Using Average Cost Methods to Estimate Encounter-Level Costs for Medical-Surgical Stays in the VA

Todd H. Wagner
VA HSR&D Health Economics Resource
Center and Stanford University
Shuo Chen
VA HSR&D Health Economics Resource Center
Paul G. Barnett
VA HSR&D Health Economics Resource
Center and Stanford University

The U.S. Department of Veterans Affairs (VA) maintains discharge abstracts, but these do not include cost information. This article describes the methods the authors used to estimate the costs of VA medical-surgical hospitalizations in fiscal years 1998 to 2000. They estimated a cost regression with 1996 Medicare data restricted to veterans receiving VA care in an earlier year. The regression accounted for approximately 74 percent of the variance in cost-adjusted charges, and it proved to be robust to outliers and the year of input data. The beta coefficients from the cost regression were used to impute costs of VA medical-surgical hospital discharges. The estimated aggregate costs were reconciled with VA budget allocations. In addition to the direct medical costs, their cost estimates include indirect costs and physician services; both of these were allocated in proportion to direct costs. They discuss the method's limitations and application in other health care systems.

Keywords: *short-stay hospitalization; cost; charges; expenditures*

The U.S. Department of Veterans Affairs (VA) maintains centralized databases containing detailed hospital discharge abstracts, but encounter-level charge or cost information has not been readily available for cost and outcome

Medical Care Research and Review, Vol. 60 No. 3, (Supplement to September 2003) 15S-36S
DOI: 10.1177/1077558703256485

analyses. This lack of data reflects the fact that VA interacts with third party payers for only a tiny percentage of the care it provides, and therefore it does not routinely generate patient bills.

As described by Barnett and Wagner (2003 [this issue]), one objective of the VA Health Economics Resource Center (HERC) has been to create patient-level cost estimates. Before HERC, researchers estimated VA costs as needed. Since no standard research methodology was in place, many of the cost estimates were not comparable (Chapko, Ehreth, and Hedrick 1991). Some progress was made by linking department-level cost and utilization data to estimate average daily rates for inpatient care (Barnett, Chen, and Wagner 2000). However, using average daily rates for medical or surgical discharges makes extreme assumptions that are generally not valid. For example, this approach assumes that appendectomies and heart transplants with the same length of stay (LOS) had equal costs. Recently, Barnett (1997) used a regression to estimate an individual's cost as a function of the deviation from a medical center's average. One problem with this approach was that the lack of institutional-level variation made it difficult to estimate precisely individual-level costs.

This article describes HERC's method for estimating the cost of VA health care encounters in fiscal year (FY) 1998 to FY 2000. Our goal was to develop a database of long-run national average costs. Intended for cost-effectiveness analysis, these data do not account for hospital market factors, nor were they designed to capture short-run fixed costs. These caveats and limitations are described in the methods and the discussion sections, but they are critical for using the data appropriately.

NEW CONTRIBUTION

Cost data are missing from VA utilization databases. In the past, researchers wanting to conduct cost-effectiveness analysis first had to estimate encounter-level costs. We use regression models to estimate the cost of inpatient medical-surgical discharges for FY 1998 to FY 2000. This method assumes that every encounter has the average cost of all encounters that share the same discharge characteristics. The cost regression exploits variation in major diagnostic category (MDC), diagnosis related group (DRG), LOS, number of diagnoses, inpatient death, sex, age, and number of intensive care unit (ICU) days. It captures

This research was supported in part by the Department of Veterans Affairs, Veterans Health Administration, and Health Services Research and Development Service (SDR-ECN-99017) and in part by the VA Cooperative Studies Program. The views expressed in this article are strictly those of the authors and do not necessarily reflect the Department of Veterans Affairs. We would like to thank Frank Lynn, Terri Menke, Ann Hendricks, Douglas Bradham, Denise Hynes, and two anonymous referees for their helpful comments.

resource variation that would be missed by using Medicare reimbursement rates based on the stay's DRG to calculate costs. The functional form allows for nonlinearities and interactions, and the final model accounts for approximately three-quarters of the variance in cost-adjusted charges. These cost estimates are in a VA database that can be merged to VA utilization records.

METHOD

OBJECTIVE AND DEFINITION OF COST

Our goal was to create an encounter-level cost database for VA medical and surgical inpatient care. More specifically, we estimated long-run national average costs, and in doing so, we treated all short-run fixed costs (e.g., capital) as long-run variable costs. We distributed VA fixed costs in proportion to VA variable costs, based on costs reported in the VA Cost Distribution Report (CDR). In addition, we did not take market-level forces or market-level input prices into account. Although market-level forces, defined as either variation in wages or availability of technology (Baker 1997; Baker and Corts 1996), can affect the supply curve, we were interested in estimating VA average costs for the nation.

Our methodological approach involved developing a cost regression for medical and surgical inpatient stays with Medicare data and using the regression coefficients to impute costs for VA inpatient encounters. Through the regression model, we estimated the relationship between cost-adjusted charges, the dependent variable, and diagnostic and demographic information, the independent variables. The beta coefficients from the regression model were then used to impute "costs" in the VA data set. All stays with the same diagnostic and demographic information were assigned an average cost, also known as a gross cost (Gold et al. 1996). Therefore, the fidelity of the cost regression was determined, in part, by the model's specifications and the independent variables.

We reconciled the costs from the regression model with the VA budget allocation. The VA has a national budget, and each local medical center has a budget. We reconciled to both, generating separate national and local estimates. We strongly encourage researchers to use the national estimates. By construction, the averages of the national and the local estimates are the same, but the local estimates have more variation and larger tails as evidenced by the ratio of local to national estimates, which ranged from 0.62 to 11.64 in FY 1999. Researchers may choose to use the local estimates to see if results hold. However, as mentioned above, the local estimates do not account for market-level effects. Therefore, differences in local estimates may reflect budget allocations, rather than input prices or the relative efficiency of production.

