Guidebook

HERC's Average Cost Datasets for VA Inpatient Care FY1998 - FY2009

Todd H. Wagner, Adam Chow and Paul G. Barnett

June 2010

HERC's Average Cost Datasets for VA Inpatient Care FY1998 – FY2008. Guidebook

Health Economics Resource Center (HERC)
VA Palo Alto Healthcare System
795 Willow Road (152 MPD)
Menlo Park, CA 94025
650-617-2630
650-716-2639 (fax)
herc@va.gov

Suggested citation: Wagner TH, Chow A. Barnett PG. HERC's Average Cost Datasets for VA Inpatient Care FY1998 - FY2008. Guidebook. Menlo Park CA. VA Palo Alto, Health Economics Resource Center; 2009.

Acknowledgements: Many programmers have made this work possible including Sam King, Jesse Velez, Matt Yeh, and Shuo Chen. Wei Yu contributed to earlier versions of this work. We would like to acknowledge Douglas Bradham, Ann Hendricks, Denise Hynes, Terri Menke, and Anne Sales. Their insight, experience and advice helped to produce a higher quality product. We would also like to acknowledge Health Services Research and Development Service (HSR&D) and the Cooperative Studies Program (CSP) for funding this research.

Contents

Tables and Figures

Terms

BEDCDR	CDA Assigned to Bedsection
BEDSECN	Bedsection Variable (synonymous with treating specialty)
BSINDAY	Bedsection admission day
BSOUTDAY	Bedsection discharge day
CDA	Cost Distribution Accounts
CMS	Centers for Medicare and Medicaid Services
COSTL	Local Cost Estimate
COSTN	National Cost Estimate
DRG	Diagnosis Related Group
DSS	Decision Support System
ESRD	End-Stage Renal Disease
FMS	Financial Management System
FY	Fiscal Year
HCUP	Healthcare Cost and Utilization Project
HERC	Health Economics Resource Center
ICU	Intensive Care Unit
LVB	Leave Days in the Bedsection
MAVERIC	Massachusetts Veterans Epidemiology Research and Information Center
NDE	National Data Extract
PRRTP	Psychosocial Residential Rehabilitation Treatment Programs
PTF	Patient Treatment File
RCC	Ratio of Costs to Charges
RUG	Resource Utilization Group
RVU	Relative Value Unit
SCRSSN	Scrambled Social Security Number
TRT	Treatment Specialty (synonymous with bedsection)
VISNs	Veterans Integrated Service Networks

Abstract

The U.S. Department of Veterans Affairs (VA) provides health care to veterans at more than 120 inpatient facilities. In 1999, the VA funded the Health Economics Resource Center (HERC) to adapt existing cost methodologies and to expand methods to estimate costs of health care encounters. This guidebook describes HERC's method for estimating the cost of VA inpatient stays from fiscal years 1998-2009; Chapter 5 provides details on how to use the data.

Inpatient stays can be classified into two categories depending on basis of admission. Acute inpatient stays include short-stay hospitalizations for acute medicine and surgical treatment, and are typically less than 60 days long. Non-acute inpatient stays encompass rehabilitation, blind rehabilitation, spinal cord injury, psychiatric, substance abuse, intermediate medicine, domiciliary, and psychosocial residential rehabilitation stays. For both types of care, we estimate costs assuming that every health care encounter has the average cost of all encounters with the same characteristics. We use length of stay as the determinant of cost in a non-acute hospitalization. This makes the assumption that every day of stay has the same cost and costs are directly proportional to length of stay. In contrast, we estimate the cost of acute medical-surgical hospital care by using an econometric cost function. This method relies on non-VA relative value weights used by Medicare to pay hospitals for providing care to Medicare patients.

The user's guide to the average cost dataset discusses methods in building the dataset, assumptions underlying the dataset, and how to use the dataset. The user's guild also discusses the data limitations and why these data may not be appropriate for every study.

1. Introduction

The U.S. Department of Veterans Affairs (VA) provides health care to veterans at more than 120 inpatient facilities. In 1999, the VA funded the Health Economics Resource Center (HERC) to extend prior methods and estimate costs for all VA encounters.[1] Our goal was to develop a set of long-term costs that could be used in cost-effectiveness analysis. We use the term long-term in the economic sense that all costs are variable. A companion User Guide on the HERC Outpatient Cost Data is also available on our web site (http://www.herc.research.va.gov/publications/default.asp).

Known as the Average Cost method, we assume that every health care encounter has the average cost of all encounters that share its same characteristics. To find the cost of rehabilitation, blind rehabilitation, spinal cord injury, psychiatric, substance abuse, intermediate medicine, domiciliary, and psychosocial residential rehabilitation stays, we found the average cost of a day of stay, and multiplied it by length of stay to estimate the cost of care. This makes the assumption that every day of stay has the same cost, that is, that costs are directly proportionate to the length of stay. Hereafter, we refer to this as rehabilitation, mental health or long-term care.

To find the cost of acute medical-surgical hospital care, we built a cost function using relative value units (RVUs) from the non-VA sector. These RVUs were the Diagnosis Related Group (DRG) weights used by the Centers for Medicare and Medicaid Services (CMS) to reimburse U.S. hospitals for the care they provide to Medicare patients. The RVUs reflect the effect of diagnosis on the relative quantity of resources used in a hospital stay. In addition to DRG weights, the cost function included length of stay, demographic and other clinical information. The method we employed makes the following assumptions: (1) that the non-VA relative value units, the Medicare DRG weights, reflect the relative costs of VA hospital stays, and (2) that all stays with the same characteristics have the same cost.

The Average Cost Data are based on care provided in the federal fiscal year, which begins on October 1 and ends on September 30 of the following year. The convention is to refer to a federal fiscal year (FY) by the year it ends, thus FY98 represents the period October 1, 1997 to September 30, 1998.

1.1. Updates

For FY98-FY00, we used resource utilization groups (RUG) to weight the costs of long-term care. Veterans with higher RUG scores have higher costs. Changes in the RUG data limit this method to FY98-FY00 only. After FY00, we calculate costs using a per diem rate and length stay.

In FY04, we switched from using the Cost Distribution Report (CDR) to a department- level summary from the Decision Support System Nation Data Extracts.

2. Cost and Utilization Data

We used the Cost Distribution Report (CDR) to create the 1998-2003 HERC average cost datasets. The CDR ceased production in 2004 and since then we have used cost data from the DSS National Data Extract (treatment specialty file), summarized to departments. Below is a brief discussion of the CDR and DSS NDE.

2.1. Cost Data

2.1.1. Distribution Report

The CDR was routinely prepared by all VA medical centers, and represented an estimate of the costs expended by each VA patient care department. The CDR was created by distributing costs reported in the Financial Management System (FMS) cost centers to the "cost distribution accounts" (CDA) of the CDR. The CDAs reflected patient care departments, such as Medical Intensive Care, or Ambulatory Care, Medicine. We created 11 groups of inpatient care and summarized the CDR into these categories (Table 1).

Table 1: Categories of Inpatient Care

Category of Care
Medicine
Rehabilitation
Blind Rehabilitation
Spinal Cord
Surgery
Psychiatry
Substance Abuse
Intermediate
Domiciliary
Long Term
Psychosocial residential rehabilitation treatment programs (PRRTP)

2.1.2. DSS Summary

The CDR ceased production in 2004. For a department-level cost dataset, we chose to create our own from the DSS National Data Extract Treatment Specialty File (TRT). The TRT is an encounter-level dataset organized by treating specialty (identical to the bedsection). By summarizing the TRT into a department level dataset, we ensure that the HERC and DSS NDE's are based on the same underlying costs. In prior years, HERC and DSS included different costs. Therefore, when researchers compared HERC and DSS costs prior to FY04, the datasets differed in both the underlying costs and the relative value units. Now, with FY04 the underlying costs are the same and the only difference between the datasets is the relative value units.

2.2. Utilization Data

The VA maintains utilization data in the Patient Treatment Files (PTF). These data do not contain cost, payment or charge data, but they do include patient demographics, length of stay, and the Diagnosis Related Group (DRG). There are three file types of PTF files: observation, extended care and acute care. The observation, extended care and acute care have a main, bedsection, observation and census file. In total, there are 12 files per fiscal year and we use all twelve in generating the Average Cost Data.

2.2.1. PTF Main

This file reports all hospital discharges within the fiscal year. This file contains one record for each hospital discharge. The main file does not use a definition of a hospital stay that is comparable to non-VA hospitals. In the non-VA sector, an acute medical-surgical hospitalization followed by a long-term care stay would be recorded as two different stays. In the PTF main file, however, this is often recorded as a single stay.

2.2.2. PTF Bedsection

The PTF Bedsection file, like the PTF main, is a discharge dataset. However, unlike the PTF Main, there is a record for each sequential bedsection. The bedsection is the "treating specialty" assigned to the physician who is responsible for the patient's care. It roughly corresponds to the location where care is delivered, such as medicine, intensive care, rehabilitation, or long-term care. The Bedsection file provides information on the number of days the patient spent in each bedsection. The PTF Bedsection and Main files have different data elements. We use both files to characterize hospital discharges.

2.2.3. PTF Census files

The PTF main and Bedsection files include information on all discharges, regardless of when they began. These files do not report on patients occupying beds at the end of the reporting period. To fill this gap, the PTF Census Files includes information on patients who are in a VA hospital at the end of the fiscal year.

2.2.4. Acute and Extended Care files

Most stays that <u>start</u> in a nursing home file are included in the extended care file, regardless of the bedsection in which the patients ends up. On the other hand, stays that do not start in the nursing home are usually listed in the acute care files.

2.2.5. Observation Bed files

The Observation Bed file was first created in 1998. If a stay includes an observation bedsection, then the observation portion of the stay is separated from the rest of the stay and included in this file. Most observation bed stays are one-day stays, with the patient being discharged from the hospital.

Observation bedsections were created at the same time as the VA was implementing managerial performance incentives to reduce the number of inpatient days per 1000 treated veterans. Observation data are not included in this performance measure.

Observation bed stays are very heterogeneous, and they present some difficulty in determining their cost. We decided that all observation stays would be given the daily cost of the marginal cost of a day. To calculate the marginal cost of day, we used a statistical model with Medicare data (see Observation Costing Method).

2.3. Merger of cost and utilization databases

We merged the cost and utilization databases, using the categories described in Table 1. In this process, we exclude some facilities and take facility mergers into account.

2.3.1. Excluded facilities

Prior to merge the cost and utilization data, we excluded the some facilities that do not provide patient care. These include records for VA Headquarters (station 101), information services centers, and other VA support facilities. A list of these facilities, and their three-digit facility number, is provided in Table 2.

Table 2: Excluded Facilities

Facility Number (sta3n)	Facility Name
101	VHA Headquarters
200	Austin Automation Center
722	Albuquerque, NM Outpatient Center
741	Denver CHAMPVA
721, 724, 742, 760, 761, 762, 763, 764, 765	
792	Prosthetics Center
794	Somerville
797	Hines (CIO)

We felt that central administration may involve activities that are more characteristic of a health care payer, rather than a health care provider. For this reason, we decided not to count these facility's costs as overhead costs that should be distributed to patient care departments.

