagnosis and resource intensity by adding four disease specific levels of risk of mortality, with 1 representing the lowest and 4, the highest risk [32]. As a result of the adjustment process, the risk-adjusted rates, which are the coefficients on the hospital fixed effects for each indicator, are the estimated performance of providers if those providers had an average case mix. This average case mix, which is included in the software, is estimated using data from 28 states in the SID databases. If the observed rate is close to zero and the provider has a more severe than average case mix, it is possible the estimated risk-adjusted rates may be negative. In this case, the adjusted rate is set to zero. The AHRQ developed this methodology to allow direct comparisons across hospitals. Thus, our research improves upon previous studies by using multiple quality indicators for common patient conditions that are risk-adjusted with the most current methodology. However, we acknowledge it is still possible that there are unmeasured patient risk factors that affect the quality indicators.

Each risk-adjusted mortality indicator is presented as a percentage for the hospital. We also include the number of patients at risk for mortality for each of the five IQIs to control for the possibility that hospital care improves with the increased experience resulting from higher volume. Table 1 summarizes all of the input and output variables for the DEA analyses.

3.4 Study hospitals

To remove the potential influence of extreme values and transfers on the risk-adjusted IQIs, we eliminated all patient observations involving transfers into or out of the hospital before calculating the risk-adjusted IQIs. In addition, we removed hospitals from the analysis if they had fewer than 30 patients at risk for an IQI. Hospitals with very few patients at risk for an IQI may have IQI rates that appear very high. We also deleted outliers for the risk-adjusted measure defined as the upper 1% of the risk-adjusted IQI rates. The distributions for the IQIs were skewed with long flat upper tails containing very high risk-adjusted mortality rates. The lowest value for each IQI is 0%.

We began with 1,497 hospitals classified as acute care general hospitals that reported at least some IQI data in 2000. Given the multiple output nature of the analysis, any hospital that did not report a rate for each of the five IQIs could not be included in the analysis (435 hospitals). We also eliminated hospitals with a high proportion of non-acute care beds, occupancy rates higher than 99% or lower than 10%, high proportions of licensed practical nurses, and an average case mix adjusted length of stay longer than 16 days (344 hospitals). These characteristics are more common among hospitals that focus on post-acute skilled nursing or rehabilitation rather than general acute care. Finally, we eliminated hospitals with missing data for outpatient and emergency room visits (51 hospitals).

The 667 hospitals retained are significantly larger than the omitted hospitals with an average of 246.7 compared to 141 beds. Correspondingly, the average number of admissions is 11,486 for the included versus 5,217 for the excluded hospitals. As a result of their size, the study hospitals have more employees, emergency room visits, and outpatient visits than the excluded hospitals. Study hospitals are also more likely to be involved in teaching activities but less likely to be public hospitals. These differences are probably due to the requirement that there are at least 30 patients at risk for each IQI, which leaves larger hospitals located in urban areas. The study hospitals are also larger than the average acute care hospital in 2000 that had 167.7 beds and 6,732 admissions [4, 33, 34].

The descriptive statistics in Table 1 show a range of size and performance among the 667 hospitals. Staffed bed size ranges from 24 to 1,049 with a standard deviation of 153.8 and full time equivalent (FTE) registered nurses (RNs) ranges from 35 to 1,683 with a standard deviation of 254.68. Some hospitals employ no licensed practical nurses (LPNs). Mean FTE LPNs is 30.89 with a standard deviation of 28.44. Similarly, hospitals vary across outputs. For example, outpatient surgeries range from 0 to 77,716 with a mean of 8,407.7. The mean number of case mix adjusted admissions is 13,097.1 with a standard deviation of 9,865.6. At the bottom of Table 1, the number of patients at risk of each inpatient quality indicator (IQI) is shown. As would be expected from the variation in the size of the other outputs, hospitals treat different numbers of these patients. The number of patients at risk for each inpatient quality indicator (IQI), has a minimum of 30, as noted previously. The highest number of patients at risk is for pneumonia, averaging 403.9 in the study hospitals, but the range is from 60 to 2,239. Finally, risk-adjusted mortality rates vary across the study hospitals. For example, the average risk-adjusted mortality rate from acute myocardial infarction is 11.80% but it ranges from 0 to 34.89%. The other four risk-adjusted mortality rates also exhibit similar variation (see Table 1).

