Published in final edited form as:

Med Care. 2009 October ; 47(10): 1098–1105. doi:10.1097/MLR.0b013e3181993191.

Development and Validation of Hospital “End-of-Life” Treatment Intensity Measures

Amber E. Barnato, MD, MPH, MS^{*,†,||}, Max H. Farrell^{*}, Chung-Chou H. Chang, PhD^{*,‡}, Judith R. Lave, PhD^{||,†}, Mark S. Roberts, MD, MPP^{*,†,||}, and Derek C. Angus, MD, MPH^{||,*,†}

^{*} Department of Medicine, University of Pittsburgh, Pittsburgh, PA

[†] Department of Health Policy and Management, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA

[‡] Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA

^{||} The CRISMA (Clinical Research, Investigation, and Systems Modeling of Acute Illness) Laboratory, Department of Critical Care Medicine, University of Pittsburgh, Pittsburgh, PA

Abstract

Background—Health care utilization among decedents is increasingly used as a measure of health care efficiency, but decedent-based measures may be biased estimates of care received by “dying” patients.

Objective—To develop and validate new measures of hospital “end-of-life” treatment intensity.

Research Design—Retrospective cohort study using Pennsylvania Health Care Cost Containment Council (PHC4) discharge data (April 2001–March 2005) and Centers for Medicare and Medicaid Services (CMS) data (January 1999–December 2003).

Subjects—Patients 65 and older admitted to 174 Pennsylvania acute care hospitals.

Measures—Hospital-specific standardized ratios of intensive care unit (ICU) and life-sustaining treatment (LST) use among terminal admissions (decedents) and admissions with a high probability of dying, and spending and use of hospitals, ICUs, and physicians among patients in their last 6 months of life.

Results—There was marked between-hospital variation in the use of the ICU and LSTs among decedents and admissions with high probability of dying. All hospital decedent and high probability of dying measures were highly correlated ($p < 00001$). In principal components factor analysis, all 4 of the last-6-months cohort-based measures, the decedent and high-risk admission-based ICU measures, and 8 of the 12 decedent and high probability of dying LST measures loaded onto a single factor, explaining 42% of the variation in the data.

Conclusions—Hospitals’ end-of-life intensity varies in the use of specific life-sustaining treatments that are somewhat emblematic of aggressive end-of-life care. End-of-life intensity is a relatively stable hospital attribute that is robust to multiple measurement approaches.

Keywords

intensive care units; terminal care; life support care; hospitals

There is substantial enthusiasm for using claims-based measures to profile the quality and efficiency of US hospitals. Recently, researchers have used claims data to measure the intensity of treatment among patients who died. Given that they found wide regional and hospital-level variations in intensity, they have concluded that high intensity care is of poor quality (1) or inefficient (2–7). The general assumption is that the provision of intensive and expensive care to those who died was a waste of resources. Further, because population-based surveys suggest the majority of Americans prefer not to have intensive hospital-based care when they die,(8, 9) such treatment appears misaligned with patient preferences.

However, these conclusions may be wrong if these measures, which are based upon the decedent follow-back (case-series) approach, are unreliable estimates of the care received by patients who providers believed were “dying” (10). Many deaths are unexpected. Intensive treatment for patients who were not expected to die (but did) has different implications for quality and efficiency than intensive treatment for patients who were expected to die. Furthermore, it is possible that intensity of treatment actually improves survival. If this were true, then the composition of the decedent cohorts used to measure each hospital’s intensity may not be comparable; that is, patients who were “saved” by the higher intensity hospital are missing from their decedent cohort.

The purpose of the current study is to describe the development of new measures of hospital “end-of-life” treatment intensity using Pennsylvania discharge data and to compare these new measures with existing measures of end-of-life intensity with the greatest public policy prominence, Dartmouth Atlas measures (www.dartmouthatlas.org). Our new measures leverage a unique feature of Pennsylvania Health Care Cost Containment Council (PHC4) discharge data, a predicted risk of death upon admission based upon key clinical findings abstracted from the medical chart within the first 48 hours of admission, to identify patients who providers might reasonably expect to die during the hospitalization. The comparative analyses herein are descriptive and do not include metrics for determining which measure of intensity is superior. Rather, we seek to explore the degree to which different measures could create divergent profiles of hospital intensity and, in turn, divergent and potentially contradictory targets for policy intervention.

