NOTE:

These materials were prepared by subcontractors for consideration by the *Committee on Geographic Variation in Health Care Spending and Promotion of High-value Care*. These analyses were commissioned and overseen by the Committee. However, the findings and views expressed in the subcontractor reports do not necessarily reflect those of the NRC/IOM or the Committee. Neither the methodology nor the subcontractor reports have been subject to formal institutional review for the Interim Report. As the committee continues to review the findings from the analyses contained herein, we invite you to provide feedback on the content of these reports. Please note that any comments will be entered into the project's public access file, and will be available for public review.

Provide Feedback

Measuring Hospital-Level Healthcare Expenditures and Outcomes with Disease-Specific Cohorts

Jonathan Skinner
Weiping Zhou
Daniel Gottlieb
Elliott S. Fisher

The Dartmouth Institute for Health Policy and Clinical Practice, Geisel School of Medicine

Final Report for the Institute of Medicine Committee on Geographic Variation in Health Care Spending

January 10, 2013

Do not quote or circulate. Correspondence: jonathan.skinner@dartmouth.edu

Abstract
There are challenges associated with measuring regional variations in health care expenditures across health care systems in the United States. In some cases, regional measures do not provide enough guidance to individual health care systems with regard to accountability – one hospitals system may attribute their region's high costs to another system across town. As well, residents of some regions, and patients in some hospitals, may be sicker on average than in others, and risk adjustment is often problematic. Finally, when we observe differences in spending across regions or hospitals, we don't always know why – what are the components of spending on otherwise similar patients that account for higher levels of reimbursements? In this paper, we address each of these three issues. First, we use "forward-looking" one-year disease cohorts at the *hospital* level for low-variation conditions such as hip fracture, acute myocardial infarction, and stroke to measure representative risk- and price- adjusted costs and outcomes, and to compare these measures with "backward-looking" measures of end-of-life expenditures. Second, we find that many problems associated with risk adjustment methods such as the hierarchical condition categories (HCCs) are less when using these cohorts, since these patients are all quick sick, making biases less dramatic. Using a 100% sample of Medicare enrollees during 2007-09, we find remarkable variation in risk-adjusted Medicare expenditures across hospitals, variations that are highly correlated across disease categories. Much of the difference is the consequence of post-acute care, such as home health care, readmission rates, and nursing home care. Disease-specific mortality was less well correlated within hospitals, poorly correlated with spending, and (unlike costs) are more sensitive to HCC risk-adjustments.

We are grateful to the Institute of Medicine and the National Institute on Aging for financial support, and to Douglas Staiger and members of the IOM Committee on the Geographic Variation in Health Care Spending and Promotion of High-Value Care for helpful comments.

I. Introduction

There are wide variations in health care expenditures across Medicare enrollees in the United States, often measured at the regional level.[1-4] But these variations are difficult to attribute to individual healthcare systems, limiting the degree of accountability in designing reform. Can such measures also be created at the level of the hospital, given that major teaching hospitals, for example, may also receive patients with more complicated illness? And to what extent are hospitals high-spending for some conditions (e.g., hip fractures) but low-spending for others (e.g., heart attacks)?

Hospitals or regions could be more expensive because their patients are sicker. Unfortunately, risk-adjustment is not so simple when it relies on physicians or other providers to measure health status. One approach of this kind has been to adjust for illness using Hierarchical Condition Categories (HCCs), which in turn are based on Medicare billing data entered by physicians and coders. The problem with HCCs, however, is that regions treating patients more intensively are also more likely to diagnose disease. Thus finding that HCCs "explain" Medicare spending may lead to reverse causality or unmeasured confounding. For example, Song et al. found that Medicare enrollees who had recently moved to high-intensity regions experienced 19% higher HCC scores than otherwise identical enrollees who had moved to low-intensity regions.[6] And the biases were not limited to simply expenditures; because high-intensity region patients were coded with so many more diseases, their risk-adjusted mortality was 15% lower.

An alternative approach to measuring costs is to consider end-of-life cohorts. In one study by the Dartmouth Atlas group, "look-back" cohorts were constructed using Medicare enrollees with serious chronic disease (as identified by Iezzoni criterion) in the last two years (or 6 months) of life.[4,7] Thus individuals in the sample both had an identified and life-threatening chronic illness, and died. These end-of-life measures have been widely used in the literature,[8,9] but even they could – in theory – be subject to potential biases, for example if high-intensity hospitals manage to save a patient who would otherwise end up in the end-of-life cohort.[10]

There also continues to be controversy as to why such variations might arise,[5] even at the level of just knowing what sources of spending account for most of the difference in overall expenditures – do more costly hospitals bill more through DRG "upcoding" of the initial condition, or are high-spending hospital systems high-spending because of a much higher use of post-discharge utilization, much of which may be outside the control of the hospital system?

In this paper, we consider each of these questions in turn. We first return to the creation of "low-variation" look-forward cohorts first developed by Fisher, Wennberg, and others.[12-14] The idea is to identify cohorts of patients with a specific disease that results in similarly ill patients whether they live in Boise, ID or in New York City. Different in reimbursement rates are adjusted so that all variations occur because of quantity differences, and not because Medicare reimbursement rates vary because of cost-of-living in the region, teaching status, or the prevalence of low-income patients.

We also want to choose disease cohorts for which admission to hospital is generally mandatory. For example, nearly every hip fracture patient will be admitted to hospital, so that the sample of hospital admissions across regions provides a reliable picture of the actual incidence of hip fractures. By contrast, patients with pneumonia or congestive heart failure might be admitted to hospital in some regions, but not in others, leading to potential selection

biases in comparing these latter cohort's costs and outcomes across hospitals or regions.[i] In addition, we are also able to adjust for income and poverty rates at the zip code level for all forward-looking cohorts.

There are several advantages to using cohorts as markers of hospital- and region-level intensity. The first is that imperfect risk-adjustment methods leads to far more modest biases in estimating costs for these cohorts of patients, all of whom have experienced serious illness. Hospitals with lower thresholds for diagnosis may code more HCCs for otherwise identical patients, but the relative importance of any potential biases are attenuated simply because *all* of the patients in the cohort are so sick to begin with.[ii]

Second, the necessity of beginning with an index event at an acute-care hospital means that we generally identify the "accountable" institution where the acute care is provided, and where (in theory) post-discharge planning takes place. That said, much of the expenditures (and the variation) occurs after the initial index discharge, which means that we are capturing utilization and expenditures in related (or completely separate) institutions such as home health care, nursing homes, and subsequent readmissions. And finally, survival rates (either one-year or shorter duration) provides an informative (albeit noisy) measure of quality provided by the hospital system.

We consider four disease cohorts. Our primary focus will be on three forward-looking cohorts: hip fracture, heart attack (acute myocardial infarction, or AMI), and stroke. These cohorts are created using 100% samples of fee-for-service Medicare enrollees between the ages of 65-99 during the years 2007-08, following forward one year (through December 31, 2009).[iii] The final cohort is a "look-back" sample of patient with serious chronic disease dying during 2003-2007, taken from the Dartmouth Atlas. The advantage of these multiple cohort measures is that one can exclude or include subsets of the 4 disease groups depending on the reader's priors regarding the reliability of each group. As it turns out, however, they are all highly correlated, suggesting that it doesn't really matter which of these measures is used to characterize high and low spending hospitals.

Briefly, we find first that price-adjustment has an important impact on spending measures; even within the same city, some teaching hospitals are paid more per procedure or visit than other teaching hospitals. Risk adjustment for health and poverty, however, are far less important in affecting our measures of utilization and expenditures -- people with a hip fracture, heart attack, and stroke are seriously ill, no matter whether they come from Salem, Oregon or New Orleans, and no matter how large are differences in diagnostic coding. And most importantly, after all risk and price adjustments, there are large differences in utilization rates and overall one-year Medicare expenditures across U.S. hospitals. For example, average one-year expenditures for a hip fracture patient at Gunderson Hospital in La Crosse, WI is $36,477, adjusted for the lower prices in Wisconsin, differences in poverty rates, age, sex, race, and HCC

[i] Even more of a problem arises when patients in some regions are more likely to be admitted multiple times for the same problem (e.g., congestive heart failure). Then almost by definition, case fatality rates will be lower for these patients.

[ii] The magnitude of biases is likely to be larger within the general population, where just a few "gray area" diagnoses causes an otherwise functional patient to be deemed chronically ill.

