Anesthesiology 2007; 106:346–55

Estimating Anesthesia and Surgical Procedure Times from Medicare Anesthesia Claims

Jeffrey H. Silber, M.D., Ph.D., Paul R. Rosenbaum, Ph.D.,† Xuemei Zhang, M.S.,‡ Orit Even-Shoshan, M.S.§*

Background: **Procedure times are important variables that often are included in studies of quality and efficiency. However, due to the need for costly chart review, most studies are limited to single-institution analyses. In this article, the authors describe how well the anesthesia claim from Medicare can estimate chart times.**

Methods: **The authors abstracted information on time of induction and entrance to the recovery room ("anesthesia chart time") from the charts of 1,931 patients who underwent general and orthopedic surgical procedures in Pennsylvania. The authors then merged the associated bills from claims data supplied from Medicare (Part B data) that included a variable denoting the time in minutes for the anesthesia service. The authors also investigated the time from incision to closure ("surgical chart time") on a subset of 1,888 patients.**

Results: **Anesthesia claim time from Medicare was highly predictive of anesthesia chart time (Kendall's rank correlation τ = 0.85, P < 0.0001, median absolute error = 5.1 min) but somewhat less predictive of surgical chart time (Kendall's τ = 0.73, P < 0.0001, median absolute error = 13.8 min). When predicting chart time from Medicare bills, variables reflecting procedure type, comorbidities, and hospital type did not significantly improve the prediction, suggesting that errors in predicting the chart time from the anesthesia bill time are not related to these factors; however, the individual hospital did have some influence on these estimates.**

Conclusions: **Anesthesia chart time can be well estimated using Medicare claims, thereby facilitating studies with vastly larger sample sizes and much lower costs of data collection.**

THE surgical and anesthesia literature commonly reports procedure time.[1–5] Although surgical procedure time may be influenced by initial severity of the patient's condition, it is generally believed that the longer a surgical procedure, the greater the probability of a complication or death.[6] This belief is fundamental and pervasive throughout the surgical literature.[7–10] For example, procedure time is used in the National Nosocomial Infection Surveillance System,‖ a scoring system in which a patient's risk of infection is predicted, in part, from procedure time. The length of a procedure also has an impact on cost, both with respect to the opportunity cost of the operative suite[11,12] and the cost of labor and anesthetic agents.[13–15]

However, because obtaining procedure time usually requires chart review, the variable is often analyzed from single-institution studies (or studies including a small number of cooperating institutions) in which such information is more easily abstracted.[1,3–5,11] One exception is the multi-institutional Veteran's Administration data base,[2,10] which has recorded and reported procedure times among institutions. Our present study explores how well the Medicare anesthesia claim ("anesthesia claim time") can also be used to estimate actual chart times associated with anesthesia ("anesthesia chart time") and surgery ("surgical chart time"). We did this by measuring both the chart times and the claim times for 1,931 Medicare patients throughout Pennsylvania. If procedure time could be well estimated using Medicare anesthesia bills, the study of the length of surgical procedures could be expanded to a much larger Medicare population, enabling the study of a vast array of research questions more appropriate for the more general Medicare population.

Materials and Methods

Patients and Databases

Medicare data for patients 65 yr and older is the most representative healthcare data for the elderly in the United States because Medicare is an entitlement program. The only significant group of elderly citizens not represented in the Medicare claims is those who opted out of the Medicare fee-for-service arrangement and joined a Medicare-approved prepaid health maintenance organization. As part of the Surgical Outcomes Study,[16–18] we obtained the Medicare Inpatient (Part A), Outpatient Standard Analytic Files, and Physician Part B files for all admissions in general and orthopedic surgical Diagnosis Related Groups in Pennsylvania during 1995 and 1996. These files represent the fee-for-service Medicare population, comprising approximately 90% of all beneficiaries for 1995–1996.[19,20] We created a longitudinal record by including all inpatient and outpatient claims and physicians' claims during that time interval for each patient. Data also included the Medicare

* Professor, The Center for Outcomes Research, The Children's Hospital of Philadelphia; Department of Anesthesiology and Critical Care and Department of Pediatrics, The University of Pennsylvania School of Medicine; and The Leonard Davis Institute of Health Economics, Philadelphia, Pennsylvania. † Professor, Department of Statistics, The Wharton School, The University of Pennsylvania. ‡ Senior Statistical Programmer, The Children's Hospital of Philadelphia. § Associate Center Director, The Center for Outcomes Research, The Children's Hospital of Philadelphia, and The Leonard Davis Institute of Health Economics.

Received from The Center for Outcomes Research, The Children's Hospital of Philadelphia and The University of Pennsylvania. Submitted for publication January 10, 2006. Accepted for publication September 20, 2006. This study was funded by a grant from the Leonard Davis Institute of Health Economics, The University of Pennsylvania, Philadelphia, Pennsylvania and grants from the Agency for Healthcare Research and Quality, Rockville, MD, R01-HS-9460, and SES-0345113 from the Methodology, Measurement and Statistics Program and the Statistics and Probability Program of the U.S. National Science Foundation, Arlington, Virginia..

Address correspondence to Dr. Silber: The Center for Outcomes Research, The Children's Hospital of Philadelphia, 3535 Market Street, Suite 1029, Philadelphia, Pennsylvania 19104. silber@email.chop.edu. Individual article reprints may be purchased through the Journal Web site, www.anesthesiology.org.

‖ National Nosocomial Infection Surveillance System. Available at: http://www.cdc.gov/ncidod/dhqp/nnis_pubs.html. Accessed February 23, 2005.

Vital Status File, American Hospital Association Annual Survey for 1996, and the Pennsylvania Health Care Cost Containment Council Hospital Discharge Database for similar years, which included the MedisGroups© (MediQual, Inc., Marlborough, MA) severity score to supplement the Medicare record.[21] Medicare electronically stores all bills submitted by hospitals, physicians, and all other caregivers, and all payments made. These billing data include limited diagnosis codes collected using International Classification of Diseases (ICD9-CM) coding and procedure specific codes using both ICD9-CM coding and the Healthcare Common Procedure Coding System (HCPCS).

Medicare claims files can be obtained through the Research Data Assistance Center, which is a contractor that provides free assistance to academic, government, and nonprofit researchers interested in using Medicare and/or Medicaid data for their research. The Research Data Assistance Center is staffed by a consortium of epidemiologists, public health specialists, health services researchers, biostatisticians, and health informatics specialists from the University of Minnesota.#

A patient's hospital bill from the Inpatient Standard Analytical file can be linked to the bills submitted by the physicians and other providers who took care of the patient during the hospital stay. Bills submitted by anesthesia providers are identified by the variable "provider specialty." For example, when this variable is "5," it identifies an anesthesiologist, and "43" identifies a Certified Registered Nurse Anesthetist. We use the variables "from date" and "through date" in the Inpatient File (Part A) and "expense date" in physician claims file (Part B) to link the information from the hospital stay to the physician bills submitted for services provided during the same stay. The inpatient file provides the ICD9-CM procedure codes for up to five procedures. For this study, we selected the principal procedure as determined in retrospect by Medicare. This procedure was identified in Part B through its corresponding HCPCS code, and its "expense date" in Part B was matched with the anesthesia HCPCS that had the same expense date as the HCPCS corresponding to the principal procedure.