METHODS

Data and Sample

The sample included all acute care hospitals in Pennsylvania (PA). Hospital-level treatment intensity information is based upon four years of Pennsylvania Health Care Cost Containment Council (PHC4) discharge data (April 2001 through March 2005), five years of Centers for Medicare and Medicaid Services (CMS) Medicare fee-for-service claims Part A hospital claims data (January 1999–December 2003), and four years of CMS fee-for-service Part B physician claims data (January 2000–December 2003). We describe the intensity measures from each of these sources in greater detail, below.

PHC4 admission-based intensity measures

We chose PHC4 data for the current study because it contains a unique data element: the predicted probability of inpatient death at the time of admission. This predicted probability is based upon key clinical findings abstracted in the first 48 hours of admission by specially trained personnel using Medi-Qual’s proprietary Atlas clinical information system. Key clinical findings include historical information, laboratory results, vital signs, clinical symptoms and signs, pathophysiologic abnormalities, and composite pathophysiological scores. As reported in Medi-Qual technical white papers and by peer-reviewed independent

assessment (11), inpatient mortality prediction using key clinical findings have better discrimination than administrative data-only models.

We calculated each hospital's standardized (observed-to-expected) ratio of ICU and life-sustaining treatments (LSTs) among two populations: 1) those who died during the admission (decedents), and 2) those who had a high probability of dying during the admission. We defined "high" as the upper 95th percentile of predicted probability of inpatient death at the time of admission, equivalent to a risk of 0.21 or higher (mean 0.41). We restricted our analyses to hospitals with at least 50 decedent and 50 high probability admissions during the four study years.

The intensity measures that are based upon treatment of decedents use the mortality follow-back, or decedent case series, approach. The intensity measures based upon treatment of patients with a high probability of dying are designed to approximate a cohort of subjects that clinicians might prospectively identify as "dying" (10).

Observed treatment—We calculated observed hospital-level rates of ICU admission and ICU length of stay (LOS) using revenue codes for general and subspecialty intensive care units (174, 175, 200–204, 206–214, and 219). We calculated observed LST use using ICD-9-CM procedure codes in any of the 6 data fields for inpatient procedures: intubation and mechanical ventilation (ICD-9 codes 96.04, 96.05, 96.7x), tracheostomy (ICD-9 codes 31.1, 31.21, 31.29), gastrostomy tube insertion (ICD-9 codes 43.2, 43.11, 43.19, 43.2, 44.32), hemodialysis (ICD-9 code 39.95), enteral or parenteral nutrition (ICD-9 codes 96.6 and 99.15), and cardiopulmonary resuscitation (CPR, ICD-9 codes 99.60, 99.63). We chose these LSTs based upon the judgment of a multidisciplinary group of internists, intensivists, emergency physicians, palliative care physicians, and health services researchers, and reinforced by semi-structured interviews with over 100 informants in 11 PA hospitals (12). To our knowledge, the reliability of coding for each of these procedures has not been specifically explored, but we expect that tracheostomy and mechanical ventilation are reliably coded because they have important implications for reimbursement and are easily verified under audit (thus, overcoding is susceptible to fraud detection and undercoding is financially inefficient).