[iii] We also considered a colon cancer cohort, but the sample sizes were so small that we did not use them for hospital-level analysis.

illness measures in the six months prior to the index admission. By contrast, similarly adjusted spending for a hip fracture patient at Edinburg Regional Medical Center, near McAllen Texas, is $79,349.

Furthermore, these risk-adjusted expenditure measures are highly correlated across disease groups – hospitals that are costly for AMI patients are also costly for stroke and hip fracture patients. We use factor analysis to estimate the common cost "factor" that best captures the propensity of a given hospital to experience high or low expenditures for a representative cohort patient. We report this "C-factor" for a selected group of hospitals, which has zero mean and a standard deviation of one.[iv] (In this case, a "6-sigma" hospital is a very expensive hospital indeed across a broad swath of patients.) All cohort-level expenditures were highly correlated with this single common "C-Factor," ranging from 0.68 (stroke) to 0.88 (end-of-life expenditures).

Why are there such large differences in expenditures? To address this question, we focus on hip-fracture patients for four different (and quite select) groups of hospitals: high versus low cost, and medical centers versus non-medical centers. We find that high-cost hospitals are higher cost because of post-acute services: more readmissions, more nursing home care, and more home health care. Relatively little is the consequence of higher DRG "upcoding" at admission.

Finally, we briefly consider an equivalent hospital level "M-Factor" or the common factor for one-year mortality by hospital across the different forward-looking cohorts.[v] And unlike the cost estimates, the survival estimates are sensitive to HCC-risk adjustment. While we argue that a simple correlation coefficient between expenditures and mortality says little about the value or efficiency of health care expenditures, we do find that the association between spending (the C factor) and outcomes (the M factor) is inconclusive and highly dependent on whether one uses HCC risk adjustment in the mortality measure.

II. Statistical Model

The simplest approach to measuring costs is to take average Medicare expenditures over the one year post-admission for a given cohort – no adjustment whatsoever. (The description of how we calculated Medicare expenditures is below in Section III.) In that case, the regression model would be

(1) $$C^*_{ij} = \mu + \sum_{j=1}^{J} \delta_{1j} D_j + \varepsilon_{1ij}$$

This equation includes a constant term μ, so that the hospital-level categorical variable δ_{1j} is the divergence of average Medicare expenditures for the jth hospital. (In practice, when we report expenditures by hospital, we will add back in the constant term μ.) The hospital indicator, D_j, is one if the patient is in hospital j, zero otherwise. Note also that C_{ij}* , expenditures for individual i in hospital j, has not been adjusted for differences across hospitals in the prices paid by Medicare for services, so we might expect such spending to be systematically higher in urban areas or for teaching hospitals.

The next estimate of hospital-level expenditures only swaps C_{ij}* with price-adjusted expenditures C_{ij}:

(2) $$C_{ij} = \mu + \sum_{j=1}^{J} \delta_{2j} D_j + \varepsilon_{2ij}$$

[iv] Remember, the "C-factor" measures how much the hospital costs Medicare for a given patient, or what expenditures would have been for the hospital had Medicare reimbursed all hospitals at the same prices.

[v] For obvious reasons, we do not calculate mortality rates for our end-of-life cohort.

Thus the ratio δ_{1j} /δ_{2j} is the price index for that specific hospital, which reflects cost-of-living, payments for graduate education, and payments for the Disproportionate Share Hospitals (DSH) program. As we show below, price-adjustments can vary across hospitals, even in the same urban market.

It is straightforward to adjust for more characteristics of the patient. Let X_{ij} comprise the basic set of risk-adjusters: age/sex/race cells (that is, five-year intervals of age interacted fully with sex and race), comorbidities coded at admission, type of diagnosis (for example, the location of the heart attack or the location of the break on the femur), and the individual's zip code median family income (by race) and poverty rates. Then once again we estimate this equation

(3) $$C_{ij} = \mu + \sum_{j=1}^{J} \delta_{3j} D_j + X_{ij}\beta + \varepsilon_{3ij}$$

which yields the partially risk-adjusted and price-adjusted measure of hospital spending, δ_{3j} for hospital j. Finally, the fully specified estimating equation adds the semi-parametric HCC score interval for each patient based on the 6 months prior to admission;

(4) $$C_{ij} = \mu + \sum_{j=1}^{J} \delta_{4j} D_j + X_{ij}\beta^* + \sum_{k=1}^{K} \alpha_k H_k + \varepsilon_{4ij}$$

Note that the HCC score is split into discrete intervals so as to allow for a fully flexible functional form. We estimate the same basic model for each of the forward-looking cohorts, even if there are minor differences in the sets of comorbidities for each disease that depend on the disease itself (e.g., location of the AMI or location of the break on the femur).

The end-of-life expenditure measures from the Atlas are estimated differently. These are a retrospective cohort (for obvious reasons) attaching patients to hospitals based on where they received the plurality of their care. In order to be included in the sample, they needed to have at least one serious chronic disease, and when there was more than one chronic disease, the individual was assigned to the highest-cost disease, and assigned an additional categorical variable denoting multiple chronic diseases. Expenditures in the two years prior to death were then stratified and adjusted for differences across regions in age, sex, race, and disease. [4] Updated data were used based on a more recent Dartmouth Atlas report.[vi] These end-of-life expenditure measures were not adjusted for differences in prices, so we calculated the price adjustment for each hospital (C*/C) using an average of our largest three cohorts: hip fracture, stroke, and AMI, and used that price-adjustment factor to create a price-adjusted end-of-life spending measure. Note finally that these end-of-life measures do not adjust for poverty or income.

Risk-adjusted mortality measures by hospital were estimated in the same way. For obvious reasons, we were not concerned with price-adjustment issues for mortality, but we did estimate three basic models: first without any adjustment (so the average one-year mortality), second with age, sex, race, type of disease (e.g., location of AMI), comorbidities at admission, zip-level poverty rates and race-specific income, and third the previous set of comorbidities plus the individual HCC index. Thus the full estimation equation is

(5) $$M_{ij} = \mu + \sum_{j=1}^{J} \kappa_{4j} D_j + X_{ij}\Gamma^* + \sum_{k=1}^{K} \upsilon_k H_k + \xi_{4ij}$$

with a less fully-specified model excluding the HCC dummy variables H_k Note that while mortality is a one-zero variable, we still estimate it using a linear regression, since the parameters

[vi] http://tdi.dartmouth.edu/documents/EOL_Trend_Report_0411%20%283%29.pdf

of interest are average adjusted mortality by hospital, the coefficients κ_j.[vii] Non-linear transformations of M, such as logit or probit models, would complicate the use of factor analysis (below) by individual hospital.[viii]

We estimate the factor model using the conventional approach. First, we estimate specific hospital-level expenditure measures δ_j for a variety of risk-adjustment methods. (In practice, we consider only risk-adjusted measures either with or without HCCs.) Then we can write

(6) $$\hat{\delta}_j = \sum_{m=1}^{M} \lambda_m L_m + \varsigma_j$$

where the m = 1,…,M vectors L_m are orthogonal and are independent of the error term ζ. The λ_m terms comprise the loading matrix; note that these are not unique given that the factor matrices and λs can be rotated. In general, this degree of invariance matters more when there are multiple factors that explain C, but as we shall see, Medicare expenditures tend to load up on one common factor. Thus just as a variety of test scores may reflect a single "math ability," so by the same token the different cohort-specific levels of Medicare expenditures reflect a common hospital-specific intensity or cost factor, which we call the "C-factor." A similar analysis is performed for mortality, resulting in a corresponding "M-factor".

III. Data

In this section, we first consider the Medicare data taken as a whole, consider the common inclusion criteria for each of the four forward-looking cohorts, examine each cohort separately, and then briefly describe the ancillary Census data used to measure income and poverty rates.[ix]

III.A Medicare Files

We first begin with the sources of files to create each cohort, which are described in more detail in Table 1.

Table 1: Description of Medicare Claims Data Using to Create Cohorts: 100% Sample, 2007-09

Enrollment Files:	*Description*
Denominator File	Annual file containing record for each Medicare beneficiary, includes Bene_ID (unique patient identifier), demographic information (age, sex, race), residence (ZIP code), program eligibility and date of death.
Claims Files:	
MedPAR File (Hospital discharge database)	Annual file containing the final claim for each hospital stay during year. Includes Bene_ID, hospital ID, dates of admission/discharge, diagnoses, procedures, DRG, charges and payments.