There is an anesthesia time unit, defined as a 15-min interval, associated with each anesthesia provider claim.**†† The time units are identified by the variable "mile/time/units/services indicator code." When this variable equals "2," it identifies anesthesia time units. In the documentation received with the electronic claims file, time units are reported as integers but should be interpreted as having one decimal. For example, a time unit value of 25 implies 2.5 time units $\times$ 15 min, or 37.5 min billed by the anesthesia provider.

According to the Medicare Claims Manual, Section 50G,** the anesthesia time "starts when the anesthesia practitioner begins to prepare the patient for anesthesia services in the operating room or an equivalent area and ends when the anesthesia practitioner is no longer furnishing anesthesia services to the patient, that is, when the patient may be placed safely under postoperative care." Time units, unlike base units, are not modified by the performance and direction of concurrent anesthesia procedures.** Base units are assigned to each anesthesia HCPCS code and reflect the relative difficulty of the anesthesia procedure.‡‡ Base units were not used for this analysis. Both the time and base units are used in determining payment from Medicare.

We obtained detailed chart abstraction data in a subset of patients from the same pool of patients for which we had claims data. The charts were abstracted as part of the Surgical Outcomes Study.[16–18,22–24]

Defining Operative Time

Chart Abstraction Algorithm. Four landmark times were abstracted from each available chart: induction, incision, closure, and entrance to recovery room. These abstraction definitions are not exactly those used by the nomenclature convention of the American Association of Clinical Directors,** nor are they consistent with "anesthesia-controlled time,"[3,11,25] because the intent of this manuscript was not to calculate operating room throughput and overall efficiency. Our landmarks were chosen because they reflect how reimbursement from Medicare is determined and because the abstractors for Medicare had used these landmarks in previous work auditing Medicare charts. We defined "anesthesia chart time" ($n = 1,931$ patients) as time of recovery room entrance minus time of induction and expressed this in minutes. We defined "surgical chart time" ($n = 1,888$ patients) as the time of closure minus the time of incision, also expressed in minutes. We excluded patients from the anesthesia time analyses if either induction or recovery room time was missing; we also excluded patients from the surgical time analysis if either incision or closure time was missing.

Claims Analysis Algorithm. Each bill for anesthesia services for a patient for a surgical procedure date may include none, one, or more than one claim from the Medicare Part B files, an associated Current Procedural Terminology or HCPCS code, and a variable identifying the specialty of the anesthesia provider. We tested nu-

Research Data Assistance Center. Available at: http://www.resdac.umn.edu. Accessed January 2, 2007.

** Medicare Claims Processing Manual, Chapter 12: Physicians/Nonphysician Practitioners; Section 50: Payment for Anesthesiology Services; Part G: Anesthesia Time and Calculation of Anesthesia Time Units, pp. 91–92. Available at: http://www.cms.hhs.gov/manuals/downloads/clm104c12.pdf. Accessed July 7, 2006.

†† Anesthesia Billing Guide, Section F: Time Units. Available at: http://www.hgsa.com/professionals/bguides/anesthesia.shtml. Accessed April 2004.

‡‡ Medicare Claims Processing Manual, Chapter 12: Physicians/Nonphysician Practitioners; Section 50: Payment for Anesthesiology Services; Part A: General Payment Rule, pp. 88–89; Part H: Base Unit Reduction for Concurrent Medically Directed Procedures, p. 92; and Part K: Anesthesia Claims Modifiers, pp. 94–110. Available at: http://www.cms.hhs.gov/manuals/downloads/clm104c12.pdf. Accessed July 7, 2006.

merous definitions for variable construction concerning the total claim minutes used in the predictive model. We defined the claim time to equal the longest time billed by any anesthesia provider on the day of the surgery. If two anesthesia bills were found for the same date as the principal procedure, the longer bill was used in the modeling. We considered several other possible definitions of the claim time, but they worked poorly. For instance, using a summation of the bills, or modeling two bills as two variables, did not improve the simple rule of using the longest anesthesia provider bill. When we included nurse anesthetist bills separately, by adding a second variable to the modeling process, we found almost no change in fit (R^2), and any coefficients on the second variable that were significant were extremely small, producing no appreciable difference in our estimates. Hence, we used a simple rule for defining claim time—the longest bill by an anesthesia provider on the same day as the principal procedure in question.

Statistical Methods. We used Kendall's τ as a measure of correlation between claim and chart times. Kendall's τ is useful because magnitudes of τ other than -1, 0, and 1 can be given a practical interpretation.[26] Consider two patients. For these two patients, the bill and the chart are said to be concordant if the patient with the longer chart time also had the longer bill time. The probability θ that two patients will be concordant is related to Kendall's rank correlation τ by the formula $\theta = (\tau + 1)/2$. A perfect correlation yields a $\tau = 1$, yielding a concordance probability of $\theta = (\tau + 1)/2 = (1 + 1)/2 = 1$. Zero correlation, $\tau = 0$, yields random or coin-flip concordance of $\theta = (\tau + 1)/2 = (0 + 1)/2 = 0.5$, or a random chance that a longer claim time will also coincide with a longer chart time. We used the claim time to predict the chart time because researchers with data from Medicare will have the claim times and our algorithm for predicting the chart times from the claim times. The question then is how well that formula will perform. Associated with Kendall's τ is a form of simple linear regression yielding Theil's slope estimate.[27] When reporting results for individual hospitals or overall results with just a single claim variable, we used Theil's slope estimate, which performs well even in small samples, such as those for individual hospitals, and is insensitive to one or two wild observations in either x or y. When performing multiple regression, we used Huber's robust m-estimation as implemented in SAS Version 9 (SAS Institute, Inc., Cary, NC) using the bisquare weight function.[28–30] In least-squares regression, R^2 is the square of the Pearson correlation between observed and predicted values of y = chart time; however, because we performed robust, outlier-resistant regression, the use of the outlier-sensitive Pearson correlation is not appropriate. Therefore, in our robust regressions, we report as R^2 (or rank R^2) the square of the Spearman rank correlation between the observed and expected y's, which is analogous to the square of the Pearson correlation between the observed and predicted ranks of y = chart time. This prevents one or two peculiar claims from increasing or decreasing the R^2.