Expected treatment—Next, we calculated expected hospital-level rates of ICU admission, ICU LOS, intubation and mechanical ventilation, tracheostomy, gastrostomy tube insertion, hemodialysis, enteral or parenteral nutrition, and cardiopulmonary resuscitation. We used data pooled across all hospitals over the four years to identify patient-level predictors of ICU admission, ICU LOS, and receipt of the 6 LSTs in the two populations (decedent and high probability of dying). Variables eligible for model inclusion included patient demographics (age, sex, race), illness severity (predicted probability of dying during the hospitalization), diagnosis (25 diagnoses most prevalent among decedents, grouped using the clinical classification software (CCS) developed by the Agency for Healthcare Research and Quality), and coronary artery bypass grafting (a common procedure associated with ICU and MV use). We used stepwise procedures for model selection to avoid biased estimations and inferences due to correlations among these covariates. The p-value of the Hosmer-Lemeshow goodness of fit test was < 0.0001 for all models. The logistic regression models each had an area under the receiver operating curve between 0.70 and 0.83 and the ICU LOS linear model had an R^2 of 0.1 and median residual of -0.15 days.

We applied the coefficients obtained from each of those models to calculate a predicted probability of ICU admission and receipt of each of 6 LSTs and a predicted ICU LOS for those with an ICU stay for each decedent and high probability of dying admission. We summed the predicted probabilities of ICU admission, individual LSTs, and ICU LOS across the decedent and high probability of dying admissions at each hospital to obtain the expected hospital-level

rates of ICU admission, LSTs, and ICU LOS for decedent and high probability of dying admissions.

Observed-to-expected treatment ratios—Finally, we calculated 16 observed-to-expected treatment ratios for each hospital (8 measures in 2 groups: decedents and high probability of dying). To address instability of estimates, particularly for smaller hospitals, we used empiric Bayes shrinkage estimation (13). The advantage of shrinkage estimation is that it reduces the likelihood of inappropriately labeling a hospital as an “outlier” relative to other hospitals. A Bayesian framework uses prior knowledge about a situation to produce estimates for the true mean that lie somewhere between the observed average and the expected mean based on prior knowledge. Empirical Bayes uses the data both as the basis for prior knowledge and to adjust this knowledge; the expected mean used to seed the iterative analysis in our calculations was the mean in the overall sample.

Dartmouth Atlas CMS cohort-based measures

We chose the Dartmouth Atlas hospital-level intensity measures for comparison to our own because they have been widely published and have garnered the attention of policymakers who seek to improve the efficiency of the US health care system. Although there is no agreed-upon “gold standard” of measuring “end-of-life” treatment intensity, the Dartmouth Atlas measures clearly have the highest public visibility (14).

The Dartmouth measures differ from our PHC4 measures in several ways. First, they are not admission-based measures; they are cohort-based measures. Specifically, they represent cumulative acute care hospital and physician spending and utilization among a cohort of chronically ill elders during in the last 6 months of life who died between 1999 and 2003. Second, the patients in the cohort are restricted to continuously enrolled Medicare fee-for-service with at least one hospital admission for one of 11 diagnoses known to confer a high risk of inpatient mortality (15) during their last 2 years of life, whereas our measures include all patients over age 65, regardless of insurance or diagnosis. Finally, the Dartmouth measures assume that patients are loyal to a particular hospital, but the PHC4 measures do not. Each member of the Dartmouth Atlas end-of-life cohort was assigned to the hospital at which they obtained the plurality of their inpatient care in the last 2 years of their life. The Atlas then attributes the beneficiaries’ acute care hospital and physician utilization and spending to that hospital, and reports hospital-specific last 6-month inpatient spending, hospital days, ICU days and physician visits for hospitals with at least 50 assigned beneficiaries over 5 years (5).

Statistical Analysis

We calculated Pearson’s correlations and conducted exploratory principal components factor analysis among the three types of end-of-life intensity measures: PHC4 decedent admission-based ICU and LST use, PHC4 high probability of dying admission-based ICU and LST use, and Dartmouth Atlas decedent cohort-based last 6 months spending, hospital, ICU, and physician utilization. We performed all analyses using STATA 9.1 (StataCorp, College Station, TX).