[vii] The corresponding mortality model in the end-of-life cohort yielded an R^2 of one – that is, everyone died. This was discovered independently by Betsy McGaughey, a former Lieutenant Governor of New York ("Treating Seniors as "Clunkers," The New York Post, October 27, 2009).

[viii] Logit or probit would provide an advantage by keeping predicted or adjusted mortality rates from being less than zero, but given the high mortality rates in these cohorts, this is not a problem.

[ix] As noted above, the end of life cohort methods are described here: http://tdi.dartmouth.edu/documents/EOL_Trend_Report_0411%20%283%29.pdf

Inpatient File (Hospital claim database)	Each annual file contains interim and final claims. In addition to MedPAR fields includes UPIN of attending and primary procedure MD. Used to link physicians to specific hospitals and procedures.
Physician Supplier Part B (Physician claims)	Annual file containing claims for physician services for 100% sample. Records include Bene_ID, performing and referring physician identifier (UPIN or NPI), dates, procedures, diagnoses, and payments.
Outpatient File (Facility bills for outpatient services)	Annual file containing claims for outpatient services at hospitals and other facilities. Includes Bene_ID, hospital ID, dates, attending UPIN/NPI, procedures, and payments.
Home Health Agency File	Annual file containing claims for home health agency services, including Bene_ID, provider ID, dates of service, types of service provided, primary diagnosis, charges and payments.
Hospice File	Annual file containing claims for hospice services, with Bene_ID, provider ID, period of service and diagnosis leading to hospice enrollment.
Skilled Nursing Facility (SNF) and Long Stay Files	Annual files containing claims for SNFs and long-stay facilities (rehabilitation). Data included are similar to that of MedPAR file.

III.B Inclusion Criteria

We next consider the common inclusion criterion for the four forward-looking cohorts in Table 2. As noted above, we include all individuals admitted during 2007-08 for one of the four index health events. These individuals are restricted to those who were in fee-for-service Medicare both during the one-year follow-up, as well as the 6-month "sweep out" period prior to the index event. The individual also needed to be at least 65 years and 6 months at the time of the index event (to allow us to collect risk-adjustment data during the 6 months prior to the index event.)

Table 2: Basic inclusion criterion for all "Forward-Looking" Cohorts

1. Admission dates occurring between and including January 1, 2007- December 31, 2008, in the 100% samples in the fee-for-service sample. Enrollees are followed through to December 31 2009 where necessary.
2. Continuous one year non-HMO from index event admission date and 6 months prior to the index event. Part A and B eligible during entire 1.5 year span of data or until date of death after the index admission. This means that the data analyzed may stretch back to July 1, 2006 (for those admitted on January 1, 2007).
3. Age 65 six months prior to index admission to age 99 on the day of admission.
4. U.S. residence
5. No previous index event in the preceding 6 months.
6. Admission must be to an acute-care hospital, although transfers are allowed to other types of hospitals or facilities

Finally, we determine characteristics of each of the cohorts, as shown in Table 3 below:

Table 3: Criterion for Disease-Specific Cohorts

1. Hip Fracture Cohort: All patients with a fracture of the neck of the femur, primary DX codes 82000-82099
2. Acute Myocardial Infarction Cohort: All patients with primary DX codes 41000-41091 (excluding diagnosis with fifth digit of "2" because those are subsequent diagnoses).
3. Stroke Cohort: For this analysis, we used only 433 (Occlusion and stenosis of precerebral arteries) and 434 (Occlusion of cerebral arteries) based on Bravada (2003) noted in aspe.hhs.gov/health/reports/09/mcperform/report.pdf, but with the fifth digit of the primary DX code of "1." We also examined primary DX codes 431,436 but did not use them in this analysis.
4. End-of-life Cohort: Died between 2003-07 with the presence of a modified Iezzoni chronic condition; see http://tdi.dartmouth.edu/documents/EOL_Trend_Report_0411%20%283%29.pdf

Finally, we considered in our statistical analysis only hospitals with at least 50 patients for the specific cohort. Thus in some hospitals, sample size is large enough for hip fracture patients, but not for stroke patients. We next turn to the specifics of the risk adjustments.

III.C Risk Adjustment

Age, sex, race. For the purpose of the risk adjustment, 5-year age brackets, interacted with sex and race (white and non-white). Because there are relatively few Hispanic and Asian enrollees noted as such in the race/ethnicity classification (fewer than 2000), additional categorical variations are included for race or ethnicity: Hispanic, Asian, Native American, and Other, but are not interacted with age and sex.

Type of disease. For example, consider the hip fracture cohort. From the DX code, by four-digit classification, we know the fracture of: unspecified intracapsular section of neck of femur, closed (8200), epiphysis (separation, upper) of neck of femur closed (8201), midcervical section of femur closed (8202), base of neck of femur closed (8203), other transcervical fracture of femur closed (8209).

In the AMI cohort, the classification of risk factors relate to physical location of the AMI: Anterolateral wall (4100), other anterior wall (4101), inferolateral wall (4102), inferoposterior wall (4103), other inferior wall (4104), other lateral wall (4105), posterior wall (4106), subendocardial (4107), other specified site (4108), unspecified site (4109). Specifications for stroke were more uniform, with primary DX code 433 (Occlusion and stenosis of precerebral arteries) and 434 (Occlusion of cerebral arteries).

Comorbidities at admission: These include vascular disease, pulmonary disease, dementia, diabetes, liver disease, renal disease, cancer (local) cancer (regional) and cancer (metastatic), previous myocardial infarction. As noted earlier, comorbidities are not included when they describe the cohort; e.g. previous myocardial infarction for the AMI cohort.

The 6-month-prior Hierarchical Condition Categories (HCC) index. This is our "best case" index that excludes potential rule-out diagnoses by requiring at least two diagnostic tests or follow-up visit to code for the specific diagnosis. This has been further broken down into k = 1,…,9 intervals (recoded to the upper interval point of .5, .75, 1.0, 1.25, 1.5, 2,3, 5, 10).

Zip code income (race specific) and poverty rates: These are based on the 2000 Census data, using deciles of income, and actual poverty rates. To avoid ecological biases, race-specific income measures are used.

IV. Results

In this section, we focus primarily on the hip fracture cohort to avoid an overload of data tables, but then consider results for all of the cohorts and the factor model.

IV.A The Hip Fracture Cohort: Risk Adjustment. We first estimated a series of regression models as described above, with the most complete risk-adjustment models reported in Appendix Table A.1 for price-adjusted expenditures (following the methods in Gottlieb et al.[16]) . Missing from the regression results are the age-sex-race cell measures, which reflect (e.g.) African-American men age 70-74. Note that non-African-American minority enrollees are adjusted for using a single categorical variable (e.g., Native American) but are included in the non-African-American age-sex cells. The sample size is 332,649.

First note that the influence of income and poverty, conditional on health status, is quite modest. The income decile variables (inc2 through the highest income category inc10, with the

lowest income group excluded) do not suggest any strong income-based pattern; none are significant at levels commensurate with a sample size of over three-hundred thousand, and quantitatively they are not large. For a mean value of nearly $50,000 in spending, these measures imply that moving 100% of patients from the lowest to the highest income group would reduce spending by about $300 per patient. Similarly, if the fraction of people living in poverty in patients' neighborhood zip codes rose from 10% to 30%, predicted expenditures would rise by roughly $600, a result consistent with previous work.[17] In any case, these income and poverty effects are adjusted for in subsequent estimates of hospital expenditures.

The HCC measures do matter for health care expenditures. Figure 1 shows the coefficients with the blue line (the left vertical axis); all standard errors are less than about $500, so these estimates are highly significant. Note that the coefficient estimates adjust for each hospital with a fixed effect, so that these estimates are based on variation in HCC and spending *within* a hospital. There is a very strong association between the HCC measure and expenditures, with estimated effects ranging as high as $31,807 for the most serious HCC code (an HCC of 5, where 1 is the national average) compared to the least serious (< 0.5). This illustrates that at the individual level, HCCs are highly predictive of expenditures.