Results

Available Data

There were 2,259 abstracted charts available for review from the Surgical Outcomes Study. The Surgical Outcome Study was a case-control study (two controls for each case) of mortality after surgery, so one third of the patients had died within 2 months of surgery. Of these, 2,007 have an anesthesia bill from Medicare Part B data. We excluded 30 charts with missing chart anesthesia time or negative chart anesthesia time and 40 charts with surgery time longer than anesthesia time. We excluded five charts for operations that were less than or equal to 10 min in length according to the anesthesia claim; we also excluded one operation with claim anesthesia bill time greater than 900 min. Of the 1,931 charts remaining, all included both induction and recovery room times, and 1,888 charts included incision and closure times as well as induction and recovery room times.

Estimating Anesthesia Chart Time

For each analysis in table 1, we let the anesthesia claim time represent the independent variable and predict anesthesia chart time as abstracted as part of the Surgical Outcome Study. We present five models: Model I uses only the claim time to predict chart time. Model I estimates chart times from the claim using a formula of the form: chart time = $\alpha + \beta$ (claim time). Model I is estimated in two ways. The first uses Theil's method,[26] the second uses Huber's m-estimation.[31] The remaining models are intended to verify the performance of Model I; they are not intended for practical use. If the bill time is to be used in lieu of the chart time, one would prefer that other patient attributes not alter the conversion formula. Models II-V use m-estimation, which allows multiple regression[28,29] and can produce a Wald test[32] on groups of coefficients analogous to an F test in linear regression. Model II uses claim time and patient procedures; Model III uses claim time and patient comorbidities; Model IV uses claim time and hospital indicators; and Model V includes all variables. Model I is the simplest model and performs as well as models that include many other predictors. Theil's estimate yields the equation (chart time) = $y = \alpha + \beta x = 0.85 + 0.97x = 0.85 + 0.97$ (bill time), or less than 1 min added to the bill time. Hence, this equation suggests that using the claim-derived anesthesia time unit alone, without the coefficients in the Theil model, would produce very similar estimates of the desired chart time.

Table 1. Regression Models to Estimate Anesthesia Chart Time Based on Anesthesia Claim Time and Other Adjustments

Model Adjustments	Intercept	Coefficient on Claim (min)	Coefficient *P* Value	Model R^2	Model Median Absolute Residual (error in min)	Contrast *P* Value (Wald test) *vs.* Model I
I. Unadjusted Theil estimation	0.85	0.97 (0.96–0.97)	< 0.0001	0.89	5.14	—
I. Unadjusted m-estimation	–1.21	0.97 (0.97–0.98)	< 0.0001	0.89	5.49	—
II. Comorbidity*	–1.67	0.97 (0.97–0.98)	< 0.0001	0.89	5.62	0.25
III. Procedure†	–1.06	0.97 (0.97–0.98)	< 0.0001	0.89	5.59	0.28
IV. Hospital‡	2.38	0.98 (0.97–0.98)	< 0.0001	0.89	5.28	< 0.0001
V. Comorbidity, procedure, and hospital indicators	2.47	0.98 (0.97–0.98)	< 0.0001	0.89	5.37	< 0.0001

n = 1,931.

* Myocardial infarct, arrhythmia, heart failure, unstable angina, angina, hypertension, aortic stenosis, pulmonic stenosis, chronic obstructive pulmonary disease, asthma, cancer, bad cancer, psychosis, electrolyte and fluid abnormality, alcoholism, liver dysfunction, renal dysfunction, renal failure, diabetes, insulin dependent diabetes, paraplegia, stroke, seizures, collagen vascular disease, coagulopathy, hemophilia, thrombocytopenia, smoking history, postinflammatory pulmonary fibrosis, and Graves disease. † Shoulder, back, knee, other orthopedic, colon resection, stomach, hernia, gall bladder, skin graft, other surgery, and hip. ‡ There were 21 hospital indicators representing 20 hospitals with at least 14 patients and 167 hospitals included as a grouped variable.

Furthermore, as can be seen in table 1, patient characteristics such as comorbid disease or procedure type did not significantly alter the predictions made by the equation that used bill time alone, and the anesthesia time units from Medicare are excellent estimates of the chart abstracted time. The R^2 statistics are unchanged to two digits regardless of the model used, and the slopes used to convert the anesthesia claims to chart minutes are almost identical. There is statistically significant variation in the bill-to-chart relationship among hospitals, but its impact on predictions for individual patients is small: the improvements in R^2 and the median absolute error of prediction were very small, and the estimated slope was essentially unchanged. The median absolute error of prediction without the hospital indicators was 5.49 min, and this decreased to 5.28 min using the hospital indicators, a difference that is statistically significant given the sample size but perhaps of little clinical relevance. When we replaced individual hospital indicators with an indicator variable for teaching status, we found that teaching was associated with a clinically small, but statistically significant, difference between claim and chart (-0.7 min, $P < 0.0001$). However, this difference did not influence the overall R^2 of the model. We also explored whether designation of urban or rural hospitals influenced the difference between chart time and claim time as determined from the anesthesia time unit information, and it did not. There were also no differences in errors when we compared procedures that were longer than the median procedure time with those that were shorter. These results suggest that using only Medicare anesthesia time units in the context of our algorithm for assigning bills to procedures is probably adequate for most comparisons focused on patients but may be more problematic when comparing, for example, two hospitals to determine which one performs a given operation more quickly.

Figure 1 displays the overall data set with Theil's regression line fit through the points. The fit was very good, corresponding to a probability of concordance of 0.93. If one compares two patients, 93% of the time, the patient with the longer bill time also has the longer chart time.

Figure 2 displays a box plot of the 1,931 converted bill times (anesthesia time units $\times$ 15 min) compared with the abstracted values of anesthesia time. Most errors are very small; in fact, the median absolute error was just 5 min, and 95% confidence interval (CI) for the median absolute error was 5-5.5 min. To the right of the box plot is the corresponding normal probability plot. Errors that were normally distributed would produce a straight line, and here we see that the data are symmetric but far from normally distributed, with long tails, as also evidenced by the fact that the Shapiro-Wilk test W for normality[26] was 0.69 ($P < 0.0001$). The lack of normality in these errors suggests that, when using the algorithm we suggest with the Medicare Part B claims, investigators should use robust regression techniques instead of ordinary least squares.[28]

Table 2 displays a set of regressions similar to those in table 1, but the dependent variable is anesthesia bill time minus chart time alone, instead of chart time, and the

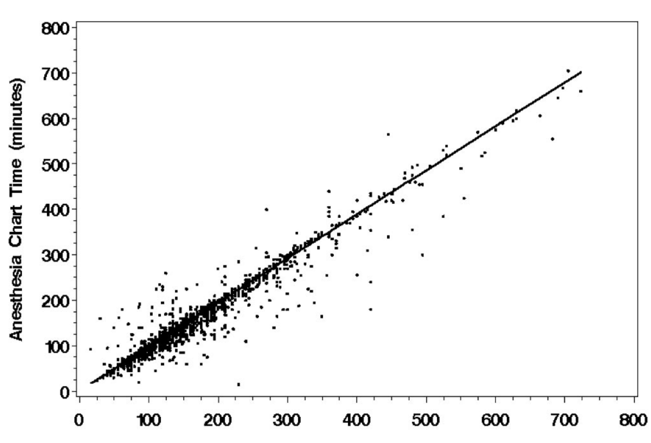

Fig. 1. Theil regression plot, n = 1,931. The independent variable is anesthesia claim minutes, and the dependent variable is anesthesia chart minutes. R^2 = 0.89.