RESULTS

Hospital Characteristics

There were 193 acute care hospitals in PA during the study years, 174 (90%) hospitals had at least 50 decedent and 50 high probability of dying admissions upon which to base the PHC4 measures, 169 of these (97%) also had Dartmouth Atlas acute care measures of end-of-life treatment intensity and 159 (91%) had Dartmouth Atlas physician measures of end-of-life treatment intensity. Study hospitals represented a broad spectrum: small to large, non-teaching

to major academic centers, and lower to higher illness acuity, as reflected in the mean probability of inpatient mortality (Table 1). There was marked variation in “end-of-life” treatment intensity among hospitals, regardless of the intensity measure (Table 2).

Admissions/Patients

We depict the predicted risk of death upon admission among all admissions in the PHC4 data and among the subset of admissions during which the patient died in Figure 1. The PHC4 decedent admission-based ICU and LST measures were based on 114,957 terminal admissions, 39% of which were also at a high probability of dying. The PHC4 high probability of dying admission-based ICU and LST measures were based on 132,585 admissions, 34% of which resulted in death. The Dartmouth Atlas decedent cohort-based last 6 months utilization measures reflect the utilization of 252,201 Medicare fee-for-service beneficiaries in their last 6 months of life.

We found that the loyalty assumption upon which the Dartmouth Atlas decedent cohort-based last 6 month utilization measures are based is highly violated in PHC4 data. Specifically, 28.2% of patients over 65 are admitted to another hospital within 30 days from their discharge from a previous hospital. This hospital “disloyalty” increases with number of admissions; for example, among those with only 2 admissions over 4 years, 36.8% were seen in 2 different hospitals; among those with 6 or more admissions in 4 years, 58.9% were seen in 2 or more different hospitals.

Relationship between PHC4 decedent and high probability of dying measures

The PHC4 decedent and high probability of dying measures of hospital-level treatment intensity were highly correlated with one another ($P < 0.0001$; lowest $\rho = 0.65$ for gastrostomy tube placement; highest $\rho = 0.95$ for ICU admission; see Appendix Table 1 and Appendix Figure). Individual hospitals’ treatment patterns among all admissions, conditional upon predicted risk of death upon admission, generally followed a consistent pattern which underscored the hospital’s practice pattern as a “higher” or “lower” intensity hospital. For example, a large urban teaching hospital in Pittsburgh with an observed-to-expected ICU admission ratio among decedents of 1.21 and among high probability of dying admissions of 1.28, admitted a greater proportion of its patients to the ICU than the state average across all deciles of risk (Figure 2). A competing large urban teaching hospital 5.5 miles away by car with an observed-to-expected ICU admission ratio among decedents of 1.05 and among high probability of dying admissions of 1.00, admitted a lower proportion of its patients to the ICU than the state average across the lower 9 risk deciles and admitted near the statewide average for patients in the top 10% of admission risk of death ($>11.4\%$).

Similarly, in Philadelphia, a large urban teaching hospital with an observed-to-expected ICU admission ratio among decedents of 1.22 and among high probability of dying admission of 1.22 also admitted a greater proportion of its patients to the ICU than the state average across all deciles of admission risk of death. Less than 3 miles away, another large urban teaching hospital with an observed-to-expected ICU admission ratio among decedents of 0.94 and among high probability of dying admissions of 0.95 admitted fewer patients to the ICU than the state average for the lower 6 deciles of admission risk of death, and about the same as the state average for the top 4 deciles of admission risk of death.

Relationships among PHC4 measures and Dartmouth Atlas measures

Among the hospitals with PHC4 decedent and high **probability of dying** intensity measures as well as Dartmouth Atlas retrospective cohort-based measures, there were statistically significant positive correlations between twelve of the sixteen PHC4 measures and the four Dartmouth Atlas measures (Appendix Table 1). The four PHC4 measures that were not

correlated with the Dartmouth Atlas measures were the standardized ratios of enteral/parenteral nutrition and CPR among decedents and high probability of dying admissions. Consistent with these findings, principal component factor analysis of the 20 intensity items yielded 2 factors with eigenvalues > 1. Sixteen of the hospital intensity measures loaded onto the first factor, the enteral/parenteral nutrition and CPR measures loaded onto the second factor (Appendix Table 2). The first factor explained 42% of the variation in the data and the second factor explained another 13% of the variation.