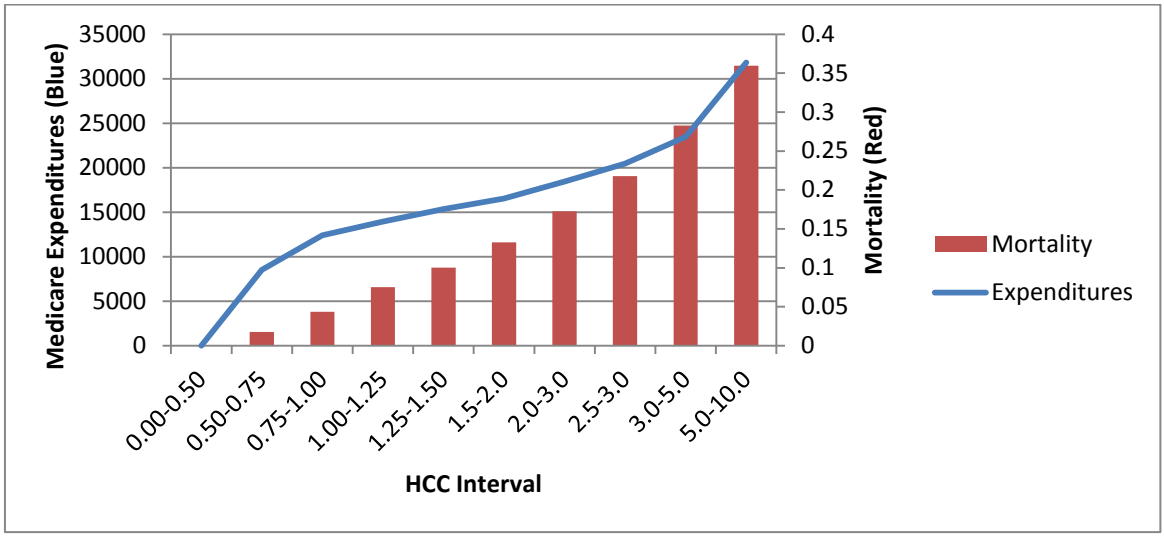

Figure 1: Estimated Association between HCC Measures and Expenditures/Mortality

Bias issues occurs typically when there are systematic differences across hospitals with regard to their aggressiveness in coding HCCs, thus leading to higher apparent average HCC levels at high-intensity hospitals. While that does appear to be true – Cedar-Sinai hospital in Los Angeles (a high-expenditure hospital) experienced 2 percent lower expenditures once adjusted for HCCs, and Gunderson Hospital in La Crosse WI experienced 4 percent greater expenditures after adjustment, the *magnitude* of these adjustments is quite modest given the enormous underlying differences in utilization between the two hospitals.

Table 4 shows average one-year expenditures for hip fracture patients among a set of hospitals, many of which have figured in the national health care reform debate, serve low-income areas, such as the University of Chicago hospital, or are located in traditionally high-expenditure or low-expenditure regions. Teaching hospitals and urban hospitals are overrepresented in our sample, so that average expenditures tend to be above the national

average of $49,171.[x] We present measures of expenditures following the methods section; first unadjusted (raw) average Medicare expenditures, then price-adjusted, then with risk-adjustment (less HCCs), and finally risk-adjustment with HCCs. Note that the hospitals are sorted according to the HCC-risk-adjustment method in the final column.

Expenditures range from about $36 thousand to $83 thousand when not adjusted for price or health or even age or sex. The gap between the highest and lowest hospital expenditures shrinks only slightly with complete price and risk adjustment, as is shown in the final column where the expenditures range between $36 thousand (again) and $79 thousand. The largest degree of adjustment occurs when introducing price-adjustment. Hospitals with large numbers of residents, those treating a large fraction of low-income patients, and in high-cost cities experience the greatest drop in adjusted expenditures.[xi] Subsequent risk adjustment appears to have only modest effects on spending measures – the ranking is essentially unchanged as greater risk-adjustments are applied. While we did not present standard-deviations explicitly, the 95% confidence intervals are roughly ± $5,000, with narrower bounds for the larger hospitals.

We also estimated a risk-adjusted model of one-year mortality, with results for the risk adjustment model that includes HCCs shown in Appendix A.2. (Once again, the age-sex-race categorical variables are not reported.) Income conditional on the HCC score has a modest (1.5 percentage point) and marginally insignificant impact on mortality only for the top income decile; the coefficient on poverty was not significant. Finding for the null hypothesis does not mean that income is irrelevant for health or that poverty does not kill – only that once a patient presents with a hip fracture, for example, income does not have a primary impact on health outcomes. This holds even more strongly when HCCs are not included in the risk adjustment model; zip-code income no longer is significant for any income decile.

It is the case, however, that HCCs matter in predicting mortality; Figure 1 shows as well the estimated coefficients for the intervals for HCCs, conditional on other factors such as age and comorbidities.[xii] The difference in predicted mortality between the lowest and the highest HCC scores is more than 35 percentage points. And HCCs matter more as well in estimating risk-adjusted mortality. Higher-expenditure hospitals – even after already having been adjusted using HCCs – show systematically "sicker" patients, given their considerably higher HCC scores. This in turn means that adding HCC risk adjustment to the risk adjustment model can have a much larger (and negative) impact on adjusted mortality risk. We return to this point below.

IV.B AMI, Stroke, and End-of-Life Cohorts

Table 5 includes expenditure estimates from the AMI, stroke, and end-of-life cohorts. (Recall that we have price-adjusted the end-of-life expenditure measures using created price indices from the hip fracture, AMI, and stroke cohorts.) The sample sizes for the new cohorts (for hospitals with at least 50 observations) are 365,168 for the AMI cohort, and 298,080 for the stroke cohort. Ocular inspection suggests a strong association among the different measures of expenditures. We make this association more rigorous next using factor analysis.

IV.C. Factor Analysis

[x] At least among hospitals with at least 50 hip fracture Medicare patients during the two year period 2007-08.

[xi] There is a debate as to whether graduate medical education supplements should be excluded from expenditures. While these payments pay for resident physicians, who visit patients but who are not allowed to bill under Medicare rules, there may also be additional costs of educating residents.

[xii] Recall that HCCs already include age, but we include age in a more nonparametric way as well.

Results from the factor analysis for both expenditures and mortality are presented in Appendix Tables A.3a and A.3b. There appeared to be one primary factor for both expenditures and mortality; the second factors were essentially unimportant. An intuitive sense of the importance of the common factor is gained from the correlation between each cohort measure and the common single factor; these are shown in Figure 2 below.

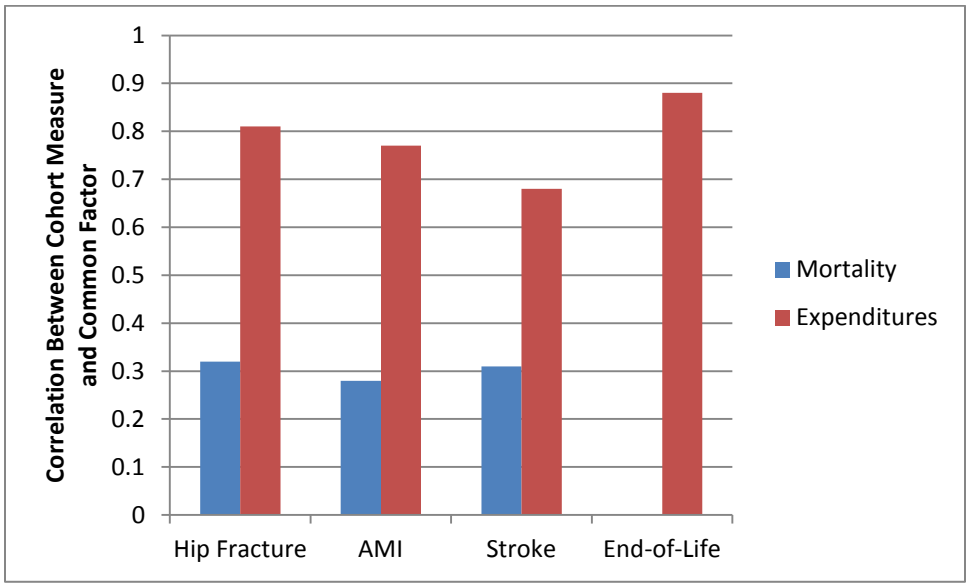

Figure 2: Correlation between the Common Factor for Expenditures (Red) and Mortality (Blue) for each cohort. (Colon Cancer cohort excluded because of small sample size)

The first factor for expenditures is strongly associated with all of the cohort expenditure measures, with the weighted correlation ranging from 0.68 for stroke to 0.88 for end-of-life expenditures. That is, end-of-life expenditure, despite its earlier time-period (2003-07) appears to be the single best predictor of the common expenditure "factor" that holds across other forward-looking cohorts.