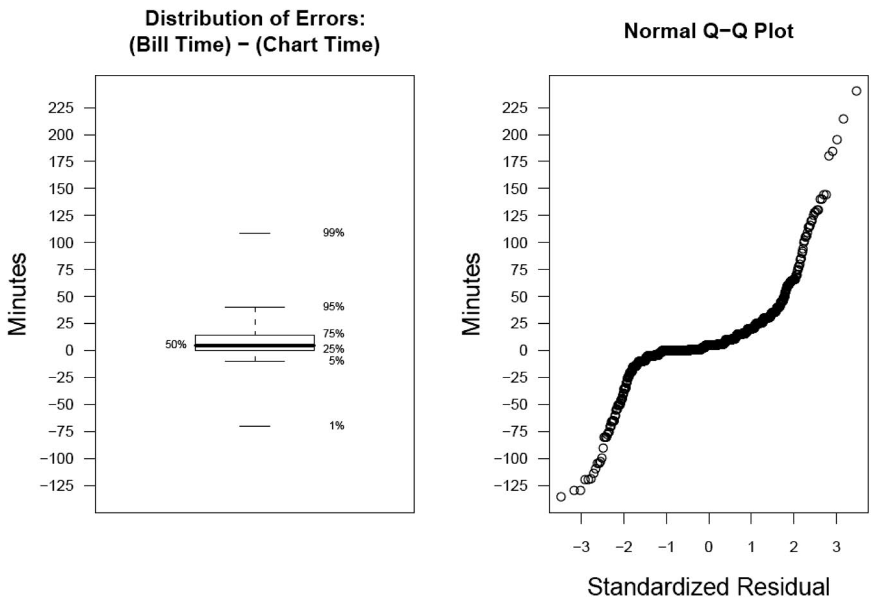

Fig. 2. Distribution of differences between data received from the Part B Medicare bill (time in minutes derived from anesthesia time units) and the chart time abstraction for anesthesia services (noting time from induction to entry to recovery room). The Shapiro-Wilk test statistic W for normality was 0.69 (*P* < 0.0001), suggesting that these differences are not normally distributed.

right side of the equations do not include bill times. In this set of analyses, we are investigating whether there are systematic errors in the billing of time data as a function of hospital characteristics such as teaching status or for-profit status. As shown in table 2, we did not see any appreciable differences among hospitals.

To better explore the differences in the estimation of chart time by claim time within and among hospitals, we estimated the relationship between anesthesia chart time and bill time separately for each hospital with at least 14 patients. Table 3 presents Kendall's τ and the associated Theil's slope and intercept for each of the 20 hospitals with 14 or more patients in the Surgical Outcomes Study and for all hospitals together. For 19 of these 20 hospitals (that is, all hospitals except hospital 11), the estimated probability of concordance between chart and bill times was greater than 0.9, so for two patients in these hospitals, more than 90% of the time, the patient with the longer bill time will also have the longer chart time. The 20 slopes and intercepts estimated separately from the 20 hospitals were generally very similar despite the small sample sizes in many of the hospitals. We next investigated how different a claim-based estimate of anesthesia time per hospital would be from the actual abstracted anesthesia chart time. For each hospital in table 3, we computed the difference between the anesthesia time based on the bill using unadjusted anesthesia time units (1 unit = 15 min) and the observed anesthesia time based on chart review. Using the estimated anesthesia times based on claims produced, on average, estimates very similar to the observed chart times. Of the 20 hospitals in table 3, only 2 had median errors greater than 10 min (hospitals 4 and 11). Each of these displayed 19.5-min median errors, or exactly 15 min (1 time unit) more than the median error associated with all hospitals combined. The *P* value in the last column of table 3 is for a hospital dummy variable in robust regression controlling for comorbidity and procedure. Although the single equation works well for converting bill to chart times, several hospitals do exhibit statistically significant deviations.

We display the nine hospitals with the most study patients in our data set in figure 3. When estimating anesthesia claim time from chart time, we consistently

Table 2. Regression Models to Estimate Difference between Anesthesia Claim Time and Anesthesia Chart Time Using Adjustments Described in Table 1

Model Adjustments	Model R^2	Model Median Absolute Residual (error in min)
II. Comorbidity*	0.012	5.50
III. Procedure†	0.003	5.21
IV. Hospital‡	0.046	5.06
V. Comorbidity, procedure, and hospital indicators	0.062	5.53

n = 1,931 patients. The unadjusted median difference between claim time based solely on the anesthesia time unit and chart time from abstraction was 4.5 min (95% CI, 4.5-4.5). The median absolute difference between claim and chart was 5.5 min (95% CI, 5.0-5.5).

*Myocardial infarct, arrhythmia, heart failure, unstable angina, angina, hypertension, aortic stenosis, pulmonic stenosis, chronic obstructive pulmonary disease, asthma, cancer, bad cancer, psychosis, electrolyte and fluid abnormality, alcoholism, liver dysfunction, renal dysfunction, renal failure, diabetes, insulin dependent diabetes, paraplegia, stroke, seizures, collagen vascular disease, coagulopathy, hemophilia, thrombocytopenia, smoking history, postinflammatory pulmonary fibrosis, and Graves disease. † Shoulder, back, knee, other orthopedic, colon resection, stomach, hernia, gall bladder, skin graft, other surgery, and hip. ‡ There were 21 hospital indicators representing 20 hospitals with at least 14 patients and 167 hospitals included as a grouped variable.

Table 3. Correlation between Anesthesia Claim Time and Anesthesia Chart Time by Individual Hospital Data and Aggregated among Hospitals