4 Methods for data envelopment analysis

Data envelopment analysis (DEA) is a non-stochastic, non-parametric estimation of a hospital's efficiency wherein a linear programming problem with multiple inputs and outputs is solved to construct a piece-wise linear best practice frontier. A hospital's relative performance is determined in relation to its position relative to the frontier.

In our analysis, the model has an output orientation because our research focuses on whether hospitals can maximize outputs (and *potentially* revenue) holding inputs (costs) fixed. We propose that hospitals maximizing outputs can increase revenues more than hospitals that do not maximize outputs. Outputs are produced holding inputs fixed and using a fixed technology with the goal of maximizing the number of outputs given the inputs (i.e., the inverse of the output distance function). We begin with a DEA model using the traditional assumption of strong disposability of outputs (SDO) wherein the expansion of all outputs is desirable even the IQI percentages. A score of 1.00 indicates that the hospital is efficient on an output basis. A score higher than 1.00 indicates inefficiency and the potential to increase outputs.

However, as more patient care outputs are produced, given fixed inputs, there may also be increased production of the bad output. Omitting undesirable outputs that can arise in the production process from an analysis may lead to mismeasurement of the economic evaluation of a hospital's input-output correspondence [5].

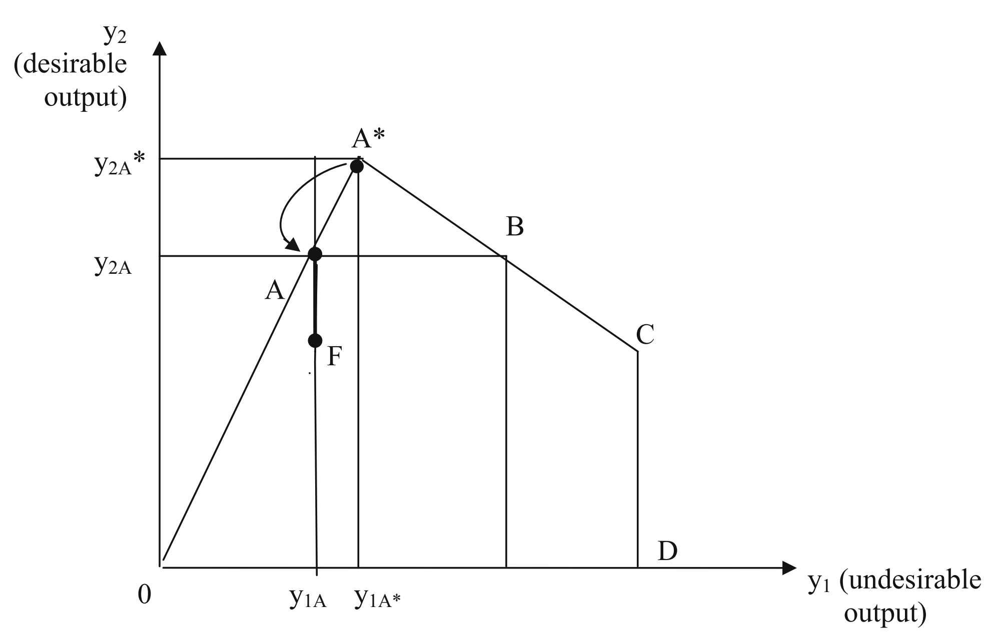
Since the purpose of the analysis is to measure the negative impact of poor outcomes on total hospital productivity, we then apply the DEA using the assumption of weak disposability of outputs (WDO) where expansion of some outputs, the IQIs, but not the number of patients at risk for an IQI, is undesirable. The ratio of the SDO to WDO efficiency scores, called congestion (C_o), reflects how much total productivity is reduced by the presence of the bad outcomes. If a hospital is more technically efficient with the WDO than SDO technology, then $C_o>1$ and output congestion exists. In other words, the bad outcomes use production capacity. Since we use the IQIs in both the SDO and the WDO models, we avoid problems with dimensionality.