2.3.2. New Facilities and Mergers

VA has been emphasizing ambulatory care and improved patient access. Consequently, VA can create new facilities and consolidate others. When one facility merges with another, they both take on a single identification number (see Table 3). If facilities consolidate into a single facility during a fiscal year, we assume the consolidation happened at the start of the fiscal year.

Table 3: Facility Mergers

	Old STA3N	New STA3N
1997		
VA Chicago Health Care System	535	537
VA Central Alabama Health Care System	680	619
VA North Texas Health Care System	522	549
Southern California System of Clinics	665,752	665
Hudson Valley VA Health Care System	533	620
VA Central Iowa Health Care System	592	555
VA Greater Nebraska Health Care System	574	597
1998		
VA Eastern Kansas Health Care System	686	677
VA Montana Health Care System	617	436
North Florida/South Georgia Veterans Health Care System	594	573
VA Greater Los Angeles Health Care System	752	691
1999		
Greater Los Angeles Health Care System	665	691
Boston VA Health Care System	525	523
2000		
NY Harbor Health Care System	527	630
Upstate NY Health Care System	532	528
Upstate NY Health Care System	670	528
VA Mid Tennessee Health Care System	622	626
Upstate NY Health Care System	500	528
VA Nebraska Western	584	636
2001		
Columbia MO Harry S Truman Memorial VA Medical Center	543	589
Eastern Kansas VA Health Care System	677	589
Marion IL VA Medical Center	609	657
Popular Blue MO John J Pershing Medical Center	647	657
2002		
VA Eastern CO Health Care System	567	554

Kansas City VA Medical Center 2003-2009 None	452	589

2.4. Definition of category of care

We created "patient care categories," which represent our best judgment about what constitutes the smallest common denominator between the cost and utilization database. A patient care category represents a group of related cost accounts and their associated utilization.

We defined eleven patient care categories based on earlier work.[1] For some categories of care at some medical centers, there were mismatches between cost and utilization data. Most mismatches were handled by assigning the costs and utilization to a similar department, creating a higher level of data aggregation.

2.5. Merger of cost and inpatient utilization data

VA databases report costs and utilization in a federal fiscal year. As mentioned above, we wanted to identify the amount of care provided during the fiscal year. Since hospital stays may span fiscal years, we developed a method to divide hospital utilization between fiscal years.

The denominator for the cost data was the fiscal year, whereas the denominator for the utilization data was discharges. These denominators are not equivalent. We could have ignored this difference. This would have been equivalent to assuming that bed occupancy was constant over the year. However, this assumption would be wrong because we know that there is a trend to shorten length of stay and to reduce hospitalization. And we did not want to assume that the same number of patients are in the hospital at the start and at the end of the fiscal year.

A better way to adjust for the difference in denominators was to use information from the Census files. With the Census files we adjusted the discharge file so that it more closely approximated utilization in the fiscal year.

For the utilization data, we included days spent during the current fiscal year by all patients. For those discharged during the fiscal year, their data came from the PTF, limiting the days to those in the fiscal year. For those patients not discharged by the end of the fiscal year, we obtained these days of stay from the PTF census files. This calculation included "leave" days, that is, days that a patient was absent from a hospital, though not yet discharged. The PTF records leave days, but it does not indicate when they occurred. We assumed that leave days are uniformly distributed throughout the stay.

The finest level of detail for the cost data is at department level; patient-level cost data do not exist. To merge the cost and utilization data, we identified 11 categories of inpatient care (see Table 4). There is an overlap between psychiatry, substance abuse and PRRTP programs, which are less intensive inpatient programs for psychiatry and substance abuse.[2] Only approved medical centers can have a PRRTP program.

Table 4: Inpatient Categories of Care

Category of Care	Bedsection / Treating Specialty
Medicine	1-19, 024, 30, 31, 34, 83, 1E
Rehabilitation	20, 35, 41, 82, 1D
Blind Rehabilitation	21, 36
Spinal Cord	22, 23
Surgery	48-63, 65, 78, 97
Psychiatry	25, 26, 28, 29, 33, 38, 39, 70, 71, 75, 76, 77, 79, 89, 91, 92, 93, 94
Substance Abuse	27, 72, 73, 74, 84, 90
Intermediate	32,40
Domiciliary	37, 85, 86, 87, 88
Long Term Care	42, 43, 44, 45, 46, 47, 64, 66, 67, 68, 69, 80, 81, 95, 96, 1A, 1B, 1C
PRRTP[⊥]	25, 26, 27, 28, 29, 38, 39

[⊥] PRRTP is less intensive psychiatry and substance use. Only approved facilities can have a PRRTP program. In FY08 PRRTP programs existed at: 459, 463, 501, 504, 515, 516, 518, 523, 528, 541, 546, 549, 554, 555, 556, 561, 568, 573, 586, 589, 590, 595, 598, 620, 622, 631, 632, 635, 640, 645, 653, 656, 658, 662, 663, 666, 676, 678, 687, 689.

2.6. Data reconciliation

After merging the cost and utilization data for each medical center, typically there are some discrepancies that need to be reconciled. A discrepancy is when there is utilization in one category but no costs, or vice versa. In general, discrepancies are quite rare. Appendix 1 describes all the reconciliations for FY98-present.

2.7. Daily rate

After reconciling the 11 inpatient categories, we divided total costs by total utilization to find the average cost for each category of care at each medical center; this is used in estimating the local costs. We also calculate the average daily rate for each of the categories for the nation; this is used in estimating the national cost.

3. The cost of rehabilitation, mental health and long-term care

3.1. What is rehabilitation, mental health and long-term care?

Most US hospitals differentiate between short-stay acute medical-surgical and non-medical/surgical hospitalizations. Short-stay acute medical-surgical hospitalizations are generally for acute medicine and surgical treatment. While over 90% of short stay hospitalizations are less than 60 days long, there are rare cases that involve a length of stay up to and over a year. In the VA, about half of the inpatient stays can be categorized as acute medical-surgical defined by their bedsections (see Table 4). The remaining stays include rehabilitation, blind rehabilitation, spinal cord injury, psychiatry, substance abuse, intermediate care, domiciliary, and nursing home. This chapter describes how we estimated the cost for rehabilitation, mental health or long-term care.

Between FY98 and FY00, we case-mix adjusted the nursing home costs. After FY00, nursing home care is based on a per diem cost. More information on the cost of nursing home care is covered elsewhere.[3]

3.2. Cost methodology for rehabilitation, mental health and long-term care

Determining costs for rehabilitation, mental health and long-term care is straightforward. We multiplied the average daily rate, discussed earlier, by the patient's length of stay. When we use the local daily rate, the result is the local cost. When we use the national daily rate, the result is the national cost.

3.2.1. Leave and pass days

For stays that began before the beginning of the fiscal year, we found the length of stay during the current fiscal year by finding the number of days between the discharge date and the beginning of the fiscal year. This calculation considered "leave" days, that is, days that the patient was absent from the hospital, though not yet discharged. Leave days are also called Absent Bed Occupant Days and are given the variable name LVB in the PTF. The PTF records leave days in a variable named LVB, but it does not record when they occurred. We assumed that leave days are uniformly distributed throughout the stay.

3.2.2. Local outlier costs

As one might expect, there is more variation in the local daily rates than the national daily rates. This raises the question about the accuracy of the local rate. To help identify inaccurate local costs, we generated a flag if a medical center had a daily rate that was 2 standard deviations from the average of all VA medical centers (for that particular care category). Part of this variation could be explained by wages or high cost procedures. Therefore, the flag variable allows the analyst to check for outliers when using the local cost estimates.

3.2.3. Why local rates at all?

Given that there is more variation in the local rates than the national rates, one may ask why do we calculate local rates at all. The answer is that sometimes the variation in the local rates is important. Wages are one factor that affects costs, as they depend on the labor market in different geographic localities. If a researcher is interested in the effect of an intervention on a local medical center or VISN, then the local rates may be more appropriate because they partly reflect the wage differentials and other local differences.

4. The cost of acute medical-surgical hospitalizations

The cost of acute medical-surgical hospital care in VA can be more accurately estimated by incorporating diagnostic information from the administrative record, and avoid the assumption that every day of stay is of equal cost.[4] We used an econometric cost function, with parameters estimated from non-VA data, to impute the costs for acute medical-surgical stays in the VA.

This method relies heavily on non-VA relative value weights. These weights, known as DRG weights, are used to pay hospitals for providing care to Medicare patients. Upon discharge, patients are assigned a Diagnosis Related Groups (DRGs) based on their primary diagnosis. This weighting system is used by the Centers for Medicare and Medicaid Services to determine Medicare payments to hospitals.

This section presents the cost function that we developed with Medicare data. Given the complexities in this chapter, a flow diagram is provided in Appendix 2 to help readers visualize the process.

4.1. Making an acute medical-surgical inpatient discharge database

The VA tracks patients using bedsection codes. Because a patient can get transferred among bedsections multiple times within a single acute medical-surgical hospital stay, keeping track of bedsections provides us with a great amount of detail that is necessary for identifying acute medical-surgical stays.

To use non-VA relative value units, we had to restructure the VA data to use the same definition of acute stays as is found outside the VA. Most non-VA databases are organized as discharge databases with each record representing an acute medical-surgical hospital discharge. While the PTF Main is a discharge database, it does not distinguish between acute medical-surgical and non-medical/surgical care. In addition, the PTF Bedsection file is a discharge file but it separates each record into bedsection stays, even if the bedsections are all part of one acute medical-surgical stay. Therefore, we had to make a database of acute medical-surgical discharges using the PTF bedsection file. Table 4 shows the bedsection codes used to identify medicine and surgery.

We then sorted the data by scrambled social security number (SCRSSN), medical center (STA3N), bedsection in day (BSINDAY) and bedsection out day (BSOUTDAY). Acute medical-surgical bedsection stays that were contiguous in time were considered to be part of the same hospitalization. Transfers within acute medical-surgical bedsections, such as from surgery to medicine, were aggregated into a single record. We adopted the rule that if a patient was transferred from an acute medical-surgical bedsection to another acute medical-surgical bedsection that this would be considered part of the same stay. Similarly, if a person was transferred from an acute medical-surgical bedsection to a non-medical/surgical bedsection, we ruled that the acute medical-surgical stay had ended. Transfers from an acute medical-surgical bedsection to a non-medical/surgical bedsection and back to an acute medical-surgical bedsection yielded one non-medical/surgical and two acute medical-surgical stays.

We created a program to accumulate contiguous acute medical-surgical bedsection stays. The program also performs a number of other important functions, such as recalculating length of stay, identifying the highest DRG weight from multiple bedsections, and calculating number of days spent in intensive care (ICU). The SAS code for creating medical/surgical discharges is available upon request.