Hospital	Patients (n)	Kendall's τ Correlation Coefficients	*P* Value for Kendall's τ	Probability of Concordance	Intercept for Individual Hospital (min)	Slope for Individual Hospital (in chart min/bill min)	Median Difference between Claim Anesthesia Minutes and Chart Anesthesia Minutes	*P* Value of Signed Rank Test of Median Difference between Claim Anesthesia Minutes and Chart Anesthesia Minutes
1	34	0.91	< 0.0001	0.95	0.82	0.99	1.5	0.0070
2	34	0.90	< 0.0001	0.95	−0.07	0.99	2.5	0.0011
3	34	0.93	< 0.0001	0.96	0.00	1.00	0	0.5458
4	33	0.84	< 0.0001	0.92	–9.68	0.97	19.5	0.0002
5	31	0.86	< 0.0001	0.93	0.00	1.00	0	0.4764
6	30	1.00	< 0.0001	1.00	0.00	1.00	0	0.3627
7	28	0.89	< 0.0001	0.95	0.75	1.00	–0.75	0.3563
8	27	0.85	< 0.0001	0.93	4.53	0.94	6.0	0.0011
9	25	0.88	< 0.0001	0.94	1.98	0.93	4.5	0.0018
10	24	0.92	< 0.0001	0.96	–7.08	1.01	4.5	0.0050
11	23	0.68	< 0.0001	0.84	–4.38	0.87	19.5	0.0114
12	22	0.86	< 0.0001	0.93	1.11	0.99	0	0.6864
13	22	0.93	< 0.0001	0.97	8.89	0.93	2.0	0.0428
14	22	0.92	< 0.0001	0.96	0.59	0.98	0.75	0.0343
15	21	0.92	< 0.0001	0.96	–2.96	0.99	4.5	0.1238
16	20	0.87	< 0.0001	0.93	–1.75	1.00	1.75	0.0007
17	20	0.91	< 0.0001	0.95	–10.10	1.00	10.0	0.0020
18	19	0.88	< 0.0001	0.94	3.02	0.90	10.0	0.0005
19	18	0.88	< 0.0001	0.94	–2.04	0.98	5.25	0.0002
20	14	0.98	< 0.0001	0.99	–0.69	1.01	0	0.6973
"Other" hospitals	1,430	0.85	< 0.0001	0.92	0.88	0.96	4.75	< 0.0001
All hospitals	1,931	0.85	< 0.0001	0.93	0.85	0.97	4.5	< 0.0001

Hospitals represented in the sample with fewer than 14 patients were included in the "other hospital" category (n = 167).

found a coefficient that was approximately 1. As shown, this relationship was quite stable among hospitals. There are some outliers, that is, individual patients who fall far from the lines; therefore, robust statistical methods that give little weight to outliers are needed when working with bill times.

Estimating Surgical Operative Time

Unlike estimating anesthesia time, in which anesthesia providers submit bills based on minutes of care, we do not have such bills from the surgeon when estimating surgical time. Nevertheless, we chose to explore how well the same anesthesia bill metric, which worked so well at estimating chart anesthesia time, would succeed at estimating surgical chart time. For each analysis in table 4, we let the anesthesia claim time represent the independent variable to predict surgical chart time as abstracted as part of the Surgical Outcome Study. We again present five models: Model I uses only the anesthesia claim time to predict surgical chart time; Model II uses claim time and patient procedures; Model III uses claim time and patient comorbidities; Model IV uses claim time and hospital indicators; and Model V includes all variables. Model I is the simplest model and performs as well as models that include many other predictors. The linear equation to estimate surgical chart time from claim time using Theil's estimate was equal to –23.81 min + 0.82*anesthesia claim time minutes, and using the robust estimate, the equation was –27.63 min + 0.84* anesthesia claim time minutes. The negative constant term, along with a slope less than 1, makes good sense because patients are under anesthesia before and throughout the operation. As shown in table 4, patient characteristics such as comorbid disease, procedure type, and hospital did influence the prediction of chart time after accounting for claim time, but the effect was very small, as evidenced by the similar R^2 for each model.

To better explore the differences in the estimation of chart time by claim time within and among hospitals, we computed Theil's regression models for each hospital individually and among all hospitals. As in table 3, table 5 presents Kendall's τ and the slope and intercept for each hospital with 14 or more patients in the Surgical Outcomes Study and also among all hospitals in the Surgical Outcomes Study. The estimates of chart time using claim time are somewhat less reliable within hospitals for surgery than for anesthesia.

Figure 4 displays the overall data set with Theil's regression line fit through the points. Anesthesia claim times are more similar to anesthesia chart time than is anesthesia claim time to surgical chart time, and there was more variability in the overall association between anesthesia claim and surgical chart time than for anesthesia chart time (compare with fig. 1). Finally, in figure 5, we display individual hospitals with their Theil fitted

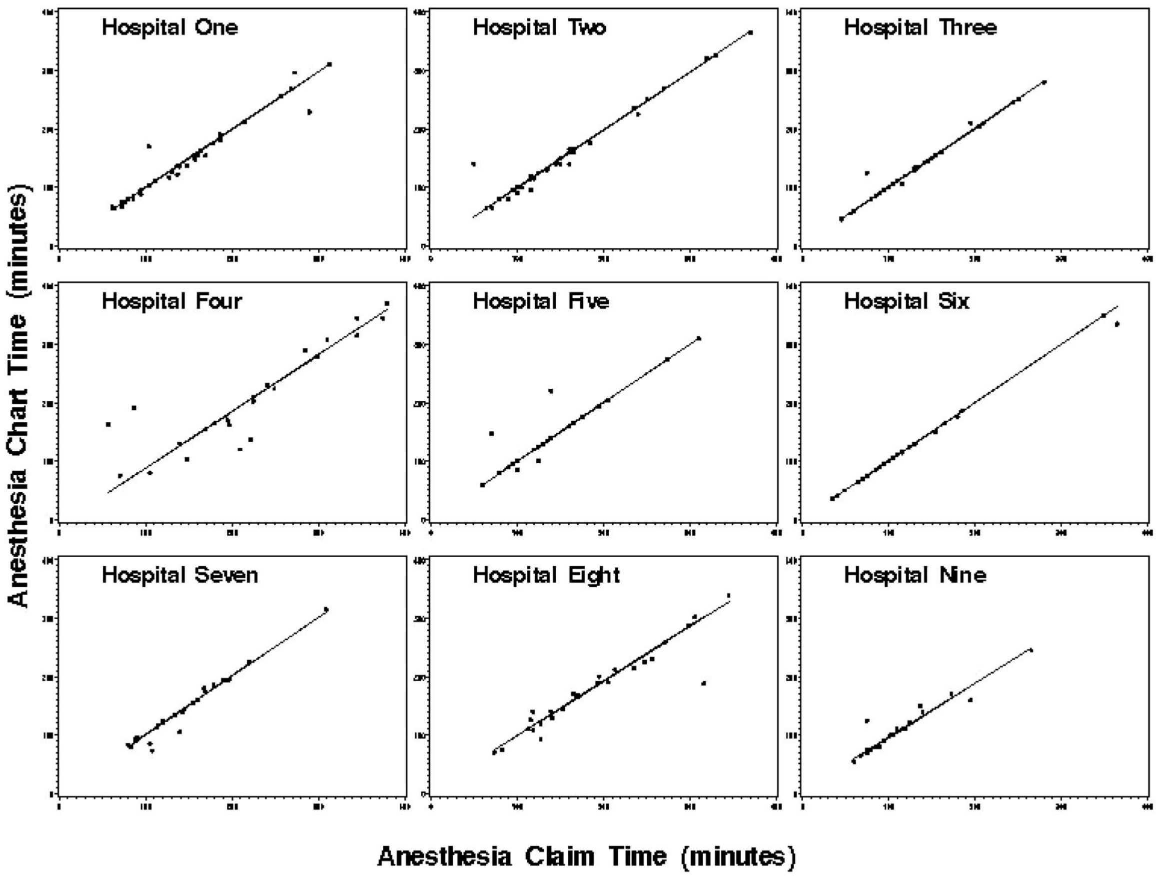

Fig. 3. Theil's regression plot using anesthesia claim time to estimate anesthesia chart time for nine hospitals in the data set with the largest sample of patients.

slopes and intercepts. Again, the relationship between claim time and chart time is quite reliable, although there are occasionally large errors in some patient predictions.