COST REGRESSION

We developed our cost regression using a subset of the 1996 Medicare Provider Analysis Review (MEDPAR) file (Centers for Medicare and Medicaid Services [CMS] 2003). Although we could have used a random sample of the 1996 MEDPAR file, we chose to use a subset of veterans with the assumption that this placed a greater weight on clinical and demographic factors that are VA relevant. A group of VA researchers has identified a cohort of Medicare enrollees who are also enrolled in VA (Wright, Hossain, and Petersen 2000; Wright, Lamkin, and Petersen 2000). The cohort contains all veterans who were users of either inpatient or outpatient VA services between 1992 and 1994 and who had their 65th birthday in 1994. The file had 372,046 hospital stays. From this cohort, we focused on hospitalizations in the continental United States. We also excluded claims for MDC 15 (i.e., newborns and other neonates with conditions originating in the perinatal period), as VA did not cover these services before 2001.

The MEDPAR data set includes a variable for total charges. Given that total charges are often greater than costs, we used the Medicare Cost Report to calculate each hospital's total cost to total charge ratio. The MEDPAR includes a hospital identifier that can be merged with the Medicare Cost Report. After linking the hospital-level cost to charge ratio to the MEDPAR data set, we were able to adjust patient-level total charges with a hospital-specific ratio of cost to charges. In our merged data set, the average of the cost-to-charge ratio was 0.60. Therefore, this adjustment tends to deflate the costs. In addition, it removes hospital-specific cost or accounting idiosyncrasies. In the cost regression, we used cost-adjusted charges as the dependent variable. We estimated the regression using ordinary least squares (OLS). Alternative models with logged costs are described in the sensitivity analysis.

We restricted our choice of independent variables to those available in both the MEDPAR and VA databases. Past literature guided our selection of independent variables (Barnett 1997). To account for resource use, we used the DRG, merged to the 1996 DRG weight file from the CMS. DRG weights are resource-based relative value weights publicly available on CMS's Web site. We captured additional variation in resource use by adding LOS as a positive integer. We also included the difference between the actual LOS and the expected LOS for a given DRG. In effect, this acts as an interaction between DRG and LOS. To allow for nonlinearities, positive and negative deviations in the actual and expected LOS were allowed to vary independently and in a nonlinear fashion (i.e., squared and cubic terms were included). In addition, we interacted the medicine MDC and surgery MDC with LOS.

VA DATA

After estimating the cost regression with MEDPAR data, we used the beta coefficients to impute VA costs. To impute meaningful estimates, we reorganized the VA data to have an equivalent structure to the MEDPAR data. A VA discharge record can include long-term care, rehabilitation, specialty substance abuse and psychiatric treatment, intermediate medicine, and domiciliary care. Many of these non-medical-surgical stays would be treated as separate stays and excluded altogether from the MEDPAR database, which includes inpatient care from short-stay hospitals. This article covers only the cost of medical-surgical care. The method for estimating the costs of rehabilitation, mental health, and long-term care is handled elsewhere in this issue (Yu et al. 2003 [this issue]).

To make a VA medical-surgical discharge data set analogous to the MEDPAR database, we worked with the VA bedsection file. A bedsection is similar to a non-VA hospital ward or department. We adopted the rule that transfers between medical-surgical bedsections were part of the same stay. If a person was transferred from a medical-surgical bedsection to a non-medical-surgical bedsection, we ruled that the medical-surgical stay ended. For example, a transfer from a medical-surgical bedsection to a non-medical-surgical bedsection and back to a medical-surgical bedsection would yield one non-medical-surgical and two medical-surgical discharge records. While combining transfers within contiguous medical-surgical bedsections (i.e., bedsection stays in which the discharge and admission dates were the same), we tracked both overall LOS and days in the ICU. Each bedsection record has an associated DRG. We merged the DRG to the CMS DRG weight public use file, and the DRG with the highest weight was retained under the assumption that this DRG more closely reflected costs and would be used to maximize payment in the non-VA sector.

POSTESTIMATION FIXES

After estimating VA costs with the cost regression, 3,032 (0.7 percent) of the 455,926 medical-surgical hospitalizations had negative costs. This result was an artifact of using a linear regression model. Rare combinations of right-hand-side variables lead to negative predictions. Although negative costs present a clear estimation problem, other cases had implausibly low costs. Forty-two hospital stays had positive costs less than $5. We decided to set a floor for the estimated discharge cost. The floor was established by using the regression model (see Table 1) to simulate the cost of staying an additional day. All other factors being equal, if a person stayed an additional day, MEDPAR cost-adjusted charges increased by an average of $684.75. A total of 9,609 (2 percent) VA stays had costs less than $684.75, and 86 percent of these cases had a 1-day LOS. These cases were all given $684.75.

RELIABILITY AND VALIDITY OF THE COST REGRESSION

We tested the model's validity and robustness using three procedures. First, we identified outliers and reestimated the model after removing approximately 1 percent, 2 percent, and 7 percent of the most influential outliers. We empirically identified outliers using Cook's distance after estimating the cost regression with the Medicare data. Conceptually, Cook's distance is an *F* test comparing the beta coefficients from the full data set to the beta coefficients from the data set excluding the one case (Cook and Weisberg 1982).

We then tested the model's fit by separating Medicare data into quartiles. Within each quartile, Medicare cost-adjusted charges and estimated costs were compared using Pearson's correlation coefficient. Finally, we also tested the model's fit using different 50 percent random samples of the Medicare data. The cost regression was estimated with one of the random samples, and then predicted costs were estimated for the other half of the sample, allowing an out-of-sample comparison of estimated costs and cost-adjusted charges.