4.2. Selecting the DRG and the relative value associated with a DRG

VA assigns a DRG to each bedsection segment of the hospital stay, and another DRG to the PTF main file, representing the DRG for the entire stay. The DRG is based on the principal diagnosis, the condition that is responsible for the patients' admission to the hospital. The Health Care Financing Administration has developed a set of weights based on the DRG (DRG weights). These DRG weights are used to pay hospitals for Medicare patients.

We decided to use the DRG weights for our relative weights in the cost function. DRG weights are not part of the VA databases and were obtained from CMS and added to the VA files. Given that we had 1996 Medicare data, we merged the 1996 DRG weights from CMS with the PTF bedsection file. Then while we were making the acute medical-surgical VA hospital discharge file, the highest DRG weight across all bedsections was maintained. The rationale for this is that a private hospital would follow the same logic to maximize reimbursement.

We considered, but did not use, other relative value systems. We decided that the weights developed by states to pay Medicaid are likely to reflect the patterns of practice in a specific state and that it would not be appropriate to apply them to the VA's national system of hospitals. Some relative value systems, such as the Severity of Illness Index, may provide some additional measure of relative cost[5], but they are not feasible for us to implement as they require data that are not available in VA utilization data at Austin. Patient Management Categories and Disease Staging are case-mix methods that can be applied to standard datasets, but they have been found to explain only 1-2% more variation than DRGs used alone.[6]

4.3. Length of stay

Length of stay is reported in the PTF bedsection file. But we had to recalculate length of stay according to our definition of acute medical-surgical stay. Consequently, length of stay represents all days the patient spent in contiguous acute medical-surgical care bedsections during the stay.

4.4. Building the cost function

In past years we used an econometric method of estimating VA acute medical-surgical care costs[4]. Starting with FY98, we developed a cost function for estimating the cost of acute medical-surgical care. The cost-function is based on non-VA data, where the hospital stay is the unit of analysis. Using the stay (rather than the average stay) as the unit of analysis provides much more variation, including observations with high DRG weights and long lengths of stay. The cost function approach allowed us to construct a more complex model that better simulates the cost of stays with characteristics that are very different from the mean.

While the mechanics of the cost function are complicated, the intuition is relatively straightforward. We built a statistical model with a hospital discharge dataset. This regression model had cost adjusted charges on the left-hand side. On the right-hand side, we included variables such as length of stay, DRG weight, whether the patient died in the hospital, age, gender, and so forth. We saved the parameters from the regression model (i.e., the beta coefficients). This vector of coefficients was used to estimate costs in the VA data. It is important to note that the only way this approach can work is for both datasets to have the exact same right-hand side variables.

4.4.1. Data

We chose to use Medicare data for the cost function. Medicare data have some limitations, namely that Medicare does do not cover non-disabled individuals under age 65. For this reason, we carefully compared Medicare data from veterans to the Health Care Cost and Utilization Project (HCUP) data.

To provide some background on these datasets, the Medicare data were a subset of the 1996 MedPar file. The MedPar file was constructed by researchers at the Massachusetts Veterans Epidemiology Research and Information Center (MAVERIC). They established a cohort of all veterans who were users of either inpatient or outpatients VA services between 1992 and 1994 and who had their 65th birthday in 1994. This cohort was then linked to Medicare denominator file to obtain Medicare enrollment. The file that we received represented 372,046 stays from hospitals in the continental US.

The HCUP data represents discharges from all types of hospitals in 22 states. Detailed information on the HCUP dataset is available on-line from www.ahrq.gov.

The primary question is, can we use the Medicare data to build a model that can estimate costs for younger veterans? Recall that Medicare data do not include non-disabled individuals under age 65. We answered this question by building a cost function with Medicare data. The function was then used to estimate the cost of stays in the HCUP sample. We then compared the estimated Medicare costs to the costs reported in the HCUP. This comparison was made for adults over 65 as well as adults under age 65. The remainder of this section describes this comparison.

First we selected a 40% random sample of non-ESRD Medicare claims in the MAVERIC cohort (125,457). With these claims, we estimated the following model:

CAC=a+b1died +b2sex +b3age+ b4npr+ b5npr2 +b6los + b7poslos + b8neglos + b9nlos2
+b10plos2 + b11nlos3 + b12drgwt +b13drgwt2 +e

where

CAC is cost adjusted charges
npr is number of surgical procedures
npr2 is number of surgical procedures squared
los is DRG specific length of stay
poslos is (average los-los) if average los > los
neglos is (average los-los) if average los < los
nlos2 nlos3 are square and cubic terms of neglos
plos2 is squared term of poslos
drgwt is CMS drgwt
drgwt2 is drgwt squared

The parameters from this model were saved and then used them to impute estimated costs for HCUP. We tried alternative model specifications, including the log transformation of cost adjusted charges and excluding people with end stage renal disease (ESRD). In all of these alternative specifications, the parameters for the older people were remarkably similar to the parameters for the younger populations. We concluded that we could use the Medicare data to estimate the costs of younger hospitalized patients. The main advantage to this approach is that the Medicare data identify the number of days spent in intensive care (ICU). Because intensive care units are resource intensive and costly, being able to estimate this parameter was a key advantage.

For the FY01 - FY04 cost estimates, we used the 1999 MedPar file of veterans for estimating costs. For FY05-present, we have used the 2003 MedPar file. We have not been able to access more recent MedPar data until VA and CMS renegotiate the data use agreement.

4.4.2. Cost adjusted charges

Utilization databases, like the Healthcare Cost and Utilization Project (HCUP) or Medicare, report charges incurred in a hospital. Yet, it is generally known that health charges usually exceed the cost of providing care. However, the degree to which charges exceed costs is not completely random. Hospitals and medical centers are somewhat idiosyncratic in how they generate bills.

Hence, we want to adjust the charges for two reasons: (1) to deflate charges so that they more closely reflect costs, and (2) to remove hospital specific idiosyncrasies. The ratio of costs to charges (RCC), described in detail below, is one way of making this adjustment.

Adjusting charges with the RCC leverages information that every hospital annually reports to Medicare in the Medicare Cost Report. The Medicare Cost Report is a very large report that hospitals are required to complete if they want to receive federal reimbursement.

In the Medicare Cost Report, there are variables for each hospital's total charges and total costs. In the most recent Medicare Cost Report (PPS version 13), the field for charges is 2135 and the field for costs is 2138. We extracted these fields along with the hospital's Medicare identification number (PPS number). The quotient (i.e., the result of dividing costs by charges) was the ratio of costs to charges (RCC). The RCC usually ranges between 0.5 and 1.0. To actually adjust charges, the RCCs were linked to the Medicare dataset with the PPS number. The charge data were then adjusted by the RCC.

For example, if we want to use the RCC to adjust charges in a dataset, such as the HCUP dataset, we must first crosswalk the RCC dataset to the HCUP dataset. This can be a complicated process, especially for crosswalking the HCUP to Medicare (for details, see http://www.herc.research.va.gov/resources/faq.asp). Once we crosswalk the files, we then multiply charges by the RCC. Recall that the RCC is a hospital-specific adjustment. In other words, within any given hospital the RCC will be constant.

4.4.3. The dependent variable

We used cost adjusted charges as our dependent variable when we built the cost function. However, the cost adjusted charges from the Medicare data are not normally distributed.

Because of the skewness, we tried transforming the cost adjusted charges. While the log transformation helped reduce the appearance of skewness, the non-logged function consistently performed better than models with logged cost adjusted charges. Using logs presents additional hurdles because the estimated costs need to be transformed back to the original metric (dollars), adjusting for retransformation bias. The usual adjustment for retransformation bias is the smearing estimator.[7] While relatively simple to implement, this adds another layer of complexity to the entire process.

4.4.4. Length of stay

There are different ways to include length of stay in a cost function. The most obvious way is to include it without making any transformations, such that length of stay is a positive integer. Variations on this approach were also considered, such as a set of dummy variables representing different lengths of stay.

A second method for including length of stay involves comparing the patient's length to the average length of stay for all patients with that DRG. This second approach requires knowing the average length of stay for each DRG. This information is conveniently provided by CMS with the DRG weight file. We found slight advantages to the second approach as the transformation turned the length of stay from a positive integer into a continuous scale. Having a continuous scale provides slightly more ability to discriminate costs based on deviations in length of stay.

We used the second approach. In addition, we relaxed the constraints of our earlier estimates, allowing the cost of marginal days of stay to vary, depending on the length of stay.

Note that we examined only those records of patients discharged during the fiscal year under study. We included days of stay in acute medical-surgical bedsections, even if they occurred in previous fiscal years, and excluded data from stays that were not complete by the end of the fiscal year. This is distinct from the rest of our method, which considered only the days of stay that occurred during the fiscal year under study. We also calculated the length of stay in ICU bedsections. For each acute medical-surgical hospital stay, we found the number of days spent in the medical and surgical ICU bedsections.

4.4.5. Individual DRG intercepts or DRG weights

We found little marginal value in including dummy variables for each DRG. When we included DRG weight (squared and cubic terms), the gain in R^2 was less than 1%. Given the additional complexity in estimating this model, we decide not to use it. Instead, we decided to use DRG weight in our cost function along with the DRG weight squared and cubed. In the final model, we also interacted the Medicine Major Diagnostic Category (MDC) and Surgery MDC with length of stay.

4.4.6. Final model

The final cost function model based on a 50% sample of the Medicare data is shown in Table 5 and the methods have been peer reviewed.[8]

4.4.7. Outliers

Outliers can have undue leverage on a regression model. After we ran the model, we found that the model fit the data reasonably well. However, the fit was based primarily on the high cost users. The model did not fit as well for low-cost users, due in part to heteroskedasticity.

One solution involves removing or "trimming" outliers. We tried this and retested the model fit. Our methods and findings are below. We first identified outliers by using the Medicare outlier designation (n=1880). This did not help the fit of the model with low-cost cases because the outlier designation typically identifies the expensive cases.

Then we empirically identified outliers by generating Cooks' distance. Cooks' distance is the leverage of case *I* on the OLS regression coefficients (*∃hat*). It can be thought of as an F test comparing the beta coefficients with and without observation *I (i.e., ∃hat* to $\exists hat_{-I}$). Large values for Cook's distance suggest that the case has a lot of leverage.

We trimmed outliers in our regression models using three exclusion criteria:[1]

1) Cooks distance >0.001 (excluded 968 observations, ~0.8%)

2) Cooks distance >0.0001 (excluded 2,101 observations, ~1.7%)

3) Cooks distance >0.00001 (excluded 8,431 observations, ~6.6%)

---

[1] We also compared logged CAC models. In every case, the log models fit significantly worse and yielded much larger differences between estimated costs and actual costs.

We found that we could estimate better fitting models if some outliers were excluded. This gain was mainly within the lowest quartile of costs. Table 6 presents correlation coefficients between actual cost adjusted charges (CAC) and estimated cost adjusted charges. Note, however, that not always did removing more outliers lead to a better fitting model. In quartile 1, only model #3 yielded higher correlations.