The C-Factor is shown for each of the hospitals of interest in Table 5. Note that the measures have mean zero and standard-deviation one, so that a measure of -1.0 means that the hospitals ranks one standard deviation below average in a normal distribution. Thus a "four-sigma" hospital, like the University of Miami Hospital (4.17) is a very expensive one indeed. There is also evidence of considerable skewness; one finds positive outliers of 3 or more, but no negative outliers of similar magnitude.

IV.D Why are High-Spending Hospitals High-Spending? A Decomposition Exercise

We have documented systematic differences across hospitals with regard to expenditures, but one might ask why they differ? Without hazarding a deeper response about incentives or physician beliefs, we can at least perform a decomposition between risk-adjusted expenditures for four different types of hospitals: high spending and low spending (for medical centers), and high spending and low spending (for non-medical centers). Rather than a systematic decomposition, we instead choose three (or at most four) representative hospitals in each

category, and average their spending measures so as to dampen individual hospital-level idiosyncrasies and improve statistical power. The specific categories are low-spending medical centers (Mayo Clinic – St. Mary's, Stanford Hospital, UCSF), high-spending medical centers (UCLA, NYU, and Cedar-Sinai[xiii]), low-spending non-medical centers (Gunderson Lutheran, Intermountain – LDS Hospital, and St. Mary's in Grand Junction CO), and high-spending non-medical centers (Doctor's Hospital at Renaissance – McAllen TX, Edinburg Regional TX, Miami Metropolitan, and Garfield Hospital in Los Angeles).

Table 6 displays the decomposition exercise for the price-adjusted data. For the medical centers, the overall difference between low-spending ($47,235) and high-spending ($62,167) hospitals is $14,932, while for non-medical centers, the difference between low-spending ($42,255) and high-spending ($81,805) is substantially larger, $39,550. Note that these measures are not risk-adjusted, because were we to risk-adjust each component of spending, they would not add up to the total.[xiv] However, the overall risk-adjusted gaps (conditional on HCCs) for the medical centers and non-medical centers, which are $13,632 and $35,787, respectively, are quite similar to the unadjusted gaps, meaning that health status, income, and poverty are not the reasons why these hospitals are so different from one another.

One hypothesis for why high-spending hospitals are so much more expensive is because they could be better at coding more expensive DRGs or billing for outlier payments. Table 6 shows Part A expenditures by group, and suggests that if anything, higher spending medical centers bill slightly less ($83) than low spending medical centers. That said, the higher spending non-medical-center hospitals show $2,906 more in spending, or about 7% of the total gap between high and low spending hospitals. This is reflected in Table 6 that shows broken out the total number of inpatient index physician visits; these are twice as high in high-spending medical centers (21) compared to low-spending medical centers (10), but the differences are even more pronounced in non-medical-centers (21 versus 7). High-spending centers do invest more in Part B expenditures during the index admission (typically physician visits), with 50% higher billing in both medical and non-medical centers, accounting for 7% and 3% of the overall gap, respectively.

The largest explanation for why high-spending hospitals are so much more expensive is post-acute (that is, after discharge for the initial admission) Part A hospital and nursing home expenditures. (Recall that post-acute ranges from the day after discharge to one-year post-admission for the index event.) As shown in Table 6, medical center spending in the high-spending hospitals is $5,389 more, or 36% of the overall gap, while for non-medical-centers, the difference is $20,881 – that is, patients in high-spending hospitals receive three times as much as those in low-spending hospitals ($31,590 versus $10,709), or 53% of the overall difference. The bottom row in Table 6 also shows the likelihood of being readmitted to the hospital within 30 days; 70% of non-medical-center patients are readmitted in the high-spending hospitals, compared to just 27% in the low-spending hospitals. These differences are mirrored as well in post-acute Part B (physician, tests) spending, with $5,285 additional spending for medical centers, and $7,235 additional spending for non-medical-centers.

Finally, Table 6 considers additional post-acute ("PA") expenditures, and finds that for both medical centers and non-medical-centers, differences in home health care spending

[xiii] Cedar-Sinai is not, strictly speaking, an integrated medical center, however.

[xiv] That is, suppose that we wanted to decompose total spending into Part A and Part B. If the risk-adjusted total difference between two hospitals is, say, $15,000, there is nothing to ensure that risk-adjusted Part A plus risk-adjusted Part B would add up to $15,000 for subsets of hospitals.

accounts for roughly one-fifth of the overall gap. The differences are far more pronounced in the non-medical-centers, with high-spending hospitals accounting for $9,083 per hip fracture patient, compared to just $2,162 for the low-spending hospitals. Finally, hospice care was higher in the lowest-cost hospitals.

In sum, most of the differences in spending occurred *after* the patient was discharged from the initial index. This makes sense; the DRG system puts some natural limits on how much can be billed to Medicare during the initial admission, although such limits are not present for physician visits. And while those initial physician visits may have laid the groundwork for subsequent utilization, most of the spending gap – whether for medical centers or non-medical-centers – took place during the time post-admission, for example in the use of readmissions, rehabilitation and nursing homes, home health care, and other ancillary services that are billed on a fee-for-service basis.

IV.E Mortality Estimates

While mortality estimates are not the primary focus of this report, we do briefly consider factor analysis of different mortality measures for a given hospital, paralleling our earlier analysis of Medicare spending patterns. (Again we drop the end-of-life cohort for obvious reasons.) The correlation across different cohorts is more modest, with correlations that range between 0.28 and 0.31 (Figure 2). Whether this common factor represents true differences in clinical quality across different hospital is not clear. More worrisome is if the factor analysis distills down any unmeasured hospital-specific confounding factor, so our "mortality factor" could be highlighting unmeasured health factors like a regional fondness for fatty foods.

While we have argued elsewhere that simple correlations between expenditures and outcomes tell us little about health systems efficiency,[18,19] it may be of interest to compare the expenditure factor (or "C-Factor") and the mortality factor (or "M-Factor), as shown in Figures 3 and 4. The two graphs may look at first to be similar – each looks very scattered. Both use expenditures that have been corrected for differences across hospitals in HCC measures. In Figure 3, we use risk adjustment without HCCs, in this case the correlation is -0.02 ($p = 0.51$) – that is, uncorrelated. When using risk adjustment with HCCs, as in Figure 4, however, the correlation is negative (-0.19) and significant ($p < .01$).

Why the difference? As noted earlier, HCC adjustment for mortality measures matters, with high-expenditure (and, most likely high-diagnosis) hospitals like Cedar Sinai experiencing a drop of as much as one-half a standard deviation in estimated mortality with the introduction of HCC measures. This finding also could explain why Ong et al. only find a negative association between expenditures and outcomes when they risk-adjust using similar claims-based risk-adjustment.[11]

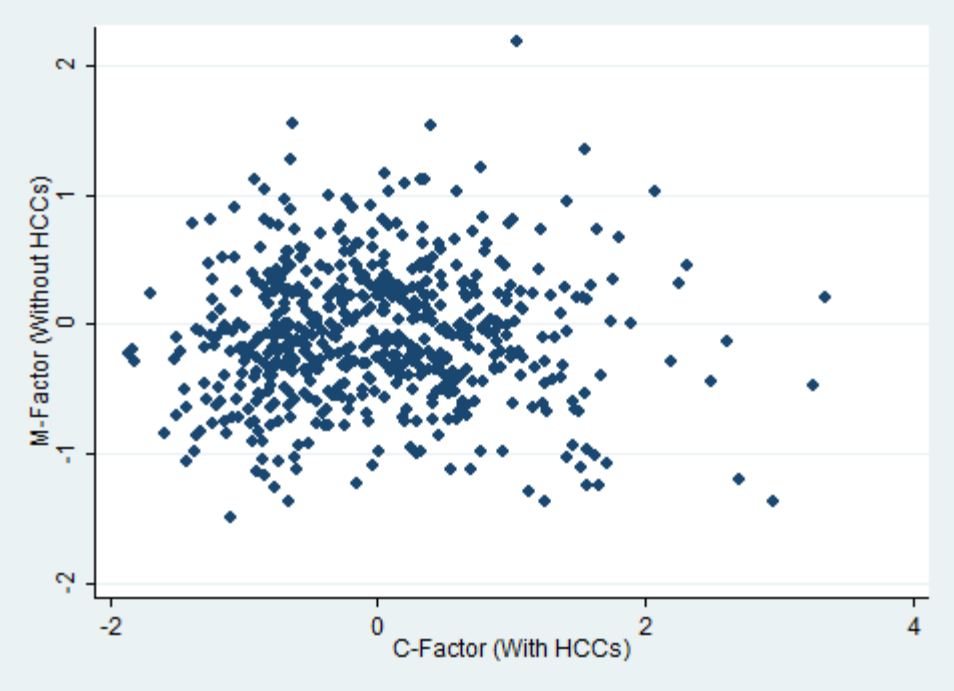

Figure 3: Hospital-Specific One-Year Expenditures and Mortality (without HCCs)