This methodology has been described elsewhere as in Färe et al. [35] and applied specifically to hospital care in Ferrier et al. [36]. Thus, the methodological details are given in the Appendix. However, we illustrate the notion of congestion graphically in Fig. 1 that denotes the relationship between producing a desirable output (y_2) and an undesirable output (y_1).

First, suppose the typical production frontier wherein both y_1 and y_2 are desirable. A hospital can operate at points D, C, B, or A* on the efficient frontier or at points A or F that are not on the efficient frontier. A hospital operating at point A, given its production technology, could increase both outputs, moving to A* and becoming efficient. To move to a different point on the efficient frontier, the hospital faces a tradeoff between y_1 and y_2. A hospital operating at point B could move to point A* only by increasing y_2 and decreasing y_1. This is what occurs under the assumption of strong disposability of outputs (SDO).

However, now consider y_1 to be undesirable. That is, assume weak disposability of outputs. If a hospital is operating at point A, the only way it can increase the amount of desirable output y_2 given its technology is to incur an additional amount of y_1 and move to point A*. Similarly, the only way to reduce the undesirable output of y_1 is to move back to point A. The production possibilities frontier is now designated by 0A*CD, where the segment 0A* represents the 'backward' bend. In this case, if a hospital is operating at point A, it is efficient *only* under the assumption of weak disposability of outputs. If a hospital is operating at point F, then the total inefficiency can be decomposed as the following. From point F to point A is congestion, and from point A to point A* is technical

Fig. 1 Relationship between desirable (y_2) and undesirable outputs (y_1)



inefficiency. To measure congestion, we take the ratio of the SDO/WDO efficiency scores. For the analyses, we use OnFront Version 2 software from the R.R. Institute of Health Economics in Malmo, Sweden.

5 Results

5.1 Efficiency model scores

The mean efficiency score for the SDO model with multiple outcome indicators is 1.17 with a standard deviation of 0.18 and a range from 1.00 to 1.86. Approximately 33% (221) of all hospitals have an SDO efficiency score of 1.0, indicating that they were located on the efficient frontier. The mean efficiency score for the WDO model is 1.06 with a standard deviation of 0.11 and a range from 1.00 to 1.52. The majority of hospitals have better WDO than SDO efficiency scores meaning that by improving their production processes, most hospitals could use their capacity to produce additional desirable outputs and fewer undesirable patient outcomes. The results indicate that poor quality outcomes are costly.

The congestion score provides an indication of how much improvement is possible if undesirable outcomes are eliminated. The mean score for the congestion ratio (C_O) is 1.10. It ranges from 1.00 for hospitals without congestion to 1.86 with a standard deviation of 0.14. As noted, 221 hospitals had SDO scores of 1.00; thus, they produced outputs with given inputs without any congestion. However, 446 hospitals (67%) had some level of congestion in their production process. The hospitals exhibiting congestion had a mean of 1.15 indicating that if the average congested hospital could eliminate excess mortality rates, its efficiency would increase by 13% ($C_O - 1/C_O$). Referring back to Fig. 1, this 13% is equal to the distance from point A to A*, and the associated percentage decrease of the desirable output (y_2) produced.

Many prior assessments of quality and cost have focused on one type of care and outcome. To assess how a model using multiple outputs differs from a model with only one output, we also evaluate congestion using only the AMI risk-adjusted indicator using the same 667 hospitals as for the analysis with five outcomes. AMI quality, and specifically hospital AMI mortality, is one of the most frequently studied hospital outcomes [37, 38]. The results from this analysis show that 103 hospitals are congested and 564 are not. Thus, efficiency assessments that focus on only one outcome indicator are likely to miss opportunities for performance improvement.