Summary index of “end of life” intensity using PHC4 measures

We conceptualized the first factor as a latent construct, “treatment intensity.” We repeated the factor analyses restricted to the PHC4 decedent-based and to the PHC4 high probability of dying admission-based measures and used the factor loadings for the standardized ratios of ICU admission, ICU LOS, intubation/mechanical ventilation, tracheostomy, gastrostomy, and hemodialysis to calculate a normalized factor score, or index of intensity, for decedent admissions and for high probability of dying admissions. The decedent and high probability of dying indices, like their underlying components, were highly correlated (Figure 3).

DISCUSSION

In this analysis of Pennsylvania hospitals, we demonstrate that hospitals vary markedly in “end-of-life” intensive care unit (ICU) and life-sustaining treatment (LST) use. Furthermore, we demonstrate correlation among measures relying on retrospectively identified decedents and measures relying on prospectively identified probability of inpatient mortality and between our narrow procedure-based measures of intensity and the broader Dartmouth Atlas utilization measures. We conclude that “end-of-life” treatment intensity is a real hospital attribute and explore the implications of these findings for measurement in the paragraphs that follow.

“End-of-life” treatment intensity is a real hospital attribute

There is no gold standard for the measurement of “end-of-life” treatment intensity. That is because it is not clear what treatments or measures should contribute to the numerator (i.e., spending? ICU use? life-support?), and because it is not easy to determine, *ex ante*, who is at the end of their lives in order to identify the denominator. Moreover, despite the intuitive appeal of the widely-adopted Dartmouth Atlas measures, the validity of these measures has never been tested. The foremost contribution of the current study is the demonstration of convergent validity of multiple “end-of-life” intensity measures. These measures varied in their construction of the numerator (the particular treatments or measurements) and the denominator (those who died vs. those who might die). Prior work has explored the robustness of hospital “end-of-life” intensity outlier status to multiple measures in the numerator among cancer decedents (16), but no prior work has assessed correlation of measures with various definitions of the both the numerator and the denominator.

Our finding that 16 of the 20 measures we studied were significantly and often highly correlated suggests that hospital “end-of-life” treatment intensity is a real construct. Moreover, because we observed relative consistency in hospitals’ use of the ICU and LSTs across groups at lower and higher probability of dying, a hospital’s “end-of-life” intensity simply may be—as the Dartmouth Atlas argues—a measure of general treatment intensity; that is, an underlying approach to all patients.

Which measure is “best”?

The answer to this question depends on whether one is trying to identify a risk-adjusted marker of general treatment intensity or whether one is fundamentally interested in end-of-life treatment. Regarding the denominator, the high probability of dying measures are theoretically

better risk-adjusted markers of intensity because they compare “apples” to “apples.” This is particularly true if more intensive treatment results in a survival advantage, which would make the composition of decedent populations non-comparable between high- and low-intensity hospitals. With respect to understanding provider behavior for dying patients, the high probability of dying measure also theoretically better reflects the “real world” of decision making under conditions of uncertainty. Yet the decedent-based measures are much more feasible to obtain from administrative data. And since decedent-based measures are robust proxies for hospital treatment intensity for patients with a high probability of dying, it may be pragmatic to use decedent-based measures.

Notwithstanding this pragmatic perspective, we note that there are hospitals that fall “off-diagonal” in correlations between decedent- and high probability of dying measures. This may reflect noise, or it could be a window into a hospital’s capacity to “discriminate.” If this is the case, one might argue that jointly reported measures could provide additional information about hospital efficiency. Specifically, a hospital with an intensity of treatment among high probability of dying patients which is lower than its intensity of treatment among decedents may be triaging some of the highest risk patients for less intensive, perhaps palliative, treatment. That is, the high(er) decedent treatment intensity is due not to intensive treatment of hopelessly ill patients who died, but of patients who were initially at relatively low probability of dying but who died (e.g., those who the providers believed might be “saved,” or, alternately who “surprised” providers by dying).