Table 5: Full model based on 50% random sample of Medicare data (FY98-00)

Source	SS	df	MS
Model	3.8009e+13	27	1.4078e+12
Residual	1.3554e+13	321555	42152405.8
Total	5.1564e+13	321582	160343662

Number of obs	= 321583
F(27,321555)	=33396.73
Prob > F	= 0.0000
R-squared	= 0.7371
Adj R-squred	= 0.7371
Root MSE	= 6492.5

cac	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
died	2671.211	57.21167	46.690	0.000	2559.077	2783.344
sex	32.90875	61.21531	0.538	0.591	-87.0715	152.889
age	-34.22324	1.851834	-18.481	0.000	-37.85278	-30.5937
ndx	619.0444	81.09738	7.633	0.000	460.0959	777.993
ndx2	-146.7017	16.61743	-8.828	0.000	-179.2714	-114.1321
ndx3	10.97541	1.022981	10.729	0.000	8.970401	12.98043
los	104.255	9.083375	11.478	0.000	86.45187	122.0582
poslos	670.9503	10.10664	66.387	0.000	651.1415	690.759
neglos	182.4991	29.68224	6.148	0.000	124.3228	240.6755
nlos2	-109.8903	7.980714	-13.769	0.000	-125.5323	-94.24832
plos2	-.7170458	.021736	-32.989	0.000	-.7596478	-.6744437
nlos3	-4.587643	.5484962	-8.364	0.000	-5.66268	-3.512606
plos3	3.32e-06	.0000198	0.168	0.867	-.0000354	.000042
drgwt	4860.036	63.69243	76.305	0.000	4735.201	4984.871
drgwt2	-255.1638	11.0401	-23.112	0.000	-276.8021	-233.5255
drgwt3	12.97284	.5057919	25.649	0.000	11.98151	13.96418
surg	1069.883	78.21631	13.679	0.000	916.581	1223.184
surlos	-42.31538	11.16155	-3.791	0.000	-64.19169	-20.43906
pl_sur	421.5315	15.61753	26.991	0.000	390.9216	452.1415
nl_sur	328.304	36.252	9.056	0.000	257.2511	399.3569
pl_sur2	-1.384451	.1793446	-7.720	0.000	-1.735961	-1.03294
pl_sur3	.001167	.0006719	1.737	0.082	-.00015	.002484
nl_sur2	47.49814	8.419396	5.642	0.000	30.99636	63.99991
nl_sur3	3.636805	.55208	6.587	0.000	2.554745	4.718866
icudays	593.0367	7.165874	82.758	0.000	578.9918	607.0816
icudays2	10.27421	.2713893	37.858	0.000	9.742298	10.80613
icudays3	-.0325464	.0017843	-18.240	0.000	-.0360436	-.0290492
_cons	413.7664	181.3739	2.281	0.023	58.27884	769.254

We decided not to remove outliers because we realized any decision about which outliers should be removed would be arbitrary and would affect the model's fit. The full model fits almost as well (and better in some instances), therefore we saw little rationale for removing outliers.

Table 6 also shows how well the model predicts costs with the other 50% of the data (out of sample). In many cases, the out-of-sample predicted costs are quite close to the actual Medicare costs. As is shown in Table 5, the overall R^2 of the model is approximately 0.74.

Table 6: Correlations between estimated costs and actual costs for the full model and for three outlier restricted models

	Actual costs							
	Quartile 1: <$2605		Quartile 2: $2605<cac<$4484		Quartile 3: $4484<cac<$8472		Quartile 4: >$8472	
	In sample	Out of sample	In sample	Out of sample	In sample	Out of sample	In sample	Out of sample
Sample size	*38304*	*38144*	*39167*	*38594*	*39939*	*40801*	*43348*	*43286*
Model with all cases	correlation coefficients							
estimated costs	0.126	0.190	0.301	0.291	0.389	0.357	0.814	0.808
Restricted models								
(1)	0.057	0.204	0.309	0.005	0.396	0.250	0.641	0.699
(2)	0.071	0.209	0.313	0.011	0.398	0.279	0.718	0.749
(3)	0.185	0.202	0.313	0.305	0.393	0.392	0.769	0.775
Model estimated with log(CAC)	0.083	0.109	0.303	0.290	0.390	0.381	0.389	0.106

Notes: (1) cost function was estimated excluding cases with a cooks' distance >.001 (least restrictive)
(2) cost function was estimated excluding cases with a cooks' distance >.0001 (more restrictive)
(3) cost function was estimated excluding cases with a cooks' distance >.00001 (most restrictive)

4.5. Observation days

Beginning in 1997, VA created 7 new codes for observation bedsections to report inpatient care provided in observation units. Most stays involving these codes are recorded in the observation PTF files, which is a new set of files in the PTF. These stays, even if there are associated with an inpatient record in the Acute PTF file, are kept in a separate observation bed file at Austin. The structure of the observation files mirror the PTF inpatient files. We found that many stays reported in this file precede or follow stays in the acute medical-surgical PTF file. When calculating length of stay, some analysts will want to regard these observation days as part of acute medical-surgical stays.[2]

For the cost of observation bed stays, for FY98 onward we costed each day at the marginal cost of an additional day (i.e., $684). This method may underestimate the cost of stand-alone observation stays. Alternatively, it may overestimate the cost of an observation stay that preceded a hospitalization. We hope to develop and test new methods for costing observation bed stays in the future.

4.6. Negative or implausible costs

After estimating FY98 VA costs with the cost function, we found that the function had imputed negative costs for 2,974 of the 541,567 (0.6%) acute medical-surgical hospitalizations. This is because the cost function was not constrained to predict non-negative estimates. Therefore, rare combinations of right-hand-side variables can lead to negative predictions. These 2,974 records were assigned the cost of a marginal day of stay ($684.75).

The cost of a marginal day of stay was calculated in a simulation with the 1996 Medicare data. Adjusting for all other covariates in a linear regression, we identified the cost for an additional day of stay. Holding all other factors at their mean, if a person stayed an additional day, they had an additional $684.75 of cost adjusted charges.

While some stays were not assigned negative costs, they were given very low costs. For instance 42 hospital stays had positive costs less than $5. We decided that any stay with a cost less than $684.75 was implausibly low and an artifact of the cost function. By setting this rule, it effectively set a floor on the estimated cost per stay. A total of 9,632 (2%) cases had non-negative costs less than $684.75. These cases were all given $684.75 per day (86% had a length of stay of one day). In the future, we will explore other methods for determining the cost of these cases, including setting constraints on the cost function.

4.7. Reconciling to the Cost Data

The cost function is based on non-VA relative value weights and non-VA cost adjusted charges. The estimated costs must be reconciled to VA costs. Reconciliation can happen at many levels including the department, medical center, and nationwide. We chose to reconcile the estimated

---

[2] Nearly 73,000 days of stay were assigned to observation bed sections in FY99 (out of 13.5 million days in VA hospitals). Most observation stays were one day long, but this was not always the case. Most observation days were in medicine, surgery, and psychiatry observation bedsections. We examined the FY99 data and found that 19,428 (26%) of the observation stays immediately preceded a stay reported the PTF bedsection files. Another 319 observations stays followed stays in the bedsection file. (Our analysis was limited to PTF bedsection file. It is also possible that observation stays precede or follow stays reported in the PTF extended care file.)

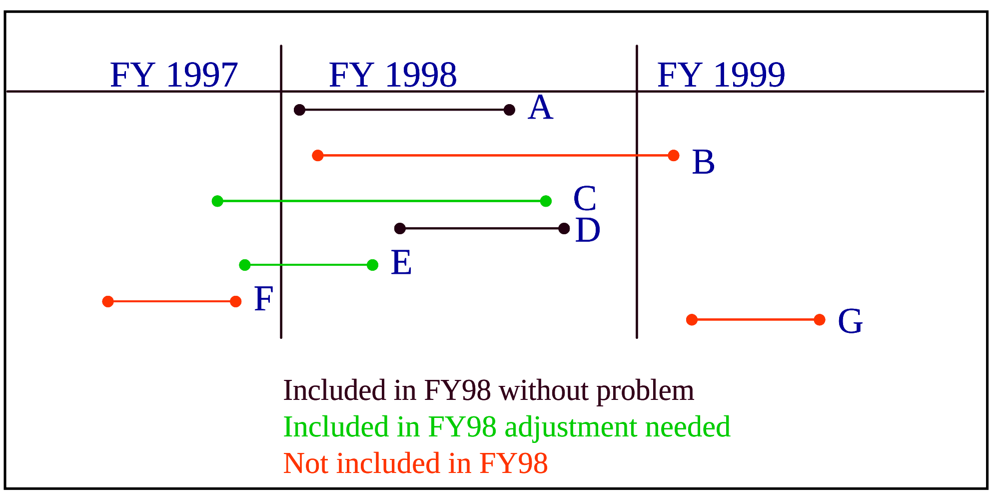
costs to the medical center and nationwide; we decided not to reconcile the estimated costs to the department. Given that the VA cost data and PTF are not reconciled against each other, our concern was that there would be too much variability in department-level costing.

Reconciling the costs to the medical center results in “local” cost estimates, while reconciling the costs for the entire VA results in “national” cost estimates. Therefore, this process results in the creation of 2 VA cost estimates: a local cost estimate (costl) and a national cost estimate (costn).

The logic behind reconciling the costs is straightforward. For the local cost estimate we sum together the estimated costs for a medical center and divide this amount by the total acute medical-surgical care costs (acute medicine and surgery) for the medical center. The quotient of this division is a scaling factor. By multiplying the estimated cost by this scaling factor, we ensure that the sum of the estimated costs is equivalent to VA costs.

Unfortunately, the reconciliation is easier said than done. Recall that CDR and DSS report costs for the fiscal year while the acute medical-surgical hospitalization data represent discharges. For FY98 data, some stays that ended in FY98 started before FY98. At the same time, there were people hospitalized in FY98 who were still in the hospital at the end of the fiscal year and are not reported in the FY98 PTF data. To illustrate this point, Figure 1 shows the hospitalization that cross the fiscal years. Cases B, C, and E all cross the fiscal years. It is not correct to assume that the cases crossing from FY97 to FY98 are equivalent in number to those cases crossing from FY98 to FY99. Due to the declining trend in inpatient hospitalization, C and E are more common than B.

Figure 1: Difference between FY view and discharge view



Note: A & D are in the med/surg file and need no adjustment
C & E are in the med/surg file and need adjustment
B, G, and F are not in the med/surg file

If no adjustment were made for this fact, then we would overestimate the number of hospitalizations, and thereby underestimate the cost of care per hospitalization. Our correction for this was to adjust the cases discharged in the fiscal year that started before the fiscal year. The FY98 adjustment factor was found by comparing the FY98 Census to the FY 97 Census (see Table 7).

After adjusting the discharge data so that it better represented the FY costs in the cost data, we reconciled the estimated costs. The national scaling factors are listed in Table 7. We multiplied every estimated cost by this scaling factor to obtain the national VA cost.