Table 2 Wilcoxon nonparametric comparisons of congested and non-congested hospitals

	Five Indicators			Acute Myocardial Infarction Indicator (AMI) Only		
	Means for			Means for		
Variable	Congested	Non-Congested	p[c]	Congested	Non-Congested	p[c]
Payer mix						
% Medicare	39.08	37.50	0.10	39.45	38.39	0.14
% Medicaid	14.90	16.86	0.15	14.15	15.80	0.18
% Private pay	37.90	38.51	0.75	38.33	38.06	0.85
% Self pay	4.90	3.97	0.01	5.17	4.49	0.01
% Other	3.20	3.16	0.07	2.88	3.24	0.25
Patient race[a]	n=420	n=212		n=98	n=534	
% White	70.03	69.12	0.34	75.97	68.57	0.02
% Black	9.73	7.53	0.06	6.78	9.40	0.10
% Other	15.09	19.83	0.002	14.30	17.11	0.35
Patient Severity[b]						
% Mortality 3 or 4	14.90	14.05	0.33	15.81	14.40	0.008
% Severity 3 or 4	18.37	17.36	0.84	18.98	17.86	0.03
Hospital Resources						
% Occupancy	63.12	66.24	0.52	62.00	64.54	0.09
ALOS	5.53	5.40	0.11	5.17	5.54	0.02
FTE RNs/adj. Admission	0.025	0.024	0.15	0.023	0.025	0.003
N	446	221	–	103	564	–

[a] Percentages do not sum to 100 because the unknown race category is not shown. Analysis also excludes hospitals in Washington State because they do not report patient race.
[b] Scale: 1=lowest severity, 4=highest severity
[c] Wilcoxon ranked sum test assesses the distribution and ranks of values rather than the difference between means

5.2 Comparisons of congested and non-congested hospitals

Comparison of congested and non-congested hospitals can provide insights regarding which hospitals may face more challenges in providing good quality of care efficiently. For example, hospitals with more elderly or severely ill patients may use relatively more resources to produce the same quality of care as other hospitals. Consequently, we examine hospital payer mix, patient racial mix, patient severity and hospital resources for differences and show these results for models using the five indicators as well as only the AMI indicator in Table 2. In addition, we examine congestion scores by hospital mission variables, namely, ownership and teaching status. Finally, because hospitals operating in different states face different political considerations and, perhaps, different physician practice patterns, we looked for differences in congestion scores across the ten states. We show the means for the congestion scores by category in Table 3. To test for differences, we use nonparametric Wilcoxon ranked sum tests, which use differences in rankings rather than means.

As shown in Table 2, when using the five indicators, hospitals with a higher percentage of self-pay as well as black (p=.06) are more likely to be congested. Previous research has shown that minority patients often receive care at poorer quality facilities [39]. These facilities are also likely to have a higher burden of self-pay patients who are often uninsured. However, in contrast, hospitals with a higher percentage of other race patients are less likely to be congested.

The statistically significant relationships do not appear to be due to differences in patient severity. Neither the percentage of patients with mortality or with severity scores of 3 or 4, are statistically significantly related to congestion. However, the percentage of Medicare patients, who, because of their age, frailty or complicating conditions may be more time and resource intensive than other patients, is marginally significant (p=0.10). Thus, it is still possible that the risk-adjustments for the IQIs do not fully control for such patient risk factors. Finally, hospital occupancy, average length of stay, and full-time equivalent RNs per adjusted admission are not significantly related to congestion.

The results for the AMI only model are somewhat different. The percentage of self-pay patients is significant and the percentage of black patients is, again, marginally significant (p=.10) but in the opposite direction. In contrast, other race is not significant but hospitals with a higher percentage of white patients are more likely to be congested. Further, both patient severity percentages are positively associated with congestion. Surprisingly, hospitals with lower occupancy (p=0.09), average length of stay and nurse staffing are more likely to be congested. This

Table 3 Wilcoxon nonparametric tests of congestion scores by ownership, teaching status and state

	N	Five Indicators	Acute Myocardial Infarction Only
Ownership		Mean congestion score	Mean congestion score
Public	65	1.10	1.01
For-profit	122	1.11	1.01
Non-profit non-church	372	1.10	1.02
Non-profit church	108	1.09	1.01
Significance for Wilcoxon test[a]		p=0.13	p=0.59
Teaching			
Major teaching	61	1.08	1.01
Minor teaching	119	1.09	1.01
Non-teaching	487	1.11	1.01
Significance for Wilcoxon test[a]		p=0.19	p=0.004
State			
Arizona	33	1.15	1.02
California	161	1.06	1.00
Colorado	22	1.15	1.02
Florida	130	1.16	1.02
Iowa	25	1.11	1.01
Massachusetts	43	1.12	1.03
New Jersey	53	1.09	1.02
New York	116	1.07	1.00
Washington	35	1.13	1.02
Wisconsin	49	1.09	1.01
Significance for Wilcoxon test[a]		p<0.001	p<0.001