We tested whether the model was highly dependent on these data and whether the estimated costs changed significantly if Medicare data from another year were used. Using 1994 and 1995 MEDPAR data for veterans who received VA care, we estimated the same regression model. The beta coefficients from these three models were compared. The 1994, 1995, and 1996 cost regressions were also used to predict 1996 Medicare costs. This allowed us to test the reliability of the cost regression, using actual 1996 cost-adjusted charges as the criterion.

We compared estimated VA costs in the different MDCs, stratified by whether the DRG was surgical or medical. The costs were then ranked from least expensive MDC to most expensive. This ranking was done for MEDPAR, as well as the Healthcare Cost and Utilization Project (HCUP) data. The HCUP data set is a nationally representative discharge data set, based on people of all ages. The HCUP data were used to verify that the method and cost estimate could be used in other circumstances. Rather than comparing the relative VA, Medicare, and HCUP costs, we compared the relative rankings of each MDC across the data sets. We did not want to directly compare costs given that they represent different years and that Medicare and HCUP include different cost

TABLE 1 Ordinary Least Squares Regression Model Estimating Discharge Cost-Adjusted Charges

Characteristic	*Beta Coefficient*	t *Statistic*
Died in hospital	2,671.211**	46.69
Sex (female = 1, male = 0)	32.909	0.54
Age in years	–34.223**	18.48
Number of diagnoses	619.044**	7.63
Number of diagnoses squared	–146.702**	8.83
Number of diagnoses cubed	10.975**	10.73
Length of stay (LOS) in days	104.255**	11.48
Positive deviation from DRG-specific average LOS (POSLOS)	670.950**	66.39
Negative deviation from DRG-specific average LOS (NEGLOS)	182.499**	6.15
NEGLOS squared	–109.890**	13.77
POSLOS squared	–0.717**	32.99
NEGLOS cubed	–4.588**	8.36
POSLOS cubed	0.000006	0.17
1996 DRG weight	4,860.036**	76.30
DRG weight squared	–255.164**	23.11
DRG weight cubed	12.973**	25.65
Surgical MDC	1,069.883**	13.68
Surgical MDC* LOS	–42.315**	3.79
Surgical MDC* POSLOS	421.532**	26.99
Surgical MDC* NEGLOS	328.304**	9.06
Surgical MDC* POSLOS squared	–1.384**	7.72
Surgical MDC* POSLOS cubed	0.001	1.74
Surgical MDC* NEGLOS squared	47.498**	5.64
Surgical MDC* NEGLOS cubed	3.637**	6.59
Days in ICU	593.037**	82.76
ICU days squared	10.274**	37.86
ICU days cubed	–0.033**	18.24
Constant	413.766*	2.28

Observations = 321,583
R^2 = .74

Note: DRG = diagnosis related group; MDC = major diagnostic category; ICU = intensive care unit.
*Significant at .05. **Significant at .01.

components compared to the VA. For example, Medicare and HCUP exclude physician services, whereas they include capital financing and malpractice.

Comparing the rankings provided a measure of agreement. To assess the statistical significance of the agreement, Kappa statistics were calculated.

OBSERVATION DAYS

Beginning in 1997, VA created seven new codes to report inpatient care provided in observation units. An observation bed stay is less intensive than a medical-surgical stay, and it does not have an associated DRG. This prevented us from including these data in the cost regression. We decided to assign each observation day at the marginal cost of an additional day in a nonobservation bedsection ($684.75). This estimate was calculated by using the regression model presented in Table 1 to estimate the additional amount that would have been incurred if the patient stayed one more day.

RECONCILING ESTIMATED AND ACTUAL VA COSTS

The VA tracks department-level expenditures in the CDR. VA expenditures are recorded in the Financial Management System (FMS). The CDR is created by distributing costs reported in the FMS to cost distribution accounts of the CDR. The distribution of costs is based on estimates prepared by the service chiefs in each medical center. At the end of each FY, a cumulative CDR is prepared, and it is reconciled to the costs reported in FMS. We adjusted our estimates so that the sum of both the national and local estimates was equal to the VA medical-surgical budget allocation reported in the CDR.

The CDR includes most VA health care costs, including the cost of physicians. We distributed physician costs across inpatient stays in proportion to facility costs. The CDR tracks capital depreciation, but it lacks information on the cost of capital financing. The CDR also lacks information on malpractice expenses. Both of these costs are covered by other federal agencies. Therefore, our cost estimates lack these two components.

When tallying the CDR costs, we excluded costs for contract care, home care programs, and benefits included in the medical or surgical cost distribution accounts because the corresponding services are often not captured in the utilization databases. We also excluded the cost of 16 facilities that do not provide patient care. These 16 sites provide central administration, which may involve activities that are more characteristic of a health care payer, rather than a health care provider. We included indirect costs by assigning them to each department in proportion to the department's share of direct costs.

Two hurdles arose when we merged the VA utilization data to the CDR cost data. First, we had to account for VA medical center mergers. If VA medical centers merged during an FY, we merged their utilization and cost data for the entire FY. It was not possible to separate accurately costs and utilization before and after the merger.

The second hurdle was that the utilization data reported all discharges ending in the FY. The data set includes stays that began in prior FYs but not those stays that end in subsequent years. In contrast, the CDR reports costs for an FY, including costs for patients not yet discharged. Ignoring this difference would be equivalent to assuming that bed occupancy is constant over time. There is a trend in VA to shorten LOS and to reduce hospitalizations. Consequently, the estimated cost of discharges that began in earlier years would be too large in current-year dollars given that current-year dollars are being spread over fewer patients each year. To adjust the dollars to more closely reflect the discharge view of the utilization data, we calculated the percentage of beds full at the end of the FY compared to the beginning of the year (0.93 for FY 1998, 0.98 for FY 1999, and 0.93 for FY 2000). We used this ratio to deflate the estimated costs for stays that started in prior FYs.