Table 7: Fiscal year adjustment and scaling factors

Fiscal year	Fiscal year adjustment	National scaling factor
FY98	0.93	1.27
FY99	0.9821	1.29
FY00	0.9290	1.41
FY01	1.0442	1.21
FY02	0.9117	1.20
FY03	1.0290	1.21
FY04	0.9990	1.26
FY05	0.9059	1.56
FY06	1.0339	1.65
FY07	1. 0061	1.81
FY08	1.1519	1.84
FY09	1.0312	1.63

4.8. Stability of the cost function over time

The cost function for FY98-FY00 was built using 1996 Medicare data. For FY01 - FY04, we used 1999 Medicare data. For FY05+, we used 2003 Medicare data. One question is whether the cost-function is robust to the input data that are being used. To answer this question, we used 1994 and 1995 MedPar data that was similar to the 1996 MedPar data. We then ran the identical cost function on all three datasets. The model coefficients from the three datasets were compared. Finally, using the regression model for each year of data, we predicted costs in 1996, using the MedPar 1996 as the criterion. We compared the estimated costs to see if differences would have occurred had they been estimated with 1994 or 1995 MedPar data.

The regression coefficients for all three models were extremely similar (Table 8). The predicted costs from the three models were also highly correlated (>0.99; Table 9). The results suggest that the cost function is highly robust to the year from which the MedPar data are used.

Table 8: Stability of regression coefficients with 1994, 1995 and 1996 MedPar data

	1994		**1995**		**1996**	
	Coeff	t-stat	coeff	t-stat	coeff	t-stat
died	2837.70	42.410	2803.32	42.650	2671.21	46.690
sex	-41.01	-0.560	-28.73	-0.400	32.91	0.540
age	-42.29	-18.590	-44.42	-19.720	-34.22	-18.480
ndx	250.36	2.740	433.47	4.710	619.04	7.630
ndx2	-80.63	-4.190	-117.71	-6.180	-146.70	-8.830
ndx3	7.44	6.150	9.60	8.120	10.98	10.730
los	50.63	4.660	52.01	4.890	104.26	11.480
poslos	656.08	54.620	666.76	54.250	670.95	66.390
neglos	272.94	9.400	338.59	11.140	182.50	6.150
nlos2	-72.45	-11.940	-71.91	-10.220	-109.89	-13.770
plos2	-1.31	-54.080	-0.62	-10.450	-0.72	-32.990
nlos3	-1.41	-4.830	-1.85	-4.490	-4.59	-8.360
plos3	0.00	30.680	0.00	2.900	0.00	0.170
Drgwt	4477.58	58.500	5149.17	69.610	4860.04	76.300
drgwt2	-161.85	-12.100	-325.22	-25.390	-255.16	-23.110
drgwt3	8.02	13.030	16.71	28.480	12.97	25.650
surg	470.37	5.280	526.47	5.890	1069.88	13.680
surlos	-48.96	-3.770	-23.43	-1.810	-42.32	-3.790
pl_sur	416.50	26.280	379.25	22.240	421.53	26.990
nl_sur	222.54	5.670	152.01	3.850	328.30	9.060
pl_sur2	-1.21	-24.520	-0.95	-8.300	-1.38	-7.720
pl_sur3	0.00	18.310	0.00	-1.250	0.00	1.740
nl_sur2	18.26	2.590	3.07	0.390	47.50	5.640
nl_sur3	0.58	1.900	0.72	1.710	3.64	6.590
icudays	395.04	47.070	553.12	67.840	593.04	82.760
icudays2	18.93	58.260	9.29	31.130	10.27	37.860
icudays3	-0.08	-37.720	-0.02	-11.440	-0.03	-18.240
_cons	1819.08	8.640	1416.06	6.650	413.77	2.280

Table 9: Pair wise Correlations in predicted costs compared to 1996 costs adjusted charges

cost94	1		
cost95	0.993	1	
cost96	0.997	0.996	1
CAC 1996	0.856	0.855	0.859

Note: CAC is cost adjusted charges

5. User's Guide

This chapter discusses how to use HERC's average cost dataset. The chapter covers four topics: 1) a brief summary of the methods, 2) assumptions underlying the dataset, 3) how to correctly use the dataset, and 4) when not to use the dataset. Although we hope that these data will be useful, we do not expect that they will be appropriate for every study. For this reason, later in this chapter we discuss limitations with these data and instances where these data are not appropriate. Appendix 3 includes the contents of the HERC Average Cost data.

5.1. Summary of methods

5.1.1. Categories of inpatient care

We categorize inpatient care into eleven categories: 0) acute medicine, 1) rehabilitation, 2) blind rehabilitation, 3) spinal cord injury rehabilitation, 4) surgery, 5) psychiatry, 6) substance abuse care, 7) intermediate medicine, 8) domiciliary, 9) nursing home care, and 10) psychosocial residential rehabilitation programs (PRRTP). These categories are defined by bedsection / treating specialty codes (see Table 4). PRRTP care can only be provided at approved medical centers. If a non-approved medical center had dollars or days in PRRTP bedsections, these were allocated back to psychiatry and substance abuse care, respectively.

5.1.2. Acute medical-surgical care

For acute medicine and surgery, we estimated costs using a cost-function from Medicare MedPar data restricted to Veteran users. To do this, we developed a VA acute medical-surgical dataset using the PTF bedsection file. Contiguous acute medical-surgical bedsection stays were aggregated into a single record. In the cost function, length of stay was entered into the model as the deviation from the expected length of stay for that DRG. We also used DRG weight as the measure of relative weight, rather than allow each DRG to have its own intercept.

For each observation day in an acute medicine or surgical bedsection, we costed it at the marginal cost per day, which we estimated at $684.75. The cost function yielded some negative and implausible costs. We set $684.75 (the marginal cost of a day), as the minimum cost possible.

Lastly, we reconciled the estimated costs to the medical center's and overall VA's costs. This yielded a local cost estimate (costl) and a national cost estimate (costn).

5.1.3. Non medical/surgical categories

Rehabilitation, psychiatry and long term care costs were estimated using a daily rate. For FY98-FY00, nursing home costs are case-mix adjusted. Since FY00, nursing home costs have been based on an unadjusted per diem.

5.2. Assumptions in the average cost dataset

Throughout this document we have tried to identify assumptions underlying the creation of the acute medical-surgical and non medical/surgical datasets. These data include indirect costs and physician costs; excluded are the cost of capital financing and malpractice. Table 10 shows the included and excluded costs.

Table 10: Included and excluded costs

Type	Notes
Excluded	
Capital financing costs	Not included, but this may be noteworthy (5%).
Malpractice expenses	Not included.
Contract provider costs	Excluded are contract services because these costs are not accurately associated with units of care
Community nursing home costs	Beginning in fiscal year 2006 (FY06), the VA introduced a set of new codes for categorizing treating specialties/bedsections for nursing home care. CNH is now identified by BEDSECN 80 or 44 and STATYP 42. Treating specialty code 44 became active on 7/1/2006, while treating specialty code 80 became inactive on 8/2/2006.
Headquarters costs	Excluded are the costs associated with VA headquarters
Prosthetics	Inpatient prosthetics billed separately are not included in the CDR accounts
Included	
Costs for physician services	These costs are included in the CDR. For every stay, physician costs are proportionate to the hospital costs.
Research & education	Included to the extent supported by the VA medical care appropriation.
Indirect costs	We assigned indirect costs to each CDA in proportion to its share of the total direct costs of its group of CDAs.

5.2.1. Data used in the cost function

The average cost estimates for acute medical-surgical stays were based on a cost function that was constructed with Medicare data. The cost function for FY98-FY00 was built using 1996 Medicare data. For FY01 - FY04, we used 1999 Medicare data. For FY05+, we used 2003 Medicare data. The Medicare data represented veteran users; excluded were cases in Hawaii, Alaska and cases related to labor and delivery. In using the Medicare data we assumed that the underlying accounting systems for non-VA hospitals could be used to impute estimates for the VA. These imputed estimates were then reconciled with the VA costs.

5.2.2. The cost of observation stays

Observation stays are a relatively new type of service provided in the VA. There is no analogous type of service provided in the private sector. To estimate the cost of the observation bed stay, we estimated a marginal daily rate and multiplied this times the length of stay. Most people stay in the observation bed for one day; a few outliers stay longer and in these cases, the cost is equivalent to this rate times the length of stay. To calculate the daily rate for observation bed stays, we developed a regression model using Medicare data. With the regression model, we simulated the marginal cost at the mean of data. We then predicted the cost if the person stayed one day longer than the mean. The difference between these two estimates was $684.75. We used this as the daily rate for the observation bed stays.

5.2.3. Costs for high and low-cost procedures

We used a cost function to estimate acute medical-surgical costs, and this method is more accurate with high-cost cases than low-cost cases. If you are assessing cases that typically have very low costs, then the average cost provided in the HERC dataset may be inappropriate. Our method does not account for very expensive inpatient procedures that are not captured by the DRG or LOS variables.

5.2.4. Implicit trimming of outliers

A byproduct of using a statically-based cost function is that the predicted costs have less variability than the true data—the method removes many of the outliers. Recall that the cost function is a linear regression model. When we calculated the cost for the VA we used the regression model to estimate costs based on averages. If you are interested in high or low-cost outliers, then the HERC dataset may be inappropriate for your use.

5.2.5. Model estimates and negative costs

Another byproduct of using a cost function is that after we imputed the VA costs we had some cases with negative or implausibly low costs. Clearly, a stay cannot have a negative cost. Therefore, we decided that we would set a floor. Any choice of a floor is somewhat arbitrary, but we chose the floor to be $684.75. Recall that $684.75 is the average cost of an additional day of stay. When you use the HERC Average Cost data, compare the length of stay to the cost. If you believe your data have low cost cases, then you may want to use other values in a sensitivity analysis.

5.2.6. VISN administrative costs

Each of the VISNs incurs administrative operating costs. We have included these costs under the assumption that they cover coordination expenses required for a large health provider. From our perspective, these costs should be distributed to all medical centers in the VISN, and it is not clear that this always happens. This may partly explain discrepancies in local costs, and if your study requires local costs, then use them carefully.

5.3. Using the average cost dataset

At Austin, we have provided three datasets. These datasets are listed in Table 11 and described below. All of the files can be found in the RMTPRD.HERC.SAS directory.

Table 11: The three average cost datasets for FY98

Dataset	Includes	Excludes
dischgXX	All persons admitted since FY98 and discharged in fiscal year. Costs for all care	Stays not completed by end of fiscal year Stays admitted before beginning of FY98 (10/1/97)
mdsrgXX	All persons discharged from an acute medical-surgical bedsection in fiscal year	Non medical-surgical bedsections People who were still in the hospital at end of FY.
nmdsrgXX	The cost of care provided in rehabilitation, mental health or long-term bedsections during the fiscal year.	The costs of care provided before the fiscal year are excluded.

Table 12: Using the three average cost datasets

Dataset	Sort and merge using	Merge data to
dischgXX	SCRSSN, ADMITDAY, DISDAY, and STA3N.	PTF main files (PM, XM and PMO)
mdsrgXX	SCRSSN, ADMITDAY, DISDAY, STA3N, and BSOUTDAY.	PTF bedsection files (PB, XB, PBO); BUT must first aggregate the bedsection file
nmdsrgXX	SCRSSN, ADMITDAY, DISDAY, STA3N, BSINDAY, and BSOUTDAY.	PTF bedsection files (PB, XB, PBO), and PTF census files.