[a] Wilcoxon ranked sum test assesses the distribution and ranks of values rather than the difference between means.

result may occur because these measures are for the hospital, not service line. However, much previous research that employs a single outcome indicator is also restricted to such aggregated hospital measures. The difference in results between our multiple and single outcome models suggests that single outcome models may produce misleading findings.

When using the five IQIs, neither ownership nor teaching status are related to significant differences in hospital congestion. However, congestion is significantly different across states in the five IQI model. In contrast, for the AMI only model, ownership is not associated with hospital congestion but teaching status and state both are. Thus, differences in mission do not appear to be related to lost efficiency but regional differences consistently are. However, the role of teaching status depends upon whether a single or set of indicators is used.

6 Discussion

We add to the literature regarding the relationship between hospital efficiency and quality by assessing the joint production of desirable and undesirable hospital outputs. In our DEA models, we use multiple inputs and outputs to analyze performance under two different sets of assumptions. Under the strong disposability of outputs (SDO) assumption, expansion of all patient care outputs, such as admissions and outpatient visits, is desirable. In contrast, the assumption of weak disposability of outputs (WDO) treats expansion of some outputs as undesirable, namely, inpatient mortality indicators for common patient conditions that are risk-adjusted for patient characteristics and comorbidities. We use data for all hospitals operating in ten geographically diverse states during a year (2000) following a period of hospital input constraint coupled with growth in outputs. Although we focus on the year 2000 in our analysis, issues of hospital productivity and output quality continue to be salient given current constraints on hospital payments [40] and policymakers' interest in restructuring payment to reward high quality care through pay-for-performance programs [41].

Our results show that technical inefficiency is associated with poorer quality of patient outcomes and that the majority of study hospitals could improve both their technical efficiency and patient care outcomes. Thus, contrary to the frequent assumption that quality costs more, our findings indicate that efficiency and quality go together. They are consistent with other studies that report higher quality is associated with lower cost and greater efficiency [25, 26]. The results are particularly interesting because they show that many hospitals (33% of our sample) were able to adapt to any resource constraints from the BBA cutbacks without harming quality of care. With our data, we were able to identify only a few hospital characteristics consistently differentiated these hospitals from the others; they are payer mix and location (state). Future research should examine other organizational characteristics such as market competition for their association with congestion.

Our study also demonstrates the application of a new empirical approach that can be used by researchers to study a variety of complex issues in health care. Of great current interest are pay for performance approaches and mandated nursing staffing ratios as in states like California. Our results suggest that paying for better performance, that is, higher quality, may not mean higher overall payments because the bonuses may be compensated for by cost savings from reduced inefficiency. Another issue of great concern is whether physicians who own specialty hospitals are steering less severe patients to their institutions because they will yield greater profits given current payment policies [42]. Using our framework, researchers may be able to conduct analyses that reveal whether hospitals that are relatively efficient under the SDO assumption are relatively less efficient once one accounts for patient severity adjusted outcomes under the WDO assumption. Such a finding would provide insights on the presence and extent of cream-skimming among specialty versus general hospitals.

Finally, the differences in findings between the models with multiple risk-adjusted quality indicators and a single indicator, AMI mortality, show that policy makers and researchers should exercise caution when examining single outcome indicators. Indeed, previous research has shown that hospitals that perform well on one indicator may not perform as well on others [43].

7 Limitations

The study has several limitations. First, although the risk-adjustment methodology has extensive support, as with previous research, it is possible that our risk-adjusted patient outcome measures did not fully adjust for patient risk factors. Thus, it is possible that some of the results, such as those for race, are related to imperfect risk adjustment.