Finally, we explored what percentage of anesthesia time could be explained by surgical time. The correlation between anesthesia chart time and surgical chart time was computed for 1,888 charts that had all four landmark times. The Spearman rank correlation was 0.904. Hence, approximately 82% of the variation in anesthesia time can be explained by surgical time.

Discussion

We found that Medicare anesthesia claims can be used to estimate anesthesia chart time using a simple algorithm that is based on the longest claim in the Medicare

Table 4. Regression Models to Estimate Surgical Chart Time Based on Anesthesia Claim Time and Other Adjustments

Model adjustments	Intercept	Coefficient on Claim (min)	Coefficient *P* Value	Model R^2	Model Median Absolute Residual (error in min)	Contrast *P* Value (Wald test) *vs.* Model I
I. Unadjusted Theil estimation	–23.81	0.82 (0.80–0.83)	< 0.0001	0.78	13.7	—
I. Unadjusted M estimation	–27.63	0.84 (0.84, 0.86)	< 0.0001	0.78	13.3	—
II. Comorbidity*	–25.89	0.84 (0.83, 0.86)	< 0.0001	0.78	13.2	0.13
III. Procedure†	–32.71	0.83 (0.82, 0.84)	< 0.0001	0.79	12.4	< 0.001
IV. Hospital‡	–29.13	0.85 (0.84, 0.86)	< 0.0001	0.78	13.2	< 0.0001
V. Comorbidity, procedure, and hospital indicators	–30.11	0.83 (0.82, 0.84)	< 0.0001	0.80	12.5	< 0.0001

n = 1,888 patients.

* Myocardial infarct, arrhythmia, heart failure, unstable angina, angina, hypertension, aortic stenosis, pulmonic stenosis, chronic obstructive pulmonary disease, asthma, cancer, bad cancer, psychosis, electrolyte and fluid abnormality, alcoholism, liver dysfunction, renal dysfunction, renal failure, diabetes, insulin dependent diabetes, paraplegia, stroke, seizures, collagen vascular disease, coagulopathy, hemophilia, thrombocytopenia, smoking history, post inflammatory pulmonary fibrosis, and graves disease. † Shoulder, back, knee, other orthopedic, colon resection, stomach, hernia, gall bladder, skin graft, other surgery, and hip. ‡ There were 21 hospital indicators representing 20 hospitals with at least 14 patients and 167 hospitals included as a grouped variable.

Table 5. Correlation between Anesthesia Claim Time and Surgical Chart Time by Individual Hospital Data and Aggregated across Hospitals

Hospital	Patients (n)	Kendall's τ Correlation Coefficients	*P* Value for Kendall's τ	Probability of Concordance	Intercept for Individual Hospital (min)	Slope for Individual Hospital (in chart min/bill min)	Median Difference between Claim Anesthesia Minutes and Chart Surgical Minutes	Median Difference between Chart Anesthesia Minutes and Chart Surgical Minutes
1	31	0.9	<0.0001	0.95	−31.82	0.97	34.5	33
2	33	0.76	<0.0001	0.88	−40.00	0.93	50.5	50
3	34	0.66	<0.0001	0.83	−30.08	0.78	60.0	60
4	33	0.8	<0.0001	0.9	−35.96	0.89	70.5	50
5	30	0.74	<0.0001	0.87	−39.16	0.91	51.0	51
6	30	0.64	<0.0001	0.82	−13.00	0.67	40.0	40
7	28	0.78	<0.0001	0.89	−30.39	0.79	53.5	54.5
8	27	0.7	<0.0001	0.85	−31.36	0.79	72.0	63
9	25	0.75	<0.0001	0.875	−24.82	0.77	50.0	45
10	24	0.84	<0.0001	0.92	−32.50	0.83	60.25	54
11	22	0.59	<0.0001	0.795	−0.61	0.64	50.25	27.5
12	22	0.81	<0.0001	0.905	−28.97	0.88	44.25	42
13	22	0.81	<0.0001	0.905	−23.94	0.81	60.75	54.5
14	22	0.78	<0.0001	0.89	−20.42	0.75	46.0	44
15	21	0.82	<0.0001	0.91	−23.60	0.88	40.5	40
16	20	0.75	<0.0001	0.875	−19.27	0.85	40.0	37
17	20	0.82	<0.0001	0.91	−40.10	1.02	38.0	25
18	17	0.86	<0.0001	0.93	−21.27	0.87	40.0	20
19	17	0.75	<0.0001	0.875	−43.75	0.83	69.0	45
20	14	0.75	<0.0001	0.875	−11.67	0.80	34.5	26.5
"Other" hospitals	1,396	0.83	<0.0001	0.915	−22.77	0.81	50.0	41
All hospitals	1,888	0.83	<0.0001	0.915	−23.81	0.82	50.0	43

Hospitals represented in the sample with less than 14 patients were included in the "other hospital" category (n = 167). Anesthesia time around surgery varies considerably among hospitals.

bill. Anesthesia and surgical times are often discussed and analyzed in the anesthesia and surgical literature, but except for the Veterans Affairs studies,[10] few large-scale studies have been reported among institutions. The potential use of easily available claim time variables from Medicare has never been formally tested on a large number of charts before this study. We can now say with some confidence that Medicare claim information, at least in the years 1995 and 1996, provide a useful proxy for anesthesia time as recorded by the anesthesia team in the chart. Given that the median anesthesia time is longer than 130 min in our study, the median absolute errors of approximately 5 min for anesthesia claims predicting anesthesia chart time and approximately 14 min for predicting surgical chart time are quite small. The possible uses of this information are manifold.

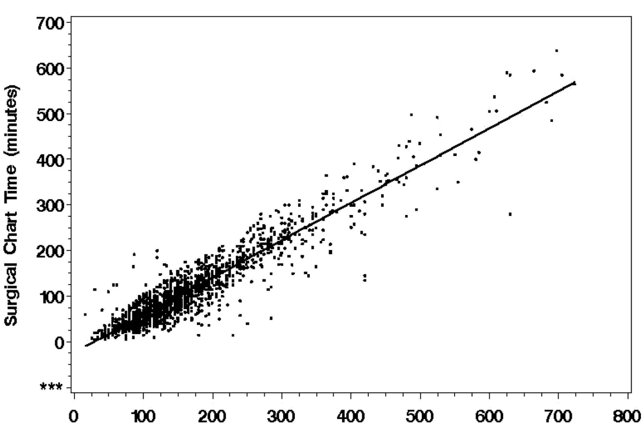

Fig. 4. Theil's regression plot (n = 1,888). The independent variable is anesthesia claim minutes, and the dependent variable is surgical chart minutes. R^2 = 0.78.