After accounting for mergers and adjusting the estimated costs to the FY, we reconciled the estimates to VA budget allocations. The reconciliation with the VA medical center produced a local cost estimate, whereas reconciling to the entire VA produced a national cost estimate. By construction, the averages of the national and the local estimates are the same, but the national and local estimates differ for any one encounter. The latter may reflect differences in input prices, but it may also reflect different accounting practices.

RESULTS

COST REGRESSION

The cost regression is presented in Table 1. The regression model is parsimonious in that it only used eight discharge descriptors, yet the model allows for interactions and nonlinearities between important variables including LOS and DRG weight. The final model accounted for almost three-quarters of the variance among veterans who used Medicare, and it was highly significant ($F_{27,321,555} = 33396.7$, $p < .0001$).

RELIABILITY AND VALIDITY OF ESTIMATED VA COSTS

We estimated the cost regression with Medicare data, saving Cook's distance. We then ran alternative models, removing an increasing percentage of outliers from the sample. The results indicated that the model's overall R^2 did not increase substantially when eliminating outliers. In fact, when we eliminated the top 1 percent of outliers, the model's R^2 decreased.

We also separated the sample into quartiles according to cost-adjusted charges. Again, eliminating the outliers did not universally improve the model's fit among the quartiles. Table 2 shows Pearson's correlation coefficients between the estimated costs and the Medicare cost-adjusted charges. Given these results, we concluded that omitting outliers would be based on an arbitrary limit, which could lead to a worse fit. Alternatively, one could identify outliers according to Medicare's outlier designator, but eliminating these cases had little effect.

An unexpected finding was that the regression model fit the large costs (quartile 4) considerably better than the low costs. Efforts to improve the model's fit with the low-cost observations often exacerbated the fit in the high-cost cases and increased the model's absolute error, measured as the difference between cost-adjusted charges and estimated costs. A semilog model, which is often used for skewed cost data, produced estimates that were weakly correlated (.106, see Table 2) with cost-adjusted charges.

The split-sample analysis confirmed the robustness of the model. As Table 2 shows, when we estimated the cost regression with a randomly selected half of the data and predicted the costs in the other half, the correlation between cost-adjusted charges and estimated costs remained consistent across quartiles.

The estimated costs were robust to the input data. Simulated VA costs using 1994, 1995, and 1996 MEDPAR data were correlated above .99. To compare estimated costs to cost-adjusted charges, we used the cost regression with 1994, 1995, and 1996 MEDPAR data restricted to veterans who had used VA services to estimate costs for the 1996 MEDPAR data. We were then able to compare estimated costs to the 1996 Medicare cost-adjusted charges, using the latter as the reference. Again the models were adept at estimating costs. Table 3 shows the correlations between the 1996 cost-adjusted charges and the estimated costs.

After estimating VA costs, we divided the sample into surgical and medical DRGs. We ranked the MDCs according to the average VA cost. We then ranked Medicare and HCUP costs in the same way. Agreement of ranks, as estimated using Kappa statistics, within the surgical and medical DRGs was statistically significant with p-values at or below .001. The agreement was slightly higher for the surgical DRGs than for the medical DRGs. Tables 4 and 5 show the rankings and the average costs for each medical and surgical MDC, respectively.

SENSITIVITY ANALYSIS

In the cost regression, we used the 1996 MEDPAR file restricted to veterans who received VA care. Other data sets, such as HCUP, could be used to estimate the cost regression. Both the MEDPAR and HCUP data sets report

TABLE 2 Correlation Coefficients of MEDPAR Cost-Adjusted Charges (CAC) and Estimated Costs by Quartile

	Quartile 1: CAC < $2,605		*Quartile 2: $2,605 < CAC < $4,484*		*Quartile 3: $4,484 < CAC < $8,472*		*Quartile 4: CAC > $8,472*	
	In Sample	*Out of Sample*	*In Sample*	*Out of Sample*	*In Sample*	*Out of Sample*	*In Sample*	*Out of Sample*
Correlation coefficients								
Full models								
ESTCOST	.126	.190	.301	.291	.389	.357	.814	.809
LGCOST	.083	.109	.303	.290	.390	.381	.389	.106
Outlier omitted models								
1.	.057	.204	.309	.005	.396	.250	.641	.699
2.	.071	.209	.313	.011	.398	.279	.718	.750
3.	.185	.202	.313	.305	.393	.392	.769	.775
Sample size	38,304	38,144	39,167	38,594	39,939	40,801	43,348	43,286

Note: MEDPAR = Medicare Provider Analysis Review; ESTCOST = estimated cost from the linear cost regression; LGCOST = estimated cost from the logged cost regression. In model 1, the cost regression excluded cases with Cook's distance greater than .001 (least restrictive). In model 2, the cost regression excluded cases with Cook's distance greater than .0001 (more restrictive). In model 3, the cost regression excluded cases with Cook's distance greater than .00001 (most restrictive).

TABLE 3 Correlations in Estimated Costs Compared to 1996 Cost-Adjusted Charges

	Model 1	*Model 2*	*Model 3*
MEDPAR cost-adjusted charges (1996)	.856	.855	.859
Model 1	1.000		
Model 2	.993	1.000	
Model 3	.997	.996	1.000

Note: MEDPAR = Medicare Provider Analysis Review. Model 1 is the 1994 MEDPAR model estimating costs for 1996 data. Model 2 is the 1995 MEDPAR model estimating costs for 1996 data. Model 3 is the 1996 MEDPAR model estimating costs for 1996 data.

medical-surgical inpatient discharges, but MEDPAR is limited to Medicare enrollees older than the age of 65 and those younger than the age of 65 with a disability, while HCUP includes people of all ages. MEDPAR reports the number of days spent in the ICU, which is not captured by HCUP.