Second, because of the data requirements for the DEA models—e.g., reporting five indicators, having at least 30 patients at risk for an indicator, etc.—our study hospitals are not representative of all US hospitals. They are generally larger and urban hospitals with a lower case mix adjusted length of stay. Thus, our results may not be generalizable to smaller or rural hospitals.

Third, DEA yields relative not absolute efficiency measures. The benefit of use the DEA approach in our study is the ability afforded to us to simultaneously measure the production of hospital care goods (outputs) while accounting for the poor outcomes that detract from overall social welfare. One of the major limitations is that the productivity of the hospitals in our sample are all assessed by a best practice frontier whose definition is limited by including all hospitals in our sample. In essence, we are gauging hospital efficiency in a relative rather than an engineering sense. Because of this limitation, we cannot make any definitive statements about absolute quality. However, we do note that the relative reduction in congestion is a necessary albeit not sufficient condition in maximizing patient and hospital welfare.

Finally, as in most other hospital studies, we faced limitations in the detail of data available. Many of the data elements, such as nurse staffing, are only available at the hospital not service line of patient care unit level. Other data elements such as capital inputs are not further disaggregated by type or amount. Unfortunately, in this analysis, we were limited to beds as an indicator of capital input. However, hospitals with a larger number of beds also typically invest more in other types of capital such as imaging equipment.

Future research may be able to resolve some of the problems noted above as better risk-adjustment methods are developed in the future and as additional data for more hospitals and in more states become available. As more data become available, refinements such as how different ownership in different states is related to congestion, may be worthy of investigation. Certainly, replication of our approach using data from more recent year and with additional subsets of quality indicators will be worthwhile to assess whether our principal findings are robust.

Acknowledgements Research for this paper was partially supported by a grant from the Agency for Healthcare Research and Quality (HS 13094-03). The authors would like to thank the anonymous reviewers for their helpful comments.

Appendix

The nonparametric production framework used to construct a piece-wise linear best practice frontier under the assumption of variable returns to scale is defined as:

$$P(x|V,S) = \left\{ y : y \leq z \cdot M, z \cdot K \leq x, z \in \Re, \sum_{j=1}^{N} z_j = 1 \right\},$$

where x inputs are used to produce y outputs. V indicates there are variable returns to scale and S shows the assumption of strong disposability of outputs (SDO). **M** is a matrix of the m outputs of each of the N hospitals and **K** is a matrix of each of the k inputs for each of the observations. In the SDO model, all outputs are considered strongly disposable—desirable and undesirable. The z represents the intensity variables required to map out the best practice frontier.

However, by relaxing the assumption of strong disposability on a vector of outputs, we define this new frontier as:

$$P(x|V,W/S) = \left\{ \left(y^W,\ y^S\right) : y^W \leq \mu \cdot z \cdot M^W,\ y^S \leq z \cdot M^S,\ z \cdot K \leq x,\ 0 \leq \mu \leq 1,\ z \in \Re, \sum_{j=1}^{N} z_j = 1 \right\},$$

where the superscript of “s” represents strong disposability and “w” denotes weak disposability. The “μ” is imposed to allow the weakly disposable outputs to move along the backward bend of the production possibility frontier. In other words, it permits for the non-linear scaling of these outputs whereas the other outputs can only be radially increased in a linear fashion. The “μ” parameter provides the measure of the “bad” output, in our case the poor health outcomes.

Computationally, we solve two linear programming models:

$$\begin{aligned} F_O(x,y|V,S) = \ & \max_{z,\theta^S} \theta^S \\ s.t.\ \ & \theta^S \cdot y \leq z \cdot M \\ & z \cdot K \leq x \\ & z \in \Re_N \\ & \sum_{j=1}^{N} z_j = 1\,. \end{aligned}$$

and

$$\begin{aligned} F_O(x,y|V,\ W/S) \ = \ & \max_{z,\theta^{W/S}} \ \theta^{W/S} \\ s.t.\ \ & \theta^{W/S} \cdot y^W \leq \mu \cdot z \cdot M^W \\ & \theta^{W/S} \cdot y^S \leq z \cdot M^S \\ & z \cdot K \leq x \\ & 0 \leq \mu \leq 1 \\ & z \in \Re_N \\ & \sum_{j=1}^{N} z_j = 1 \end{aligned}$$