Figure 4: Hospital-Specific One-Year Expenditures and Mortality (with HCCs)

One may view the two alternative mortality measures as "bounds" with the true measure somewhere between the two.[xv] More to the point, both Figure 3 and Figure 4 – which represent smoothed measures of both expenditures and mortality by hospital – display a remarkable lack of association between health care inputs and outputs for these cohorts of patients. (Put another way, even a correlation coefficient of 0.2 means that expenditure differences across hospitals can explain at most 4% of the variance in mortality outcomes.) An interesting topic for future research is to attempt to measure and quantify the factors which appear to have the greatest impact on mortality and other outcomes.

V. Discussion

In this paper, we have argued that the use of low-variation cohorts with well-defined index events (and hence start times) are a viable approach to measure the intensity of care; they reflect both intensity in the initial index event, as well as post-discharge patterns of care in the community. Because these cohorts represent individuals who are objectively sick, they avoid the common problems inherent in using claims-based risk adjustment, [6,21] at least with regard to measuring utilization.

Indeed, we found that when using these cohorts, much of the unadjusted expenditure differences across hospitals were largely unaffected by age, sex, race, and illness adjustment. The rankings did change considerably with price-adjustment; Medicare pays far different amounts of money to different hospitals depending on their graduate medical education, cost-of-living, and fraction of patients from low-income regions. Thus these cohort-based measures of Medicare expenditures appear to sidestep the otherwise intractable issues surrounding biases introduced by HCC-level adjustments.

We also introduced a factor analysis approach allows for the characterization of hospital-specific intensity that is common across cohorts. We show that there is a very strong association across all measures, including end-of-life care from an earlier period (2003-07). This suggests some degree of temporal stability for measures of institutional health care intensity measures.

We also show that these results are consistent with earlier work showing differences in expenditures tend to reflect specific components of what are likely to be more discretionary components of care: the frequency of readmission initially and over time, differences in the intensity of post-acute care, and the use of evaluation and management services and associated tests and diagnostic procedures.[13,14,22] And while we show that the general patterns are similar for medical centers and non-medical-centers, the sheer magnitude of the differences in Medicare spending between high- and low-spending hospitals are substantially greater for the hospitals that are not integrated medical centers.

We also find variations in risk-adjusted mortality. These differences may arise from different clinical processes, or even from systematic confounding across hospitals, but at best are very poorly explained by expenditures. As well, hospitals may be high performing on clinical quality and safety for one condition, but not for others. Simple correlations between spending

[xv] As noted above, even if we knew the "true" correlation, this would not tell us whether more is better or worse. For example, suppose that greater expenditures (particularly post-discharge) did nothing for health outcomes, but that hospitals with more intensive treatment patterns also were the ones using surgical quality reforms such as checklists; see for example deVries, EN, et al., (NEJM 2010: 363:1928-37). Mortality would be lower because of the surgical checklists, not because patients are more likely to be seen by multiple physicians.

and outcomes may not tell us much about the effectiveness of spending more, but they do provide an insight into the potential biases arising from risk-adjusting mortality rates using HCC adjustments.

In sum, we would argue that the current technology and statistical approaches are at the point where we can accurately characterize Medicare expenditures and utilization intensity for individual hospitals and the post-discharge services provided to their patients. At this stage, we may not be able to yet accurately characterize risk-adjusted mortality, given that the most popular risk-adjustment measure exhibited biases which in this paper has been shown to be large in magnitude. More accurate measures of health, perhaps with the help of biomarkers and self-reported data, represent a future research task to better measure outcomes.

References

1. Goodman DG, Fisher ES, Chang C-H, Morden NE, Jacobson JE, al. e. Quality of End-of-Life Cancer Care for Medicare Beneficiaries. Lebanon, NH; 2010
2. Service NH. The NHS Atlas of Variation in Healthcare. In. London; 2010.
3. Wennberg JE, Cooper MM. The Dartmouth Atlas of Health Care in the United States. Chicago: American Hospital Publishing, Inc.; 1999.
4. Wennberg JE, Fisher ES, Goodman DC, Skinner J. Tracking the Care of Patients with Severe Chronic Illness: The 2008 Dartmouth Atlas of Healthcare. Lebanon NH: Dartmouth Institute for Health Policy and Clinical Practice; 2008.
5. Bach PB. A map to bad policy--hospital efficiency measures in the Dartmouth Atlas. The New England journal of medicine 2010;362:569-73; discussion p 74.
6. Song Y, Skinner J, Bynum J, Sutherland J, Wennberg JE, Fisher ES. Regional variations in diagnostic practices. The New England journal of medicine 2010;363:45-53.
7. Wennberg JE, Fisher ES, Stukel TA, Skinner JS, Sharp SM, Bronner KK. Use of hospitals, physician visits, and hospice care during last six months of life among cohorts loyal to highly respected hospitals in the United States. BMJ 2004;328:607.
8. Silber JH, Kaestner R, Even-Shoshan O, Wang Y, Bressler LJ. Aggressive treatment style and surgical outcomes. Health services research 2010;45:1872-92.
9. Romley JA, Jena AB, Goldman DP. Hospital spending and inpatient mortality: evidence from California: an observational study. Annals of internal medicine 2011;154:160-7.
10. Bach PB, Schrag D, Begg CB. Resurrecting treatment histories of dead patients: a study design that should be laid to rest. JAMA : the journal of the American Medical Association 2004;292:2765-70.
11. Ong MK, Mangione CM, Romano PS, et al. Looking forward, looking back: assessing variations in hospital resource use and outcomes for elderly patients with heart failure. Circ Cardiovasc Qual Outcomes 2009;2:548-57.
12. Fisher ES, Welch HG. Hospital readmission rates for cohorts of Medicare beneficiaries in Boston and New Haven. New England Journal of Medicine 1994;331:989-95.
13. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care. Annals of internal medicine 2003;138:273-87.
14. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The implications of regional variations in Medicare spending. Part 2: health outcomes and satisfaction with care. Annals of internal medicine 2003;138:288-98.
15. McCaughey B. Treating seniors as 'clunkers'
The New York Post 2009 October 27.
16. Gottlieb DJ, Zhou W, Song Y, Andrews KG, Skinner JS, Sutherland JM. Prices don't drive regional Medicare spending variations. Health Aff (Millwood) 2010;29:537-43.
17. McClellan M, Skinner J. The Incidence of Medicare. Journal of Public Economics 2006;90:257-76.
18. Fisher E, Skinner J. Comment on Silber et al.: Aggressive treatment styles and surgical outcomes. Health services research 2010;45:1893-902; discussion 908-11.
19. Skinner J, Staiger D. Technology Diffusion and Productivity Growth in Health Care. Cambridge MA: National Bureau of Economic Research; 2009.
20. de Vries EN, Prins HA, Crolla RM, et al. Effect of a comprehensive surgical safety system on patient outcomes. The New England journal of medicine 2010;363:1928-37.
21. Welch HG, Sharp SM, Gottlieb DJ, Skinner JS, Wennberg JE. Geographic variation in diagnosis frequency and risk of death among Medicare beneficiaries. JAMA : the journal of the American Medical Association 2011;305:1113-8.
22. Fisher ES, Wennberg JE, Stukel TA, Sharp SM. Hospital readmission rates for cohorts of Medicare beneficiaries in Boston and New Haven. The New England journal of medicine 1994;331:989-95.