In deciding whether the cost regression should be estimated with MEDPAR or HCUP data, we ran a number of regressions with the MEDPAR and HCUP data and looked at model fit and absolute mean error. Including ICU days increased the R^2 from approximately .69 to .74. In comparison to ICU days, age is a relatively poor predictor of hospital costs and adds very little to the model's fit (R^2 increased .002). The absolute mean error between Medicare cost-adjusted charges and imputed costs was $2,662 and $2,825 for the models including and excluding ICU days, respectively. When age was excluded from the MEDPAR cost regression, the absolute mean error increased by $0.21. The age effect was statistically significant, but the results suggest that after controlling for other variables, it is more important to be able to adjust for ICU days than to include a wider distribution of ages. Consequently, we chose to estimate the cost regression using the MEDPAR data.

The cost regression's dependent variable was cost-adjusted charges. The unadjusted cost-adjusted charges were highly right skewed, raising questions about the appropriateness of using OLS regression. We used the log transform with the smearing estimator (Duan 1983) to test whether this produced a model with a better fit and whether this reduced error in the residuals. Although the log transform helped reduce the appearance of skewness, the nonlogged cost regression consistently performed better than models with logged cost-adjusted charges. OLS models had a higher R^2 than the semilog model. In addition, when we used a randomly selected half of the MEDPAR sample to predict with OLS and semilog models the cost of the other half, the OLS models consistently had substantially lower absolute mean error (i.e.,

TABLE 4 Ranks of Medical Major Diagnostic Categories from Least to Most Expensive

	VA (1998)			MEDPAR (1996)			HCUP (1996)					
Major Diagnostic Category	n	*Average Cost*	*Average DRG Weight*	n	*Average Cost*	*Average DRG Weight*	n	*Average Cost*	*Average DRG Weight*	*VA Rank*	*MEDPAR Rank*	*HCUP Rank*
14. Pregnancy and childbirth	13	2,874	0.609	3	961	0.383	6,619	2,084	0.437	1	1	1
23. Factors influencing health status	7,080	3,391	0.695	7,226	9,620	1.255	915	7,710	1.172	2	20	20
13. Female reproductive system	190	3,565	0.742	19	4,379	1.002	154	3,371	0.865	3	7	3
2. Diseases of the eye	1,127	3,584	0.634	162	3,910	0.669	57	3,581	0.654	4	3	5
9. Skin, subcutaneous tissue, and breast	11,862	4,282	0.743	3,663	4,604	0.793	811	3,920	0.732	5	9	7
21. Injuries, poisonings	4,994	4,294	0.714	1,891	4,238	0.804	601	3,497	0.703	6	6	4
3. Ear, nose, mouth, and throat	5,082	4,620	0.816	1,567	3,539	0.677	310	3,060	0.663	7	2	2
16. Blood and Immunological disorders	5,444	4,730	0.934	2,082	5,681	0.988	453	5,682	0.944	8	13	15
10. Endocrine, nutritional, and metabolic	15,499	4,749	0.776	7,651	4,095	0.816	1,301	4,150	0.789	9	4	8
19. Mental diseases and disorders	4,465	4,841	0.767	10,119	7,413	0.784	2,283	4,943	0.758	10	18	13
8. Musculoskeletal system	11,139	4,945	0.767	6,768	5,504	0.775	1,395	4,592	0.757	11	12	11
12. Male reproductive system	2,778	5,015	0.773	902	4,175	0.782	57	3,640	0.735	12	5	6

(continued)

TABLE 4 (continued)

Major Diagnostic Category	*VA (1998)* n	*VA (1998)* *Average Cost*	*VA (1998)* *Average DRG Weight*	*MEDPAR (1996)* n	*MEDPAR (1996)* *Average Cost*	*MEDPAR (1996)* *Average DRG Weight*	*HCUP (1996)* n	*HCUP (1996)* *Average Cost*	*HCUP (1996)* *Average DRG Weight*	*VA Rank*	*MEDPAR Rank*	*HCUP Rank*
6. Digestive system	32,464	5,028	0.845	19,477	4,454	0.898	3,295	4,163	0.850	13	8	9
5. Circulatory system	100,794	5,325	0.866	72,322	4,836	0.978	8,341	4,638	0.935	14	10	12
11. Kidney and urinary tract	19,269	5,925	0.918	8,620	5,016	0.978	1,408	4,578	0.897	15	11	10
17. Myeloproliferative diseases	9,452	6,400	1.137	3,288	7,433	1.403	672	7,275	1.247	16	19	19
1. Nervous system	27,709	6,491	0.970	21,877	6,187	1.048	3,026	5,620	1.035	17	15	14
7. Hepatobiliary and pancreas	14,109	7,075	1.191	4,782	5,786	1.194	953	6,404	1.162	18	14	17
4. Respiratory system	60,941	7,930	1.236	51,919	6,771	1.329	5,377	6,193	1.250	19	16	16
18. Infectious and parasitic diseases	6,003	9,052	1.238	6,995	6,830	1.381	1,004	6,685	1.327	20	17	18
25. HIV	2,468	9,313	1.482	90	10,751	1.534	346	9,690	1.600	21	21	21
24. Multiple significant trauma	58	15,416	1.954	148	15,326	1.954	47	11,460	1.954	22	22	22

Note: VA = U.S. Department of Veterans Affairs; MEDPAR = Medicare Provider Analysis Review; HCUP = Healthcare Cost and Utilization Project; DRG = diagnosis related group. Dollars are nominal and not adjusted for inflation. Comparison of VA rank to Medicare rank = 18.18 percent agreement, Kappa = .1429, p = .0011. Comparison of VA rank to HCUP rank = 22.73 percent agreement, Kappa = .1905, $p < .0001$.

cost-adjusted charges minus estimated charges) than the semilog models. Lipscomb et al. (1998) suggested that the ability to predict costs should be the primary concern when choosing the specification of the statistical model.