In earlier work, Abouleish *et al.*[1] have demonstrated an excellent correlation between minutes as reported from billing departments at four different hospitals and their "surgical time" recorded from the chart (corresponding to our "anesthesia chart time"). They reported a correlation of 0.85 and displayed a figure that shows an apparent slope near 1 relating the two quantities. Our results, using Medicare claims obtained from Medicare and chart review of anesthesia and surgical time, display a similar correlation and slope but are derived from Medicare claims data and not directly obtained from hospital billing departments. As we described earlier, such Medicare claims are readily available through the Centers for Medicare and Medicaid Services as facilitated by the Research Data Assistance Center.#

In our data, we had both anesthesia bill times from Medicare and anesthesia chart times from chart abstraction, whereas in typical applications, only the bill times will be available. In our data, we were able to predict the chart time with good accuracy from the bill time by a simple linear equation. To predict anesthesia chart time,

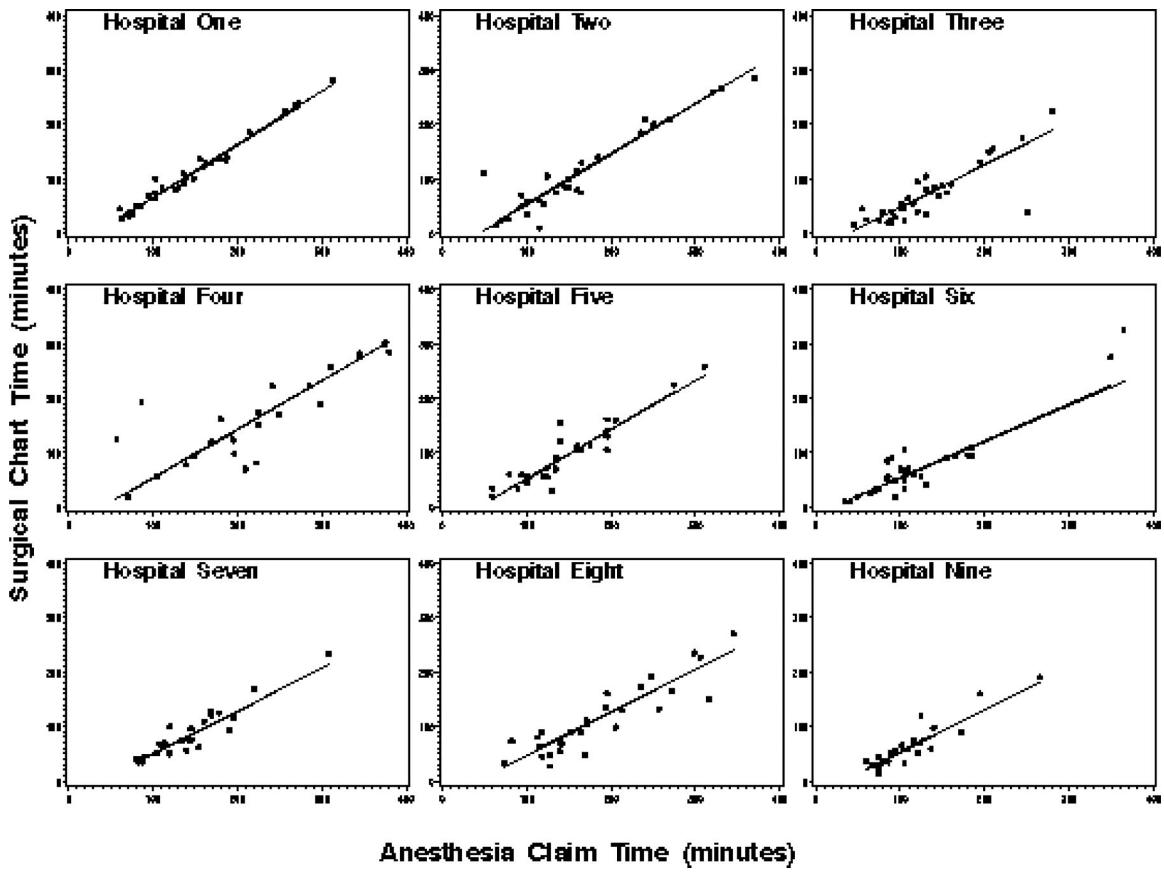

Fig. 5. Theil regression plot using anesthesia claim time to estimate surgical chart time for nine hospitals in the data set with the largest sample of patients.

further adjustments for comorbidities and procedures were not needed; they did not improve the prediction compared with using bill times alone. Two cautions about the use of anesthesia bill times for chart times are important. First, the relationship between chart and bill times had a moderate number of outliers, that is, isolated patients with incommensurate chart and bill times. For this reason, when working with bill times in place of chart times, one should use robust statistical procedures, such as m-estimation,[28,33,34] that limit the influence of individual patients. In an analysis of thousands of bill times from Medicare, it is very likely that one or two patients will have bizarre bill times, which convert to bizarre estimated chart times, even though there was nothing bizarre in the underlying clinical practice. Robust statistical methods will give limited weight to these few bizarre times. Second, we did find some differences among hospitals in the relationship between chart and bill times: these differences were only moderate in size or clinical significance, but they were often statistically significant. This has two consequences. First, if one finds a clinically small but statistically significant difference between two hospitals in their chart times predicted from their bill times, one should not jump to the conclusion that there is a corresponding difference in clinical practice at the two hospitals; it may instead reflect a small but highly systematic difference in billing practices at the two hospitals. Second, in studying patient or clinical issues related to anesthesia time, it is safest to incorporate into the analysis adjustments for hospitals, such as matching patients from the same hospital, stratification on the hospital, suitable modeling, or combinations of these.

In summary, noting these caveats, it seems that the study of Medicare anesthesia claims has the potential to provide new insights into the practice of anesthesiology and surgery and variations among hospitals, providers, and their patients undergoing surgical procedures.

We thank William J. Greeley, M.D., M.B.A. (Anesthesiologist-in-Chief, Department of Anesthesiology and Critical Care Medicine, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania), and Lee Fleisher, M.D. (Chair, Department of Anesthesiology and Critical Care Medicine, The University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania), for their aid in conducting this research. We thank the anonymous reviewers for their insightful comments and suggestions regarding this manuscript. We also thank Traci Frank (Administrative Coordinator) and Andrea Millman, B.A. (Research Assistant), from the Center for Outcomes Research, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, for aid in preparing this manuscript.