The first model measures output efficiency where all outputs are considered positively, i.e., strongly disposable. The second model measures the technology in which case some of the outputs are considered as weakly disposable (WDO). Finally, by taking the ratio of the results from these two models:

$$C_O(x,y|V) = \frac{F_O(x,y|V,S)}{F_O(x,y|V,W/S)} = \frac{\theta^{S*}}{\theta^{W/S*}} \geq 1$$

we can derive a measure of congestion. If the hospital’s efficiency measures are equal under both assumptions, there is no congestion. In contrast, if a hospital is more efficient in the SDO than the WDO model, congestion exists and $C_o > 1$. C_o reflects how much of total productivity is reduced by the presence of these “bad” hospital outcomes.

References

1. Bazzoli GJ, Lindrooth RC, Hasnain-Wynia R, Needleman J (2004/2005) The Balanced Budget Act of 1997 and US hospital operations. Inquiry 41:401–417
2. Lindrooth RC, Bazzoli GJ, Needleman J, Hasnain-Wynia R (2006) The effect of changes in hospital reimbursement on staffing decisions at safety net and non-safety net hospitals. Health Serv Res 41:701–720
3. Bazzoli GJ, Clement J, Lindrooth RC, Chen HF, Aydede S, Braun B, Loeb, J (2007) Hospital financial condition and operational decisions related to the quality of hospital care. Med Care Res Rev 64:148–168
4. American Hospital Association (AHA). (2001) TrendWatch 2001 Volume 3 Number 3. American Hospital Association, Chicago, IL
5. Balk B (2003) The residual: on monitoring and benchmarking firms, industries, and economies with respect to productivity. J Prod Anal 20:5–47
6. Aiken LH, Smith, HL, Lake ET (1994) Lower Medicare mortality among a set of hospitals known for good nursing care. Med Care 32:771–787
7. Blegen MA, Goode CJ, Reed L (1998) Nurse staffing and patient outcomes. Nursing Research 47:43–50 (January–February)
8. Blegen MA, Vaughan T (1998) A multi-site study of nurse staffing and patient occurrences. Nurse Economics 16:196–203 (July–August)
9. Kovner C, Gergen PJ (1998) Nurse staff levels and adverse events following surgery in US hospitals. Image J Nurs Scholarsh 30:315–321
10. Lichtig LK, Knauf RA, Miholland DK (1999) Some impacts of nursing on acute care hospital outcomes. J Nurs Adm 29:25–33 (February)
11. Manheim LM, Feinglass J, Shortell SM, Hughes EFX (1992) Regional variation in Medicare hospital mortality. Inquiry 29:55–66 (Spring)
12. McCloskey JM (1998) Nurse staffing and patient outcomes. Nurs Outl 46:199–200 (September–October)
13. Provonost P, Dorman T, Jenckes M, Garret E, Breslow M, Rosenfield B, Lipsett P, Bass E (1999) Organizational characteristics of intensive care units related to outcomes of abdominal aortic surgery. JAMA 281:1310–1317
14. Schultz MA, van Servellen G, Chang B, McNeese-Smith D, Waxenberg, E (1998) The relationship of hospital structural and financial characteristics to mortality and length of stay in acute myocardial infarction patients. Outcomes Manag Nurs Pract 2:130–136
15. Schultz MA, van Servellen G, Litwin MS, McLaughlin EJ, Uman GC (1999) Can hospital structural and financial characteristics explain variations in mortality caused by acute myocardial infarction? Appl Nurs Res 12:210–214 (November)
16. Silber JH, Rosenbaum PR (1995) Measuring quality of hospital care. JAMA 273:21–22
17. Taunton RL, Kleinbeck SV, Stafford R, Woods CQ, Bott MJ (1994) Patient outcomes: are they linked to registered nurse absenteeism, separation, or work load? J Nurs Adm 24:48–55 (July–August, Suppl)
18. van Servellen G, Schultz MA (1999) Demystifying the influence of hospital characteristics on inpatient mortality rates. J Nurs Adm 29:39–47