5.3.1. Discharge dataset

Combining the acute and non-acute datasets yielded the discharge dataset. It represents a discharge dataset, and as such it only has cases that were discharged in the FY. In addition, only people admitted since the beginning of FY98 are included in the discharge datasets. Patients that were admitted prior to FY98 are excluded

The discharge dataset includes additional variables that track cost subtotals, length of stay subtotals, DRG weight, and ICU days.

Table 13: Discharge dataset

Scrssn	Numeric field. Identifies a patient's scrambled social security number
sta3n	3-digit numeric field. Represents the VA medical center's station number. These can change when facilities merge.
Adtime	Admission time for an inpatient stay.
Admitday	Admission day for an inpatient stay (SAS date)
Disday	Discharge day for an inpatient stay (SAS date).
b4fy98	Flag that identifies inpatient stays that began prior to FY98. The numeric variable is either 0 or 1. In FY03, we started providing costs of rehabilitation, mental health and long-term care for these discharges. Note these costs are incomplete and exclude costs prior to FY98 and any med/surg care.
costl	Total local cost. Represents the entire cost of the stay, reconciled with the local medical center's expenditures.
costl_0	Local cost for medicine and surgery
costl_1	Local cost for rehabilitation
costl_2	Local cost for blind rehabilitation
costl_3	Local cost for spinal cord injury
costl_4	Does not exist; this category is included with 0
costl_5	Local cost for psychiatry
costl_6	Local cost for substance use treatment
costl_7	Local cost for intermediate medicine
costl_8	Local cost for domiciliary
costl_9	Local cost for nursing home care
costl_10	Local cost for psychosocial residential rehabilitation treatment programs
costn*	Total national cost. Represents the entire cost of the stay, reconciled with expenditures from all VA medical centers. Same categories as local costs.
los*	Length of stay overall and for the different categories of care. Same categories as local cost.
Flag	An indicator for local costs that deviate +/- 2 standard deviations from the national costs.
Flagnh	A flag for community nursing home. HERC does not estimate the community nursing home costs. Other costs may be reported for these individuals if they were transferred to a facility.
Flagext	A flag to identify cases where the costs were recalculated because HERC length of stay differed from PTF main length of stay.

A single discharge record provides important subtotals. For example, if a researcher is interested in mental health costs, he/she can now identify the mental health costs for every inpatient encounter. This is particularly helpful for those patients who receive care in many different categories during a stay. Again, note that these changes only pertain to the inpatient discharge datasets.

5.3.2. Acute medical-surgical dataset

This dataset is best described as a discharge dataset for persons who were discharged or transferred from an acute medical-surgical bedsection in the fiscal year. The key to understanding this dataset is that we aggregated the bedsection files to make a discharge file that is analogous to the MedPar dataset.

The first step of the process involved identifying acute medical-surgical bedsections. If, during a stay,[3] a person was in three acute medical-surgical bedsections, we combined these bedsections. Transfers within acute medical-surgical bedsections, such as from surgery to medicine, were aggregated into a single record. We adopted the rule that if a patient was transferred from an acute medical-surgical bedsection to another acute medical-surgical bedsection that this would be considered part of the same acute medical-surgical stay. Similarly, if a person was transferred from an acute medical-surgical bedsection to a non-medical/surgical bedsection, we ruled that the acute medical-surgical stay had ended. Transfers from an acute medical-surgical bedsection to a non-medical/surgical bedsection and back to an acute medical-surgical bedsection were treated as one non-medical/surgical and two acute medical-surgical stays.

You will want to link this file to the PTF bedsection files. But before you merge those files with this cost file, you will need to aggregate the bedsection file. Please contact HERC if you would like an electronic version of this SAS code.

[3] Stays were defined by five variables: scrssn, sta3n, admitday, adtime, disday.

Table 14: Variables in the Medical Surgical Dataset

Scrssn	Numeric field. Identifies a patient's scrambled social security number
sta3n	3-digit numeric field. Represents the VA medical center's station number. These can change when facilities merge.
Adtime	Admission time for an inpatient stay.
admitday	Admission day for an inpatient stay (SAS date)
Disday	Discharge day for an inpatient stay (SAS date).
bsoutday	Discharge day for the bedsection
bsinday	Does not exist; creating the dataset alters this variable. If you really need it, consider making a pseudo-bsinday by subtracting LOS from the bsoutday. However, this may be imperfect for merging.
Source*	numeric field that identifies the source of the data. 1=XB census 2=XB discharge 3=PB census 4=PB discharge 5=OBS discharge 6=OBS census
Los	Length of stay.
Drgwt	Diagnostic related weight created by Centers for Medicare and Medicaid Services for reimbursing inpatient Medicare stays. Numeric field.
Icudays	Length of stay in the ICU; 0 if none.
Drg	Diagnostic related group created by Centers for Medicare and Medicaid Services for reimbursing inpatient Medicare stays. Each group has an associated drgwt– see above. Numeric field.
costl	Total local cost. Represents the entire cost of the stay, reconciled with the local medical center's expenditures.
Costn	Total national cost. Represents the entire cost of the stay, reconciled with expenditures from all VA medical centers.
Flag	An indicator for local costs that deviate +/- 2 standard deviations from the national costs.

*Not included after FY04.

5.3.3. Rehabilitation, mental health or long-term dataset

This dataset contains costs for people who had a non medical-surgical stay. Only costs for stays during the fiscal year are included. If a person was admitted and discharged in FY05, then the total cost of their stay is in the FY05 dataset. However, if a person was admitted in FY04 and discharged in FY05, then only costs for the portion of the stay during FY05 is reported in the FY05 dataset. One of the reasons for doing this is that there are some people in long-term care who have been there for 30+ years. It would be extremely difficult to identify the entire cost of these stays. For information on costs prior to FY98, see HERC Technical Report 1.[1]

Table 15: Rehabilitation, Mental Health and Long-Term Care Dataset

scrssn	Numeric field. Identifies a patient's scrambled social security number
sta3n	3-digit numeric field. Represents the VA medical center's station number. These can change when facilities merge.
adtime	Admission time for an inpatient stay.
admitday	Admission day for an inpatient stay (SAS date)
disday	Discharge day for an inpatient stay (SAS date).
bsoutday	Discharge day for the bedsection
bsinday	Admit day for the bedsection
bedsection	Lists the bedsection of the treating physician. For more information see www.virec.research.med.va.gov
Lsb	Length of stay in bedsection.
distype	Type of discharge; identifies death in hospital. See www.virec.research.med.va.gov
source	numeric field that identifies the source of the data. 1=XB census 2=XB discharge 3=PB census 4=PB discharge 5=OBS discharge 6=OBS census
Cat	HERC category of care 0= Medicine and Surgery 1= Rehabilitation 2= Blind rehabilitation 3= Spinal cord injury 4= Surgery (category does not exist; we combined it with 0) 5= Psychiatry 6= Substance use treatment 7= Intermediate medicine 8= Domiciliary 9= Nursing Home 10= Psychosocial residential rehabilitation programs
Drg	Diagnostic related group created by Centers for Medicare and Medicaid Services for reimbursing inpatient Medicare stays. Each group has an associated drgwt– see above. Numeric field.
costl	Total local cost. Represents the entire cost of the stay, reconciled with the local medical center's expenditures.
costn	Total national cost. Represents the entire cost of the stay, reconciled with expenditures from all VA medical centers.
Flag	An indicator local costs that deviate was 2 standard deviations from the national costs.

5.3.4. Flag

An important variable is the flag variable. This variable indicates when the local cost estimate (costl) is > 2 standard deviations above or below the national cost estimate. Flag is an indicator or dummy variable; use the costl with caution when the flag variable is one.

5.4. When not to use the average cost dataset

5.4.1. Effects not detected in this cost estimate

It is not always appropriate to use the Average Cost data in your analysis. The average cost method assigns the same cost to all inpatient stays with the same demographic and discharge information. Stays that have the identical characteristics will have the same cost. If you are interested in assessing the cost consequences of a new procedure, then these data are likely to be inappropriate unless the cost of the procedure is entirely reflected by variables in the cost function. If the procedure saves money, but it does not affect one of the variables in the cost function, such as DRG weight or length of stay, then these stays will all get the average cost.

For example, let us assume that we had a new procedure for transfusing blood during a heart transplant. We are interested in whether this new procedure saves money. First, let us assume that this intervention would not affect the patient's DRG. In this case, it is also likely that the intervention would not affect other variables in the cost function, such as length of stay. Therefore, the estimated cost of care for people who received this new procedure would be the same estimated cost of care for people receiving the usual therapy. This does not mean that there was not a cost difference from this new therapy. It only means that any differences were not reflected in the HERC Average Cost data.

5.4.2. Comparison of medical center efficiency

The economic definition of efficiency is to use fewer inputs to make the same level of output, or conversely, to use the same number of inputs to make more output. The relative value weights we use DO NOT capture differences in the quantity or price of the inputs. In addition, the CDR costs (FY98-FY03) and DSS costs (FY04+) exclude the cost of capital financing. Finally, we distribute other short-term fixed costs in proportion to the variable costs. Although these issues may not be critical for cost-effectiveness analysis, they may be problematic and potentially fatal for efficiency analysis.

5.4.3. Point estimates versus variance estimates

We believe the average cost method produces relatively accurate point estimates for the costs. However, a consequence of estimating costs with a cost function is that the variance of the estimated costs is biased downwards. The reason for this is that many factors that affect costs are not included in the cost function, and if the stays are identical on all observed factors then these cases receive the same estimated cost. In Table 16 we show the costs reported by Medicare (1996) for five DRGs. We also show the estimated costs from our cost function (estcost). As is clear from this Table, the standard deviation is smaller in the estimated costs. Also, the minimum and maximum are attenuated toward the mean.

Table 16: The cost function's effect on the variation of the estimated costs

	Obs	Mean	Std. Dev	Min	Max
DRG14 Specific cerebrovascular disorders except TIA					
Cost	10534	6829	7587	7	175346
estcost	10534	7377	7476	685	147135
DRG79 Respiratory infections & inflammations age >17 w cc					
Cost	7767	7923	8445	16	213967
estcost	7767	8210	6423	685	198091
DRG88 Chronic obstructive pulmonary disease					
Cost	15428	4786	5525	5	203877
estcost	15428	4535	4269	685	128695
DRG89 Simple pneumonia & pleurisy age >17 w cc					
Cost	12905	5468	8863	8	662916
estcost	12905	5238	4675	685	160280
DRG127 Heart failure & shock					
cost	21463	4941	4979	10	109945
estcost	21463	5224	4479	685	190673

Note: cost is cost adjusted charges and estcost is the estimated cost adjusted charges.

If you are interested in evaluating the variation of these cost estimates, then use the Average Cost data carefully. If you use these cost estimates in a statistical model, most statistical tests will be biased toward the null. If you are trying to identify cases on the fringe of the cost distribution (high or low), then you will almost certainly miss some using these data.