Transforming the dependent variable presents additional hurdles because the estimated costs need to be transformed back to the original metric (dollars). Although one can use retransformations, such as the smearing estimator (Duan 1983), this often requires arbitrary assumptions about the error distributions, and the point and variance estimates can still be biased (Mullahy 1998; Ai and Norton 2000; Manning 1998; Manning and Mullahy 2001). Some researchers have used cost regressions with heteroscedastic smearing estimates (Andersen, Andersen, and Kragh-Sorensen 2000). Given these limitations and our empirical evidence, we used OLS without transformation.

Variation in cost-adjusted charges is associated with variations in LOS and the DRG. We faced several options for including these variables in the cost regression. We could have included LOS without making any transformations, such that LOS would be a positive integer. Variations on this approach (e.g., using dummy variables) were also considered, but in every case, these approaches yielded a lower R^2 and a higher absolute mean error than the current model.

Although we used DRG weights to account for the approximately 500 DRGs, we could have used dummy variables. The gain in R^2 from using DRG-specific intercepts was approximately 1 percent greater than the models in which we included DRG weight. Given the complexity and instability of estimating a model with more than 500 collinear covariates, we chose to use DRG weights instead of DRG-specific intercepts.

DISCUSSION

The cost regression we estimated with 1996 MEDPAR data accounted for almost three-quarters of the variance in cost-adjusted charges. The cost regression did a better job of predicting high-cost stays than low-cost stays, and it proved to be highly robust to outliers. It was also robust to the year of input data: when the cost regression was run with 1994, 1995, and 1996 MEDPAR data, the estimated costs were correlated above .99 with the cost-adjusted charges. These findings suggest that the cost regression produced reliable cost estimates.

To assess the validity of the cost regression, we ranked the medical and surgical MDCs. Tables 4 and 5 show that the rankings are relatively consistent, and the agreement between VA, HCUP, and Medicare data is statistically significant. There was slightly more concordance in the surgical categories of care compared to the medical categories. This might be because the cost regression did a better job estimating high-cost cases than low-cost cases. These checks provide limited evidence that the average cost data for medical-surgical stays are valid and reliable.

A virtue of this method is that long-run average costs can be estimated with only eight variables from discharge records: MDC, DRG, LOS, number of diagnoses, death in hospital, sex, age, and number of ICU days. When we estimated costs with MEDPAR data, we accounted for 74 percent of the variance. Unfortunately, some data sets, such as HCUP, lack ICU days. This model could be used without ICU days, but the model's R^2 decreased from .74 to .69. Future research is needed to explore these cost estimates in more detail. In particular, comparing these costs to the VA Decision Support System will provide additional feedback on the validity and reliability of these cost estimates.

LIMITATIONS

One limitation of using MEDPAR data to estimate VA costs is that only hospital charges are reported. Physician charges are not included. Instead, they are reported on the Medicare physician/supplier part B files. Including the cost of physician services is important in determining VA costs. Physician costs are reported in the CDR; therefore, reconciling the estimated MEDPAR costs to the CDR distributes physician costs to each record in direct proportion to the hospital costs. Future research will look at alternative ways to estimate VA physician costs for inpatient stays. One option involves using resource-based weights calculated by Welch and Larson (1989) as an alternative to obtaining the physician services part B file and laboriously calculating these weights. Nevertheless, at this time, the VA costs include physician services, and these costs are allocated proportionately in accordance with the hospital costs.

Another limitation is that the cost regression did not capture all of the variance. A consequence of this is that the estimated costs have less variance and fewer outliers than the true VA costs. This limitation has two important implications. First, it suggests that researchers may not want to use the estimated costs for identifying high- or low-cost outliers. Second, it implies that the cost regression biases the variance of the estimated costs downwards. The reason for this is that many factors that affect costs are not included in the cost regression. Stays that may differ in cost but have identical observed factors are assigned the same estimated cost. In Table 6, we show the costs reported by 1996 MEDPAR for five DRGs, along with the estimated costs from our regression. As is clear from this table, the standard deviations for estimated costs are smaller than the actual costs. Also, note that the minimum and maximum values are attenuated toward the mean.

TABLE 5 Ranks of Surgical Major Diagnostic Categories from Least to Most Expensive