19. Picone GA, Sloan FA, Chou SY, Taylor DH (2003) Does higher hospital cost imply higher quality of care? Rev Econ Stat 85(1):51–62
20. Burstin HR, Lipsitz SR, Udvarhelyi IS, Brennan TA (1993) The effects of hospital financial characteristics on quality of care. JAMA 270:845–849 (August 18)
21. Dranove D, White WD (1998) Medicaid-dependent hospitals and their patients: how have they fared?. Health Serv Res 33:163–186 (June, Part I)
22. Lindrooth RC, Bazzoli GJ, Clement J (2007) Hospital reimbursement and treatment intensity. South Econ J 73:575–587
23. Rapoport J, Teres D, Lemeshow S, Gehlbach (1994) A method for assessing the clinical performance and cost-effectiveness of intensive care units: a multicenter inception cohort study. Crit Care Med 22(9):1385–1391
24. Bradbury RC, Golec JH, Steen PM (1994) Relating hospital health outcomes and resource expenditures. Inquiry 31:56–65 (Spring)
25. Carey K, Burgess JF (1999) On measuring the hospital cost/quality trade-off. Health Econ 8:509–520
26. Deily ME, McKay NL (2006) Cost inefficiency and mortality rates in Florida hospitals. Health Econ 15:419–431
27. Fleming ST (1990) The relationship between the cost and quality of hospital care: a review of the literature. Med Care Res Rev 47:487–502 (Winter)
28. Hoerger TJ (1991) 'Profit' variability in for-profit and not-for-profit hospitals. J Health Econ 10:259–289
29. California Healthcare Association (2003) Implementation of California's Nurse Ratio Law, Summary of the law. Available at http://www.calhospital.org. Accessed January 12, 2007
30. Donabedian A (1988) The quality of care: how can it be assessed? JAMA 260:1743–1748
31. Agency for Healthcare Research and Quality (2002) Inpatient quality indicators: software documentation, version 2.1. Department of Health and Human Services, Washington, DC. AHRQ Pub No. 02-R0205, Revision 1. July 3, 2002
32. Averill, RF, Goldfield N, Hughes JS, Bonazelli J et al (2003) All patient refined DRGs (APR-DRGs), Version 20.0. Available at http://www.hcup-us.ahrq.gov/db/nation/nis/APR-DRGsV20MethodologyOverviewandBibliography.pdf , Accessed March 6, 2007
33. American Hospital Association (2006) AHA Hospital statistics 2006 Edition. Health Forum, LLC, Chicago, IL
34. American Hospital Association (2006) TrendWatch Chartbook 2006. American Hospital Association, Chicago, IL. Available at http://www.aha.org/aha/research-and-trends/health-and-hospital-trends/2006.html, Accessed March 6, 2007
35. Fare R, Grosskopf S, Lovell CAK (1994) Production frontiers. Cambridge University Press, New York
36. Ferrier GD, Rosko MD, Valdmanis VG (2006) Analysis of uncompensated hospital care using a DEA model of output congestion. Health Care Manage Sci 9:181–188
37. Volpp KGM, Williams SV, Waldfogel J, Silber JH, Schwartz JS, Pauly MV (2003) Market reform in New Jersey and the effect on mortality from acute myocardial infarction. Health Serv Res 38(2):515–533
38. Gowrisankaran G, Town R (2003) Competition, payers, and hospital quality. Health Serv Res 38(6 Pt 1):1403–1422
39. Mayberry RM, Mili F, Ofili E (2000) Racial and ethnic differences in access to medical care. Med Care Res Rev 57 (Supplement 1):108–145
40. Medicare Payment Advisory Commission (2006) Data Book: Healthcare Spending and the Medicare Program, June 2006
41. Medicare Payment Advisory Commission (2007) Report to the Congress: Medicare Payment Policy, March 2007
42. Medicare Payment Advisory Commission (2006) Report to the Congress: Physician-Owned Specialty Hospitals Revisited, August 2006
43. Jha AK, Li Z, Orav EJ, Epstein AM (2005) Care in U.S. hospitals—the hospital quality alliance program. N Engl J Med 353(3):265–274