5.5. Duplicates

Researchers who want to merge VA utilization data to our average cost estimates need to be aware that the PTF files have duplicates. There are duplicates within each file (e.g., PB discharge file) and between files (e.g., PB discharge file and XB discharge file). We excluded duplicates when we created the average cost datasets, and then we added the duplicate records back into the dataset to ensure the data had the same number of records. These duplicates have missing costs, so they can be easily excluded.

When merging records:

1) drop HERC records with missing values. This includes duplicate records, community nursing home records and patients admitted prior to FY98.

2) Delete duplicates from the Austin data that you are working with. One way to do this is to run the following command in SAS. Note that these commands only identify records that have duplicate values of the sort variables. The records may differ in other respects.

```
proc sort data=<indata> out=<outdata> nodupkey;
  by scrssn admitday adtime disday sta3n bsinday bsoutday;
```

References

1. Barnett PG, Chen S, Wagner TH. Determining the Cost of VA Care with the Average Cost Method for the 1993-1997 Fiscal Years. *HERC Technical Report 1*. 2000.

2. Wagner TH, Chen S. An economic evaluation of inpatient residential treatment programs in the Department of Veterans Affairs. *Med Care Res Rev.* Apr 2005;62(2):187-204.

3. Yu W, Wagner TH, Chen S, Barnett PG. Average cost of VA rehabilitation, mental health, and long-term hospital stays. *Med Care Res Rev.* Sep 2003;60(3 Suppl):40S-53S.

4. Barnett PG. Research without billing data. Econometric estimation of patient-specific costs. *Med Care.* Jun 1997;35(6):553-563.

5. Averill RF, McGuire TE, Manning BE, et al. A study of the relationship between severity of illness and hospital cost in New Jersey hospitals. *Health Serv Res.* Dec 1992;27(5):587-606; discussion 607-512.

6. Calore KA, Iezzoni L. Disease staging and PMCs. Can they improve DRGs? *Med Care.* Aug 1987;25(8):724-737.

7. Duan N, Manning Jr W, Morris C, Newhouse J. A Comparison of Alternative Models for the Demand for Medical Care. *Journal of Business & Economic Statistics.* 1983;1(2):115-126.

8. Wagner TH, Chen S, Barnett PG. Using average cost methods to estimate encounter-level costs for medical-surgical stays in the VA. *Med Care Res Rev.* Sep 2003;60(3 Suppl):15S-36S.

Appendices

Appendix 1

Table A1: Reconciliations for FY98-FY09.

sta3n	Old cat	98	99	00	01	02	03	04	05	06	07	08	09
402	6	5	5	5	5	5	5						
405	3				7								
436	7		9										
437	6				5								
438	6	5	5	5	5								
442	5			7									
452	1	7											
452	3			7									
452	6	5											
459	0				9								
459	4				9						9		
459	7	9											
463	0				8	8		8					
463	9						8						
500	2	1											
500	7	9											
503	4						0		0	0			
504	6	1											
504	7		9										
506	1		9										
508	6	5	5	5	5	5	5						
508	7			9									
509	6		5	5	5						5		
512	1		9	9	9	9							
515	1		9										
515	6	5	5										
515	8	9	9				9						
516	1	9	9	9	9	9							
516	2		9										
516	6	5	5	5	5	5							
518	0					8			9				
518	6					5							
520	6					5							
520	8									9			
521	3		2	2									
521	8											0	
521	9						2						
523	7	5											
526	1		9	9	9	9	9						
526	6	5							5				
528	1		9										
528	3			1									
528	6			5		5	5						
529	4		9										
529	7	9											

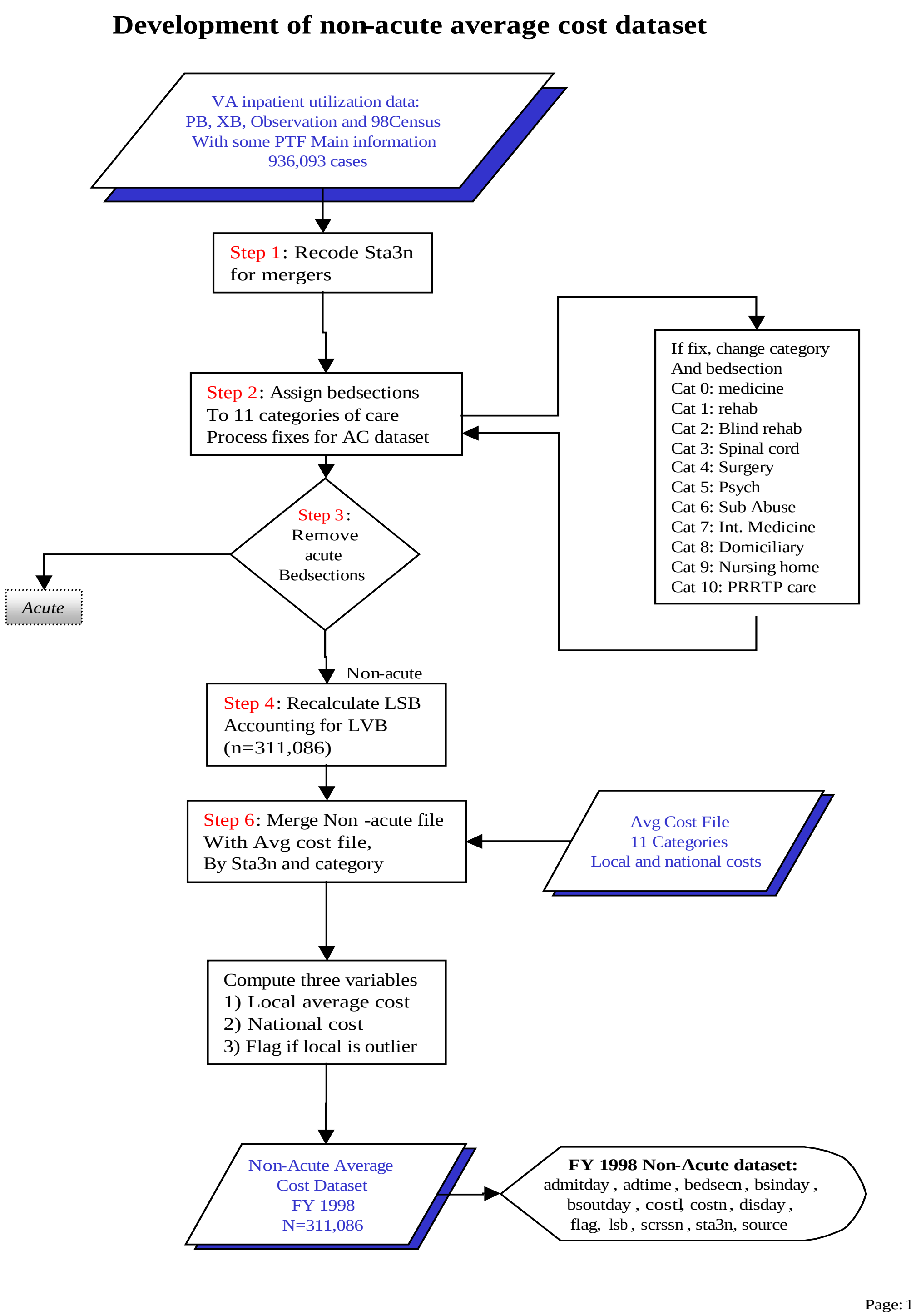
sta3n	Old cat	98	99	00	01	02	03	04	05	06	07	08	09
531	3			9									
531	6			5									
534	9	0											
537	2	1											
537	7					5	5						
537	8								1				
538	6	5											
539	1		9										
540	9						7						
541	7						9						
542	6				5								
543	7		9										
544	7					9		9					
546	2					1							
546	6							5					
549	1				9								
549	6	5	5	5									
550	6					5	5						
552	1						9						
552	6						5						
552	7						9						
553	1		9										
554	7		9	9	9	9	9						
555	1	9	9										
555	6	5	5										
555	10	5											
556	4							0	0				
556	6	5	5	5	5	5							
556	7					9							
557	1		9	9									
557	6				8	8							
558	6	5											
558	7	9	9	9	9	9	9						
561	2	1											
561	6		5	5	5		5						
562	5	7											
564	9											7	
567	0				9								
570	6				5	5	5						
570	7					9	9						
573	6	5											
573	10	5											
578	7					9							
580	6						5						
581	5	7									7		
583	9									7			
585	6			5	5	5	5						
586	6					5	5						5
586	7						9						
586	1										9		
589	6	5	5			5							
589	7	5	9	5									
590	6				5								

sta3n	Old cat	98	99	00	01	02	03	04	05	06	07	08	09
596	6												5
596	1			9									
596	7										9		
597	6		5										
598	6	5			5								
600	6						5						
603	1	7											
603	3	7											
603	9						7						
605	1	9	9	9	9	9	9						
608	0											9	
608	1			9	9								
608	4			0	0	0	0						
608	7						9						
609	7	9											
610	1		9	9									
610	3		0										
610	4			0	0	0	0						0
610	7	9	9										
612	1			9									
612	5			9	9							9	
612	7	9	9	9									
614	1	7					7						
614	9					7	7						
619	6	5	5										
619	8	9	9	9									
620	4	0											
620	6				5								
620	10				5								
621	6	5	5	5	5	5	5						
621	7			9									
622	6	5											
623	1		7										
623	9						7						
626	6					5	5						
629	8				9								
629	0									4		9	
630	6			5	5								
631	6	5	5	5									
631	7					9							
631	8					9							
632	2	1				9	9						
632	6				5	5							
635	6			5									
636	1			9	9	9	9						
637	6					5							
642	1		9	9									
642	6	5	5	5		5	5			5			
642	7	9				9							
644	1	9	9	9	9								
644	3				9								
644	6					5							
646	6	5											