Major Diagnostic Category	*VA (1998)* n	*VA (1998)* Average Cost	*VA (1998)* Average DRG Weight	*MEDPAR (1996)* n	*MEDPAR (1996)* Average Cost	*MEDPAR (1996)* Average DRG Weight	*HCUP (1996)* n	*HCUP (1996)* Average Cost	*HCUP (1996)* Average DRG Weight	*VA Rank*[a]	*MEDPAR Rank*	*HCUP Rank*
1. Nervous system	1	3,273	0.600	1	3,491		2,044	4,091	0.822	1	1	1
2. Diseases of the eye	1,379	4,726	0.672	435	4,339	0.706	110	4,220	0.716	2	2	2
3. Ear, nose, mouth, and throat	857	6,135	0.993	301	5,021	1.051	1,745	4,987	1.049	3	4	3
4. Respiratory system	4,750	6,400	1.021	6,406	4,764	0.955	498	5,835	1.078	4	3	4
5. Circulatory system	2,333	7,592	1.179	470	6,934	1.235	253	8,147	1.556	5	5	8
6. Digestive system	5,661	10,165	1.597	4,206	9,138	1.643	596	10,303	1.706	6	7	12
7. Hepatobiliary and pancreas	3,018	10,926	1.378	1,189	8,266	1.638	576	6,450	1.209	7	6	5
8. Musculoskeletal system	717	11,085	1.131	210	14,049	1.131	54	8,481	1.131	8	15	9
9. Skin, subcutaneous tissue, and breast	19,288	11,361	1.706	15,085	10,051	1.873	3,644	9,164	1.295	9	9	10
10. Endocrine, nutritional, and metabolic	1,768	12,046	1.759	706	9,635	1.784	225	7,641	1.693	10	8	7
11. Kidney and urinary tract	5,458	12,572	1.963	5,253	10,563	1.934	749	15,189	2.835	11	10	16
12. Male reproductive system	212	12,942	2.356	104	13,430	2.551	43	14,849	2.911	12	14	15
13. Female reproductive system	76	13,658	2.371	93	14,290	2.371	8	13,827	2.371	13	17	14
14. Pregnancy and childbirth	895	14,958	2.296	559	12,536	2.448	119	18,644	3.531	14	13	19
16. Blood and Immunological disorders	3,628	15,008	2.210	3,378	11,533	2.255	912	9,622	1.999	15	12	11

(continued)

TABLE 5 (continued)

Major Diagnostic Category	*VA (1998)* n	*VA (1998)* Average Cost	*VA (1998)* Average DRG Weight	*MEDPAR (1996)* n	*MEDPAR (1996)* Average Cost	*MEDPAR (1996)* Average DRG Weight	*HCUP (1996)* n	*HCUP (1996)* Average Cost	*HCUP (1996)* Average DRG Weight	*VA Rank*[a]	*MEDPAR Rank*	*HCUP Rank*
17. Myeloproliferative diseases	11,851	15,282	2.150	7,381	14,153	2.664	1,680	10,643	2.371	16	16	13
18. Infectious and parasitic diseases	1,216	15,601	1.808	684	10,989	1.978	153	7,004	2.131	17	11	6
19. Mental diseases and disorders	4,146	17,565	2.827	2,631	15,176	2.881	435	17,419	4.153	18	18	18
20. Alcohol/drug use	27,207	19,610	3.323	31,201	17,713	3.658	3,415	17,026	3.652	19	19	17
21. Injuries, poisonings	1,071	22,234	3.554	1,216	19,005	3.554	210	20,785	3.877	20	20	21
22. Burns	44	25,345	4.789	9	19,600	4.789	22	20,061	5.661	21	21	20
23. Factors influencing health status	22	26,503	4.240	173	30,361	4.372	81	32,782	4.414	22	22	22

Note: VA = U.S. Department of Veterans Affairs; MEDPAR = Medicare Provider Analysis Review; HCUP = Healthcare Cost and Utilization Project; DRG = diagnosis related group. Dollars are nominal and not adjusted for inflation. Comparison of VA rank to Medicare rank = 27.27 percent agreement, Kappa = .2381, $p < .0001$. Comparison of VA rank to HCUP rank = 18.18 percent agreement, Kappa = .11429, $p = .0011$.

a. Ranked by ascending cost.

TABLE 6 The Cost Regression's Effect on the Variation of the Estimated Costs

DRG	n	*Average Cost*	SD	*Minimum*	*Maximum*
14. Specific cerebrovascular disorders except transient ischemic attack					
Cost	10,534	6,829	7,587	7	175,346
Estimated cost	10,534	7,377	7,476	685	147,135
79. Respiratory infections and inflammations, age older than 17 with complications and comorbidities					
Cost	7,767	7,923	8,445	16	213,967
Estimated cost	7,767	8,210	6,423	685	198,091
88. Chronic obstructive pulmonary disease					
Cost	15,428	4,786	5,525	5	203,877
Estimated cost	15,428	4,535	4,269	685	128,695
89. Simple pneumonia and pleurisy, age older than 17 with complications and comorbidities					
Cost	12,905	5,468	8,863	8	662,916
Estimated cost	12,905	5,238	4,675	685	160,280
127. Heart failure and shock					
Cost	21,463	4,941	4,979	10	109,945
Estimated cost	21,463	5,224	4,479	685	190,673

Note: DRG = diagnosis related group. Cost is the Medicare Provider Analysis Review cost-adjusted charges for 1996. Estimated cost is the estimated cost-adjusted charges for fiscal year 1998. Dollars are nominal and not adjusted for inflation.

USING THE AVERAGE COST DATA

The medical-surgical average cost databases are available for VA researchers. To merge these data with the VA utilization files, researchers need to reconfigure the VA utilization files, as we did to create the database. As an easy-to-use alternative, we created a discharge data set that combines the medical-surgical, rehabilitation, mental health, and long-term care stays and can be easily merged to the VA patient treatment file (main). For more details, see the HERC average cost guidebook (Wagner et al. 2001).

Finally, users should remember that these cost estimates reflect costs listed in the CDR, which does not include the cost of capital financing or malpractice because they are covered by other federal agencies. Therefore, the HERC cost estimates may not be appropriate to use when a health care program requires additional space or affects malpractice claims or when VA costs are compared to those of non-VA providers.

The average cost method assigns the same cost to all inpatient stays with the same demographic and discharge information. Patients with identical observed characteristics are assigned the same cost. It is important to note that it is not always appropriate for researchers to use the average cost data. Although these data were created with cost-effectiveness analysis in mind, if researchers are interested in assessing the cost effectiveness of close substitutes, then these data are likely to be inappropriate unless one of the interventions affects one of the variables in the cost regression (e.g., LOS). When these data are not helpful, micro-costing methods, such as pseudo-bills or direct measurement, would be necessary (see Smith and Barnett 2003 [this issue]).