Table 4: One-Year Hip-Fracture Medicare Expenditures, 2007-09: Selected Hospitals

Hospital	HRR City	State	N of Obs.	Unadjusted (1)	Price-Adjusted (2)	Price-Ratio	Risk-Adjusted except HCC (3)	Risk-Adjusted with HCC (4)
Gundersen Lutheran Med Center	La Crosse	WI	186	$35,036	$37,046	0.95	$35,826	$37,239
St. Mary's Hospital	Rochester	MN	375	$41,585	$42,026	0.99	$42,305	$43,390
St. Mary's Hosp & Med Center	Grand Junction	CO	118	$40,103	$42,634	0.94	$43,030	$43,809
Intermountain Medical Center	Salt Lake City	UT	163	$44,055	$47,894	0.92	$48,724	$48,706
Geisinger Medical Center	Danville	PA	95	$51,232	$52,490	0.98	$51,384	$49,810
Dartmouth-Hitchcock Med Center	Lebanon	NH	212	$55,906	$50,091	1.12	$50,004	$50,506
LDS Hospital	Salt Lake City	UT	102	$43,556	$48,681	0.89	$50,636	$51,129
New York-Presbyterian Hospital	Manhattan	NY	541	$69,697	$52,485	1.33	$53,104	$52,940
University of Chicago Hospital	Chicago	IL	59	$64,918	$56,281	1.15	$53,874	$53,679
Cleveland Clinic Foundation	Cleveland	OH	60	$53,011	$55,951	0.95	$54,209	$54,056
Boston Medical Center	Boston	MA	71	$71,051	$57,312	1.24	$55,971	$54,319
Stanford Hospital and Clinics	San Mateo County	CA	181	$74,835	$53,189	1.41	$53,936	$54,331
UCSF Medical Center	San Francisco	CA	98	$81,261	$55,962	1.45	$55,516	$55,136
Massachusetts General Hospital	Boston	MA	304	$65,798	$58,742	1.12	$58,892	$58,788
NYU Medical Center	Manhattan	NY	197	$72,991	$61,116	1.19	$62,109	$61,272
Cedars-Sinai Medical Center	Los Angeles	CA	362	$70,133	$61,839	1.13	$62,578	$61,339
UCLA Medical Center	Los Angeles	CA	73	$78,319	$66,320	1.18	$66,130	$65,546
University of Miami Hospital	Miami	FL	76	$79,708	$75,791	1.05	$72,729	$70,294
Doctor's Hospital at Renaissance	McAllen	TX	61	$77,575	$81,446	0.95	$78,292	$77,687
Edinburg Regional Medical Center	McAllen	TX	184	$83,605	$82,332	1.02	$79,853	$79,349

Table 5: One-Year Medicare Expenditures for Five Cohorts: Selected Hospitals

Hospital	HRR City	State	Hip Fracture	AMI	Stroke	Colon Cancer	End-of-Life (2 yrs)	C-Factor
Gundersen Lutheran Med Center	La Crosse	WI	$37,239	$39,168	$30,419	$45,602	$43,422	-1.76
St. Mary's Hospital	Rochester	MN	$43,390	$40,427	$32,132		$59,517	-0.67
St. Mary's Hosp & Med Center	Grand Junction	CO	$43,809	$32,887	$27,888		$49,142	-1.53
Intermountain Medical Center	Salt Lake City	UT	$48,706	$41,601	$35,432		$61,300	-0.25
Geisinger Medical Center	Danville	PA	$49,810	$40,929	$34,874	$55,353	$51,877	-0.68
Dartmouth-Hitchcock Med Center	Lebanon	NH	$50,506	$38,803	$37,991	$55,295	$53,004	-0.61
New York-Presbyterian Hospital	Manhattan	NY	$52,940	$48,360	$53,218	$67,994	$72,695	1.13
University of Chicago Hospital	Chicago	IL	$53,679	$44,235	$45,736	$68,727	$69,233	0.67
Cleveland Clinic Foundation	Cleveland	OH	$54,056	$49,304	$41,315	$58,225	$64,834	0.57
Boston Medical Center	Boston	MA	$54,319	$46,309	$56,423		$70,068	1.05
Stanford Hospital and Clinics	San Mateo County	CA	$54,331	$42,232	$49,614	$69,818	$56,027	0.08
UCSF Medical Center	San Francisco	CA	$55,136	$36,936	$56,602	$72,283	$60,314	0.27
Massachusetts General Hospital	Boston	MA	$58,788	$51,554	$46,918	$58,772	$75,072	1.46
NYU Medical Center	Manhattan	NY	$61,272	$50,486	$56,207	$72,411	$92,840	2.57
Cedars-Sinai Medical Center	Los Angeles	CA	$61,339	$49,881	$51,065	$66,575	$103,898	2.95
UCLA Medical Center	Los Angeles	CA	$65,546	$49,189	$56,133		$80,458	2.11
University of Miami Hospital	Miami	FL	$70,294	$64,499	$60,988		$105,241	4.17
Doctor's Hospital at Renaissance	McAllen	TX	$77,687	$49,808	$51,995	$69,499	$93,447	3.14
Edinburg Regional Medical Center	McAllen	TX	$79,349	$54,639	$62,081		$87,468	3.35
U.S. Average for Sample			*$49,171*	*$42,656*	*$35,452*	*$50,562*	*$59,565**	*0.00*

* End-of-life national average weighted by hip fracture sample sizes in each hospital (for hospitals with N > 50 patients).

Table 6: Composition of One-Year Medicare Spending, by Component, for High/Low Cost Medical Centers and Non-Medical Centers

	Medical Center				*Non-Medical Center*			
	Low Cost	**High Cost**	**Diff.**	**%**	**Low Cost**	**High Cost**	**Diff.**	**%**
Total Spending	$47,235	$62,167	$14,932	100	$42,255	$81,805	$39,550	100
Index Part A	21,475	21,392	($83)	-1%	20,358	23,264	$2,906	7%
Index Part B	2,131	3,223	$1,092	7%	1,912	3,148	$1,236	3%
Post-Acute Part A	14,373	19,762	$5,389	36%	10,709	31,590	$20,881	53%
Post-Actute Part B	4,650	9,935	$5,285	35%	3,668	10,903	$7,235	18%
PA: Outpatient & DME	2,092	2,327	$235	2%	2,172	3,469	$1,297	3%
PA: Home Health Care	1,952	5,049	$3,097	21%	2,162	9,083	$6,921	17%
PA: Hospice	562	478	($84)	-1%	1,266	348	($918)	-2%
Memo:			Diff.				Diff.	
Index Inpatient visits	10	21	11		7	25	18	
30-Day Readmission	29%	51%	22%		27%	70%	43%	

Appendix Table A.1: Risk Adjustment Estimate for Hip Fracture Cohort: Price-Adjusted One-Year Expenditures

Number of obs = 332649
R-squared = 0.1061
Adj R-squared = 0.1001
Root MSE = 30614

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Vascular	1880.825	218.8398	8.59	0.000	1451.906	2309.745
Pulmonary	1310.503	130.1873	10.07	0.000	1055.34	1565.666
Dementia	-3036.578	149.7124	-20.28	0.000	-3330.01	-2743.146
Diabetes	3830.303	136.329	28.10	0.000	3563.102	4097.504
Liver dis.	1020.71	685.7457	1.49	0.137	-323.3317	2364.752
Renal fail.	7213.674	170.5085	42.31	0.000	6879.483	7547.866
Cancer_local	-916.975	241.7041	-3.79	0.000	-1390.708	-443.24
cx_regional	2846.192	2829.392	1.01	0.314	-2699.335	8391.719
cx_distant	-8662.222	490.6593	-17.65	0.000	-9623.9	-7700.544
Native	-77.45791	958.5093	-0.08	0.936	-1956.108	1801.193
Hispanic	633.7289	544.6968	1.16	0.245	-433.861	1701.319
Other	-27.95769	794.6733	-0.04	0.972	-1585.495	1529.579
Asian	-2463.789	655.0596	-3.76	0.000	-3747.687	-1179.891
Income	.003924	.0090092	0.44	0.663	-.0137338	.0215818
inc2	328.393	292.354	1.12	0.261	-244.6124	901.3984
inc3	74.5339	317.5355	0.23	0.814	-547.8266	696.8944
inc4	192.9977	339.3566	0.57	0.570	-472.1314	858.1269
inc5	-683.8335	361.6623	-1.89	0.059	-1392.681	25.01409
inc6	-217.8218	387.8532	-0.56	0.574	-978.003	542.3593
inc7	-231.7685	426.7753	-0.54	0.587	-1068.236	604.6987
inc8	-313.4018	471.9653	-0.66	0.507	-1238.44	611.6367
inc9	-66.30068	543.8409	-0.12	0.903	-1132.213	999.6119
inc10	-324.2786	719.418	-0.45	0.652	-1734.317	1085.76
Poverty	3141.014	1253.215	2.51	0.012	684.7476	5597.28
ICD 82092	-2325.452	2447.856	-0.95	0.342	-7123.18	2472.276
ICD 82009	-37.78872	158.3173	-0.24	0.811	-348.0862	272.5087
ICD 82019	-104.2963	2722.398	-0.04	0.969	-5440.118	5231.525
ICD 82029	1367.236	138.4914	9.87	0.000	1095.796	1638.675
ICD 82039	423.4285	1729.983	0.24	0.807	-2967.289	3814.146
hcc2	8505.218	274.7198	30.96	0.000	7966.775	9043.661
hcc3	12387.84	275.1824	45.02	0.000	11848.49	12927.19
hcc4	13942.51	286.6334	48.64	0.000	13380.72	14504.31
hcc5	15342.53	301.7525	50.84	0.000	14751.1	15933.95
hcc6	16538.44	290.5905	56.91	0.000	15968.89	17107.99
hcc7	18439.4	318.9843	57.81	0.000	17814.2	19064.6
hcc8	20440.36	350.49	58.32	0.000	19753.41	21127.31
hcc9	23491.78	320.3102	73.34	0.000	22863.98	24119.58
hcc10	31805.57	516.2382	61.61	0.000	30793.76	32817.38
Constant	37734.43	510.8254	73.87	0.000	36733.23	38735.63
Provider	F(2168, 330419) =		6.868	0.000	(2169 categories)	