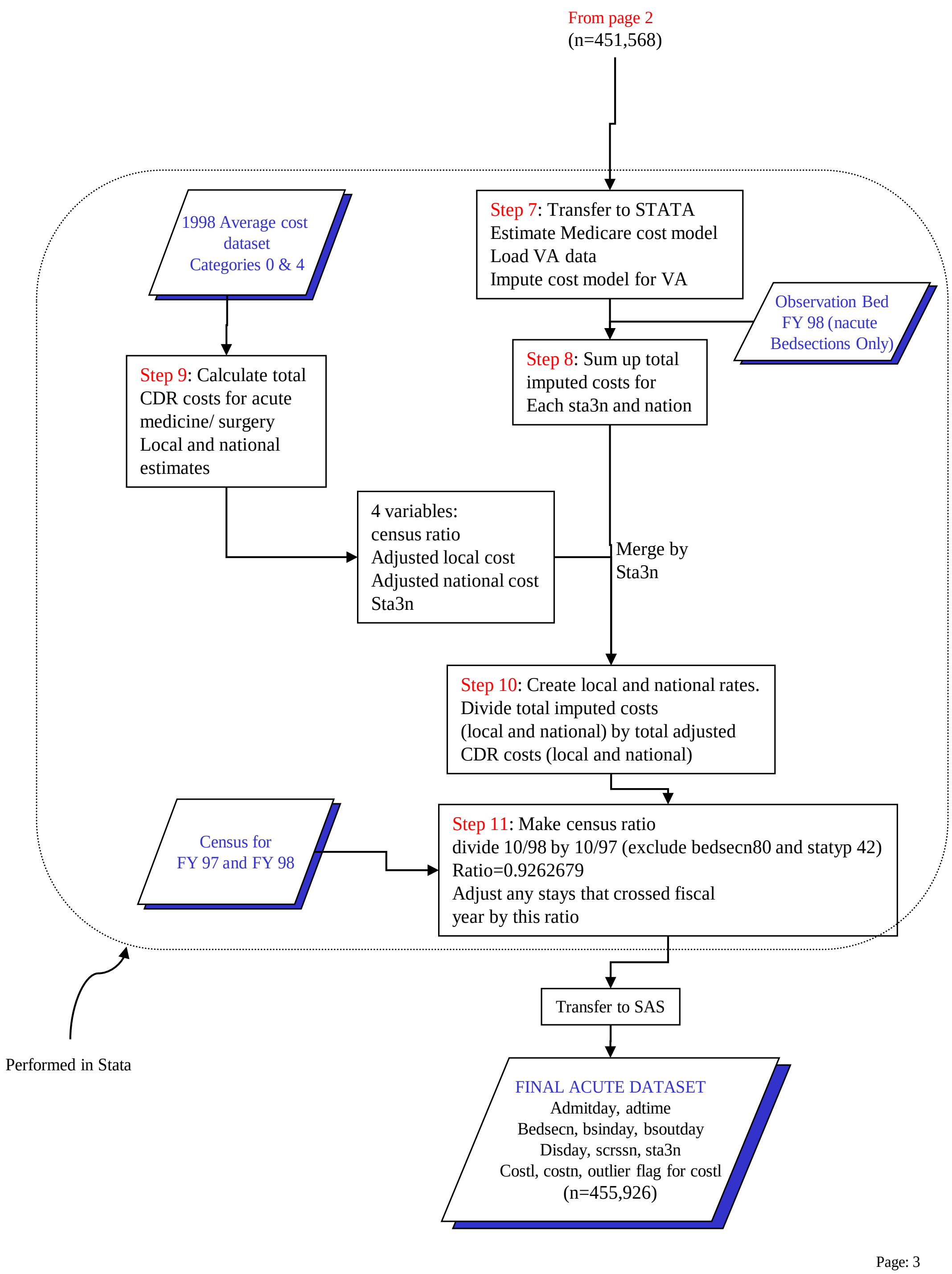
sta3n	Old cat	98	99	00	01	02	03	04	05	06	07	08	09
647	5	9											
648	1						9						
648	8						9						
648	2									9			
649	1							9					
652	6			5									
653	6		5	5	5	5	5						
654	1			9	9	9	9						
654	6			5	5	5	5						
655	3		0										
655	4	0		0	0	0	0						
656	6			5									
656	7				9								
657	1							9	9				
657	6				5								5
658	6		5										
660	9					7	7						
662	2						9						
662	3					9							
662	7					9	9						
662	8					9		9					
662	10						5						
664	1	9			9								
664	6		5	5	5	5	5						
667	7	5											
668	1			9	9								
668	6	5											
670	3		1										
671	1		9	9					9				
671	7		9	9	9	9							
672	7						9						
673	6					5							
674	1	9	9	9	9	9	9						
674	6	5	5	5	5								
678	1				9								
678	6	5	5	5	5	5	5						
678	7		9	9		9							
679	0				7	7							
679	1				9								
679	8	9	9										
687	4	0											
688	6	5	5			5	5						
689	3						7		7	7			
689	7	9	9	9									
689	8						9						
691	3	1											
691	6	5	5										
692	0	8	8	8	8	8							
692	9						8						
693	6					5	5						
695	7					9							

Note: if the cell is blank for a new category year, then there were no reconciliations made
Each column represents the new category of care.

Appendix 2: Flow diagram for inpatient care



Scaling the average cost dataset for acute inpatient costs in FY1998



Appendix 3: Contents of HERC DATASET at Austin

Discharge Dataset

```
              -----Alphabetic List of Variables and Attributes-----

 # Variable Type Len Pos Format  Informat Label
--------------------------------------------------------------------------------
 4 ADMITDAY Num    5 305 DATE9.  7.       DATE OF ADMISSION (SASDATE)
 3 ADTIME   Num    5 300                  TIME OF ADMISSION
 6 B4FY98   Num    8   0                  FLAG if Admitted Prior to Fiscal Year 98
10 COSTL_0  Num    8  32                  cost (local) for cat 0:acute med/surg
11 COSTL_1  Num    8  40                  cost (local) for cat 1:rehab
12 COSTL_2  Num    8  48                  cost (local) for cat 2:blind rehab
13 COSTL_3  Num    8  56                  cost (local) for cat 3:spinal cord
14 COSTL_5  Num    8  64                  cost (local) for cat 5:psych
15 COSTL_6  Num    8  72                  cost (local) for cat 6:substance abuse
16 COSTL_7  Num    8  80                  cost (local) for cat 7:intermed. med
17 COSTL_8  Num    8  88                  cost (local) for cat 8:domiciliary
18 COSTL_9  Num    8  96                  cost (local) for cat 9:nursing home
19 COSTL_10 Num    8 104                  cost (local) for cat 10:PRRTP
30 COSTN_0  Num    8 192                  cost (national) for cat 0:acute med/surg
31 COSTN_1  Num    8 200                  cost (national) for cat 1:rehab
32 COSTN_2  Num    8 208                  cost (national) for cat 2:blind rehab
33 COSTN_3  Num    8 216                  cost (national) for cat 3:spinal cord
34 COSTN_5  Num    8 224                  cost (national) for cat 5:psych
35 COSTN_6  Num    8 232                  cost (national) for cat 6:subst. abuse
36 COSTN_7  Num    8 240                  cost (national) for cat 7: intermed. med
37 COSTN_8  Num    8 248                  cost (national) for cat 8: domiciliary
38 COSTN_9  Num    8 256                  cost (national) for cat 9: nursing home
39 COSTN_10 Num    8 264                  cost (national) for cat 10: PRRTP
 5 DISDAY   Num    5 310 DATE9.  7.       DATE OF DISCHARGE (SASDATE)
41 FLAGEXT  Num    8 280                  FLAG if Observation Days/Cost Extrapolated
40 FLAGNH   Num    8 272                  Community Nursing Home Discharge
20 LOS_0    Num    8 112                  length of stay for cat 0:acute med/surg
21 LOS_1    Num    8 120                  length of stay for cat 1:rehab
22 LOS_2    Num    8 128                  length of stay for cat 2:blind rehab
23 LOS_3    Num    8 136                  length of stay for cat 3:spinal cord
24 LOS_5    Num    8 144                  length of stay for cat 5:psych
25 LOS_6    Num    8 152                  length of stay for cat 6:substance abuse
26 LOS_7    Num    8 160                  length of stay for cat 7:intermed. med
27 LOS_8    Num    8 168                  length of stay for cat 8:domiciliary
28 LOS_9    Num    8 176                  length of stay for cat 9:nursing home
29 LOS_10   Num    8 184                  length of stay for cat 10:PRRTP
 1 SCRSSN   Num    7 288 SSN11.  11.      SCRAMBLED SOCIAL SECURITY NUMBER
 2 STA3N    Num    5 295 STA3NL.          STATION (PARENT)
 8 costl    Num    8  16                  case-mix adj local cost
 7 costn    Num    8   8                  case-mix adj national cost
 9 flag     Num    8  24                  Cost Estimate +/- 2 Std. from Average


Sortedby:      SCRSSN ADMITDAY ADTIME DISDAY STA3N
```

Rehabilitation, Mental Health and Long Term Care Dataset

```
              -----Alphabetic List of Variables and Attributes-----

 # Variable Type Len Pos Format    Label
-------------------------------------------------------------------------------
 1 ADMITDAY Num    8   0 MMDDYY10. DATE OF ADMISSION (SASDATE)
 7 ADTIME   Num    8  48           TIME OF ADMISSION
 8 BEDSECN  Num    8  56 BEDSECN.  BED SECTION
 3 BSINDAY  Num    8  16 MMDDYY10. DAY ADMITTED TO  BEDSECT (SASDATE)
 4 BSOUTDAY Num    8  24 MMDDYY10. DAY TRANSFERED FROM BEDSECT (SASDATE)
 2 DISDAY   Num    8   8 MMDDYY10. DATE OF DISCHARGE (SASDATE)
10 DISTYPE  Num    8  72 DISTYPEL. TYPE OF DISCHARGE
 9 LSB      Num    8  64           LENGTH OF STAY IN BEDSECTION
 5 SCRSSN   Num    8  32 SSN11.    SCRAMBLED SOCIAL SECURITY NUMBER
 6 STA3N    Num    8  40 STA3NL.   STATION (PARENT)
12 cat      Num    8  88           Category of Care
14 costl    Num    8 104           Local-level Cost Estimate
15 costn    Num    8 112           National-level Cost Estimate
13 flag     Num    8  96           Cost Estimate  +/- 2 Std. from Average
11 source   Num    8  80           Categorical Indicator of Type
                                   Bedsection File Input

Sortedby:      SCRSSN ADMITDAY ADTIME BSINDAY BSOUTDAY DISDAY STA3N
```

Medical Surgical Care Dataset

```
              -----Alphabetic List of Variables and Attributes-----

 # Variable Type Len Pos Format    Label
--------------------------------------------------------------------------------
 4 ADMITDAY Num    8  24 MMDDYY10. DATE OF ADMISSION (SASDATE)
 3 ADTIME   Num    8  16           TIME OF ADMISSION
 6 BSOUTDAY Num    8  40 MMDDYY10. DAY TRANSFERED FROM BEDSECT (SASDATE)
12 COSTL    Num    8  80           Local-level Cost
13 COSTN    Num    8  88           National-level Cost
 5 DISDAY   Num    8  32 MMDDYY10. DATE OF DISCHARGE (SASDATE)
11 DRG      Num    8  72           Diagnostic Relate Groupings(DRG)
14 FLAG     Num    8  96           Cost Estimate  +/- 2 Std. from Average
 1 SCRSSN   Num    8   0 SSN11.    SCRAMBLED SOCIAL SECURITY NUMBER
 7 SOURCE   Num    8  48           Categorical Indicator of Type Bedsection File Input
 2 STA3N    Num    8   8 STA3NL.   STATION (PARENT)
 9 drgwt    Num    8  56           Diagnostic Related Groupings(DRG) Weights
10 icudays  Num    8  64           Number of days in an Intensive Care Unit
 8 los      Num    5 104           LENGTH OF STAY IN BEDSECTION

Sortedby:      SCRSSN ADMITDAY ADTIME BSOUTDAY DISDAY STA3N
```