CONCLUSION

This article reports on the methods we used to develop a VA cost database for medicine and surgery inpatient care. The cost estimates are generated from a regression model based on MEDPAR data. The regression model does not account for market-level factors or input prices. This strategy reflected our goal of generating long-run average VA costs. In particular, we generated national VA costs by reconciling the estimated costs with the VA national budget. Although we also generated local VA costs by reconciling the estimated costs with local VA budgets, we strongly encourage researchers to use the national cost estimates. The local cost estimates may be appropriate for a sensitivity analysis in a cost-effectiveness analysis. However, variation in local cost estimates reflects local budget allocations, not underlying differences in input prices, market factors, or production efficiency.

With relative ease, these methods could be adapted to estimate the cost of care in other health care systems. An important factor to consider is the data set on which the cost regression is estimated. We used MEDPAR data restricted to VA patients, but researchers could use the 5 percent MEDPAR data set, HCUP data, or other hospital discharge data. Some of these data sets, such as HCUP, do not have all eight independent variables, thus limiting the model's fit. The researchers would need to determine whether the cost estimates should be reconciled to an accounting data set to reflect system-specific costs. As we have discussed above, caution should be used in applying these cost estimates in a research project. Nevertheless, these methods can produce robust estimates.

REFERENCES

Ai, C., and E. C. Norton. 2000. Standard errors for the retransformation problem with heteroscedasticity. *Journal of Health Economics* 19 (5): 697-718.

Andersen, C. K., K. Andersen, and P. Kragh-Sorensen. 2000. Cost function estimation: The choice of a model to apply to dementia. *Health Economics* 9 (5): 397-409.

Baker, L. C. 1997. The effect of HMOs on fee-for-service health care expenditures: Evidence from Medicare. *Journal of Health Economics* 16 (4): 453-81.

Baker, L. C., and K. S. Corts. 1996. HMO penetration and the cost of health care: Market discipline or market segmentation? *American Economics Review* 86 (2): 389-94.

Barnett, P. G. 1997. Research without billing data. Econometric estimation of patient-specific costs. *Medical Care* 35 (6): 553-63.

Barnett, P. G., S. Chen, and T. H. Wagner. 2000. *Determining the cost of VA care with the average cost method for the 1993-1997 fiscal years*. HERC technical report no. 2. Palo Alto, CA: Health Economics Resource Center.

Barnett, P. G., and T. H. Wagner. 2003. Preface. *Medical Care Research and Review* 60 (3 Suppl.): 7S-14S.

Centers for Medicare and Medicaid Services. 2003. Medicare provider analysis and review (MEDPAR) of short-stay hospitals. Retrieved 10 March 2003 from http://cms.hhs.gov/statistics/medpar/default.asp.

Chapko, M. K., J. L. Ehreth, and S. Hedrick. 1991. Methods of determining the cost of health care in the Department of Veterans Affairs medical centers and other non-priced settings. *Evaluation and the Health Professions* 14 (3): 282-303.

Cook, R. D., and S. Weisberg. 1982. *Residuals and influence in regression*. London: Chapman and Hall.

Duan, N. 1983. Smearing estimate: A nonparametric retransformation method. *Journal of the American Statistical Association* 78 (383): 605-10.

Gold, M. R., J. E. Siegel, L. B. Russell, and M. C. Weinstein, eds. 1996. *Cost-effectiveness in health and medicine*. Oxford, UK: Oxford University Press.

Lipscomb, J., M. Ancukiewicz, G. Parmigiani, V. Hasselblad, G. Samsa, and D. B. Matchar. 1998. Predicting the cost of illness: A comparison of alternative models applied to stroke. *Medical Decision Making* 18 (2 supp.): S39-56.

Manning, W. G. 1998. The logged dependent variable, heteroscedasticity, and the retransformation problem. *Journal of Health Economics* 17 (3): 283-95.

Manning, W. G., and J. Mullahy. 2001. Estimating log models: To transform or not to transform? *Journal of Health Economics* 20 (4): 461-94.

Mullahy, J. 1998. Much ado about two: Reconsidering retransformation and the two-part model in health econometrics. *Journal of Health Economics* 17 (3): 247-81.

Smith, M. W., and P. G. Barnett. 2003. Direct measurement of health care costs. *Medical Care Research and Review* 60 (3 Suppl.): 74S-91S.

Wagner, T. H., S. Chen, W. Yu, and P. G. Barnett. 2001. *Guidebook for HERC's inpatient average cost datasets for VA care*. Menlo Park, CA: VA Health Economics Resource Center. Available from http://www.herc.research.med.va.gov/ACM.htm.

Welch, H. G., and E. B. Larson. 1989. Rapid estimation of hospital charges in patients with leukemia. Validation of a multivariate prediction model. *Medical Care* 27 (9): 900-904.

Wright, S. M., M. Hossain, and L. A. Petersen. 2000. Enrollment in the Medicare program among veterans turning age 65. Paper read at VA HSR&D Annual Conference, 22-24 March 2000, Washington, D.C.

Wright, S. M., R. P. Lamkin, and L. A. Petersen. 2000. Medicare enrollment status among VA users by VISN. Paper read at VA HSR&D Annual Conference, 22-24 March 2000, Washington, D.C.

Yu, W., T. H. Wagner, S. Chen, and P. G. Barnett. 2003. Average cost of VA rehabilitation, mental health, and long-term hospital stays. *Medical Care Research and Review* 60 (3 Suppl.): 40S-53S.