NOTES: inc(j) refers to the jth zip code income decile, Poverty is the fraction of the zip code under the poverty line, and hcc(j) is the jth interval of the HCC measure. Provider dummy variables included.

Appendix Table A.2: Risk Adjustment Estimate for Hip Fracture Cohort: One-Year Mortality

Number of obs = 332649
R-squared = 0.1293
Adj R-squared = 0.1234
Root MSE = .4226

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
vascular	-.0034082	.0030209	-1.13	0.259	-.0093291	.0025127
pulmononary	.0507638	.0017971	28.25	0.000	.0472414	.0542861
dementia	.1398672	.0020667	67.68	0.000	.1358166	.1439178
diabetes	-.0361835	.0018819	-19.23	0.000	-.039872	-.032495
liver dis.	.1822793	.0094662	19.26	0.000	.1637259	.2008328
renal fail.	.0778764	.0023537	33.09	0.000	.0732631	.0824896
cancer_local	.0617708	.0033365	18.51	0.000	.0552313	.0683103
cx_regional	.1760597	.0390575	4.51	0.000	.0995081	.2526113
cx_distant	.3743226	.0067732	55.27	0.000	.3610474	.3875979
Native	-.031209	.0132315	-2.36	0.018	-.0571423	-.0052757
Hispanic	-.0304886	.0075191	-4.05	0.000	-.0452258	-.0157513
Other	-.0193902	.0109698	-1.77	0.077	-.0408908	.0021103
Asian	-.0394791	.0090426	-4.37	0.000	-.0572023	-.0217559
Income	7.45e-08	1.24e-07	0.60	0.549	-1.69e-07	3.18e-07
inc2	-.0041173	.0040357	-1.02	0.308	-.0120272	.0037926
inc3	-.0049746	.0043833	-1.13	0.256	-.0135658	.0036165
inc4	-.0004546	.0046845	-0.10	0.923	-.0096362	.008727
inc5	.0012611	.0049925	0.25	0.801	-.008524	.0110462
inc6	-.0022796	.005354	-0.43	0.670	-.0127733	.0082141
inc7	-.0075776	.0058913	-1.29	0.198	-.0191243	.0039692
inc8	-.0097504	.0065151	-1.50	0.135	-.0225199	.003019
inc9	-.0051175	.0075073	-0.68	0.495	-.0198316	.0095966
inc10	-.0153471	.009931	-1.55	0.122	-.0348116	.0041174
Poverty	.0133211	.0172996	0.77	0.441	-.0205857	.0472279
ICD 82092	.0787049	.0337907	2.33	0.020	.0124761	.1449337
ICD 82009	-.0003878	.0021854	-0.18	0.859	-.0046712	.0038956
ICD 82019	.0007312	.0375805	0.02	0.984	-.0729256	.074388
ICD 82029	.0067747	.0019118	3.54	0.000	.0030277	.0105217
ICD 82039	-.011202	.0238811	-0.47	0.639	-.0580082	.0356042
hcc2	.0177289	.0037923	4.67	0.000	.0102961	.0251617
hcc3	.0434644	.0037987	11.44	0.000	.0360191	.0509097
hcc4	.0749946	.0039567	18.95	0.000	.0672395	.0827497
hcc5	.1003246	.0041655	24.08	0.000	.0921604	.1084888
hcc6	.1325448	.0040114	33.04	0.000	.1246826	.140407
hcc7	.1723575	.0044033	39.14	0.000	.1637271	.1809879
hcc8	.2177084	.0048382	45.00	0.000	.2082256	.2271912
hcc9	.2824854	.0044216	63.89	0.000	.2738191	.2911517
hcc10	.3594636	.0071263	50.44	0.000	.3454964	.3734309
_cons	.0520845	.0070515	7.39	0.000	.0382637	.0659053
provider	F(2168, 330419) =		1.690	0.000	(2169 categories)	

NOTES: inc(j) refers to the jth zip code income decile, Poverty is the fraction of the zip code under the poverty line, and hcc(j) is the jth interval of the HCC measure. Provider dummy variables included.

Appendix Table A.3a: Factor Analysis for Expenditures

Note: These are for hip fractures, AMI, stroke, and end-of-life spending

Expenditures

Factor analysis/correlation — Number of obs = 1706
Method: principal factors — Retained factors = 2
Rotation: (unrotated) — Number of params = 6

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor1	2.48591	2.48565	1.0996	1.0996
Factor2	0.00026	0.07997	0.0001	1.0997
Factor3	-0.07971	0.06606	-0.0353	1.0645
Factor4	-0.14577	.	-0.0645	1.0000

LR test: independent vs. saturated: chi2(6) = 3655.99 Prob>chi2 = 0.0000

Factor loadings (pattern matrix) and unique variances

Variable	Factor1	Factor2	Uniqueness
End of Life	0.8788	-0.0048	0.2278
Hip Fracture	0.8115	-0.0098	0.3413
AMI	0.7666	0.0086	0.4122
Stroke	0.6837	0.0081	0.5325

Scoring coefficients (method = regression)

Variable	Factor1	Factor2
End of Life	0.43113	-0.01201
Hip Fracture	0.26078	-0.02195
AMI	0.21667	0.02155
Stroke	0.15762	0.01518

Appendix Table A.3b: Factor Analysis for Mortality

Note: These are for hip fractures, AMI, and stroke.

Mortality

```
.
. factor hpadjmort1 amadjmort1 stadjmort1 [aw=wtfactor]
(sum of wgt is   3.1850e+05)
(obs=1711)

Factor analysis/correlation                       Number of obs    =     1711
    Method: principal factors                     Retained factors =        1
    Rotation: (unrotated)                         Number of params =        3

    --------------------------------------------------------------------------
         Factor  |   Eigenvalue   Difference        Proportion   Cumulative
    -------------+------------------------------------------------------------
        Factor1  |      0.64263      0.77672            1.9142       1.9142
        Factor2  |     -0.13409      0.03874           -0.3994       1.5148
        Factor3  |     -0.17283            .           -0.5148       1.0000
    --------------------------------------------------------------------------
    LR test: independent vs. saturated:  chi2(3)  =  327.99 Prob>chi2 = 0.0000

Factor loadings (pattern matrix) and unique variances

    ----------------------------------------
        Variable |  Factor1 |   Uniqueness
    -------------+----------+--------------
    Hip Fracture |   0.4787 |      0.7709
    AMI          |   0.4390 |      0.8073
    Stroke       |   0.4698 |      0.7793
    ----------------------------------------

. predict m1
(regression scoring assumed)

Scoring coefficients (method = regression)

    ------------------------
        Variable |  Factor1
    -------------+----------
    Hip Fracture |  0.31657
    AMI          |  0.28158
    Stroke       |  0.30839
    ------------------------
```