References

1. Abouleish AE, Prough DS, Zornow MH, Hughes J, Whitten CW, Conlay LA, Abate JJ, Horn TE: The impact of longer-than-average anesthesia times on the billing of academic anesthesiology departments. Anesth Analg 2001; 93:1537–43
2. Khuri SF, Najjar SF, Daley J, Krasnicka B, Hossain M, Henderson WG, Aust

JB, Bass B, Bishop MJ, Demakis J, DePalma R, Fabri PJ, Fink A, Gibbs J, Grover F, Hammermeister K, McDonald G, Neumayer L, Roswell RH, Spencer J, Turnage RH: Comparison of surgical outcomes between teaching and nonteaching hospitals in the Department of Veterans Affairs. Ann Surg 2001; 234:370–82

3. Schuster M, Standl T, Reissmann H, Kuntz L, Am Esch JS: Reduction of anesthesia process times after the introduction of an internal transfer pricing system for anesthesia services. Anesth Analg 2005; 101:187–94

4. Sandberg WS, Daily B, Egan M, Stahl JE, Goldman JM, Wiklund RA, Rattner D: Deliberate perioperative systems design improves operating room throughput. Anesthesiology 2005; 103:406–18

5. Torkki PM, Marjamaa RA, Torkki MI, Kallio PE, Kirvela OA: Use of anesthesia induction rooms can increase the number of urgent orthopedic cases completed within 7 hours. Anesthesiology 2005; 103:401–5

6. Brunicardi FC, Andersen DK, Billiar TR, Dunn DL, Hunter JG, Pollock RE: Schwartz's Principles of Surgery, 8th edition. New York, McGraw-Hill, 1994.

7. Haley RW, Culver DH, Morgan WM, White JW, Emori TG, Hooton TM: Identifying patients at high risk of surgical wound infection: A simple multivariate index of patient susceptibility and wound contamination. Am J Epidemiol 1985; 121:206–15

8. Culver DH, Horan TC, Gaynes RP, Martone WJ, Jarvis WR, Emori TG, Banerjee SN, Edwards JR, Tolson JS, Henderson TS: Surgical wound infection rates by wound class, operative procedure, and patient risk index. National Nosocomial Infections Surveillance System. Am J Med 1991;91:152S–7S

9. Scott CF Jr: Length of operation and morbidity: Is there a relationship? Plast Reconstr Surg 1982;69:1017–21.

10. Grossmann EM, Longo WE, Virgo KS, Johnson FE, Oprian CA, Henderson W, Daley J, Khuri SF: Morbidity and mortality of gastrectomy for cancer in Department of Veterans Affairs Medical Centers. Surgery 2002; 131:484–90

11. Dexter F, Coffin S, Tinker JH: Decreases in anesthesia-controlled time cannot permit one additional surgical operation to be reliably scheduled during the workday. Anesth Analg 1995; 81:1263–8

12. Dexter F: A strategy to decide whether to move the last case of the day in an operating room to another empty operating room to decrease overtime labor costs. Anesth Analg 2000; 91:925–8

13. Abouleish AE, Prough DS, Vadhera RB: Influence of the type of anesthesia provider on costs of labor analgesia to the Texas Medicaid program. Anesthesiology 2004; 101:991–8

14. Abouleish AE, Dexter F, Whitten CW, Zavaleta JR, Prough DS: Quantifying net staffing costs due to longer-than-average surgical case durations. Anesthesiology 2004; 100:403–12

15. Watcha MF, White PF: Economics of anesthetic practice. Anesthesiology 1997; 86:1170–96

16. Silber JH, Rosenbaum PR, Trudeau ME, Even-Shoshan O, Chen W, Zhang X, Mosher RE: Multivariate matching and bias reduction in the surgical outcomes study. Med Care 2001; 39:1048–64

17. Silber JH, Rosenbaum PR, Trudeau ME, Chen W, Zhang X, Rapaport Kelz R, Mosher RE, Even-Shoshan O: Changes in prognosis after the first postoperative complication. Med Care 2005; 43:122–31

18. Silber JH, Rosenbaum PR, Trudeau ME, Chen W, Zhang X, Lorch SA, Kelz RR, Mosher RE, Even-Shoshan O: Preoperative antibiotics and mortality in the elderly. Ann Surg 2005; 242:107–14

19. US Department of Health and Human Services: Statistical Supplement 1997. Health Care Finance Rev 1997:31.

20. US Department of Health and Human Services: Statistical Supplement 1998. Health Care Finance Rev 1998:177.

21. Steen PM, Brewster AC, Bradbury RC, Estabrook E, Young JA: Predicted probabilities of hospital death as a measure of admission severity of illness. Inquiry 1993; 30:128–41

22. Rosenbaum PR, Silber JH: Matching and thick description in observational study of mortality after surgery. Biostatistics 2001; 2:217–32

23. Ming K, Rosenbaum PR: A note on optimal matching with variable controls using the assignment algorithm. J Comput Graph Stat 2001; 10:455–63

24. Ming K, Rosenbaum PR: Substantial gains in bias reduction from matching with a variable number of controls. Biometrics 2000; 56:118–24

25. Williams BA, Kentor ML, Williams JP, Figallo CM, Sigl JC, Anders JW, Bear TC, Tullock WC, Bennett CH, Harner CD, Fu FH: Process analysis in outpatient knee surgery: Effects of regional and general anesthesia on anesthesia-controlled time. Anesthesiology 2000; 93:529–38

26. Hollander M, Wolfe DA: Nonparametric Statistical Methods, 2nd edition. New York, John Wiley & Sons, 1999.

27. Hollander M, Wolfe DA: Regression problems, Nonparametric Statistical Methods. New York, John Wiley & Sons, 1973, pp 421–8.

28. Huber PJ: The basic types of estimates, Robust Statistics. New York, John Wiley & Sons, 1981, pp 43–55.

29. Hampel FR, Ronchett EM, Rousseeuw PJ, Stahel WA: Linear models: Robust estimation, Robust Statistics: The Approach Based on Influence Functions. New York, John Wiley & Sons, 1986, pp 315–28.

30. Huber PJ: Robust regression: Asymptotics, conjectures and the Monte Carlo. Ann Stat 1973; 1:799–821

31. Hampel FR, Ronchett EM, Rousseeuw PJ, Stahel WA: One-dimensional estimators, Robust Statistics: The Approach Based on Influence Functions. New York, John Wiley & Sons, 1986, pp 100–7.

32. Hampel FR, Ronchett EM, Rousseeuw PJ, Stahel WA: Linear models: Robust testing, Robust Statistics: The Approach Based on Influence Functions. New York, John Wiley & Sons, 1986, pp 364–6.

33. SAS Institute Inc. : The ROBUSTREG Procedure. SAS/STAT User's Guide, Version 9.1.3, Cary, NC, SAS Institute, Inc., 2003.

34. Hampel FR, Ronchett EM, Rousseeuw PJ, Stahel WA: Robust Statistics, The Approach Based on Influence Functions. New York, John Wiley & Sons, 1986.