Regarding the numerator, the Dartmouth Atlas measure of Medicare spending is attractive because it reflects the financial implications for the US Treasury. But for those interested in the treatment of dying patients, the underlying contributors to those costs (acute care admissions and physician visits) do not necessarily produce negative social costs; that is, these broad categories of treatment may be desirable, even among patients expected to die soon. On the other hand, particular life-sustaining treatments, such as mechanical ventilation, that most Medicare beneficiaries indicate they prefer to avoid when faced with less than a year to live, even if it briefly extends life (9), may produce negative social costs.

Finally, regarding whether to use a cohort-based approach like the Dartmouth Atlas or an admission-based approach like the one we used here, the correlation among these measures suggests that the frequently-violated assumption of loyalty may not produce substantial misspecification in the cohort-based measures.

Limitations of our study

Our study relies upon a predicted probability of dying estimate that we did not create ourselves, but which is derived from a proprietary model. Although available information suggests that the model is robust, the lack of transparency may raise doubts among some readers. Furthermore, even if the predicted probability of dying estimate were entirely trustworthy, model-based predictions of mortality risk and physician predictions are only modestly aligned (17), calling into question whether this measure is a reliable proxy for clinicians’ decision making under conditions of uncertainty that we sought to capture. Indeed, it is physician’s estimates of likelihood of survival in the intensive care unit and a high likelihood of poor cognitive function -- not objective severity of illness and organ dysfunction scores--that predict withdrawal of mechanical ventilation (18), one of the LSTs we studied.

We chose the 95th percentile of predicted probability of dying as our cut-off for patients at high probability of dying because the size of the population was approximately equivalent to the population of decedents. This group included many patients whose risk of survival was greater than 50/50 and for whom intensive treatment likely is warranted. Hypothetically, using the 99th percentile (> 50% predicted probability of dying) might have been better, but such

estimates would have been far too unstable, even with over 2 million observations. We did not explicitly compare the distribution of risk in the high probability of dying group between hospitals; however, we empirically adjusted for selection effects by standardizing our ratios based upon models that included patient-level risk of death data. It must be acknowledged that our approach to creating standardized ratios assumed that the overall Pennsylvania cohort represented expected “end-of-life” intensity. However, the norms of behavior in Pennsylvania are significantly different than in other regions of the US with lower (Oregon) and higher (New York) “end-of-life” utilization rates as measured by the Dartmouth Atlas.

Of concern, neither the Dartmouth Atlas measures or our PHC4 measures accounted separately for intermediate care and intensive care beds. Reassuringly, the most-recently released Dartmouth Atlas data (2008) parses out intermediate- and high-intensity ICU use and finds them to be correlated; however, variation in intermediate care beds accounts for a greater proportion of the total variation than high-intensity beds. Another concern is that our approach assumed homogeneity in each hospital’s intensity; however, it is likely that within-hospital variation in end-of-life intensity exists between ICUs (e.g., cardiothoracic surgery vs. medical). Finally, no studies have validated the reliability of coding for mechanical ventilation of < 96 hours duration and other life-sustaining treatments using chart-based audits, so these measures may be subject to coding inconsistencies. Indeed, the finding that CPR and enteral/parenteral nutrition did not correlate with the other LST measures suggests either that decision making for these LSTs is governed by different forces, or, more likely, is unreliably coded. This aberrancy in our findings deserves further study.

Conclusions

From a research and measurement perspective, we conclude that hospital “end-of-life” treatment intensity is a real attribute and that mortality follow-back studies of hospital behavior produce reasonably unbiased estimates of hospitals’ intensity among patients with a high probability of dying, at least in the current delivery environment in which hospitals do not appear to selectively target patients for whom intensive treatment is more likely to result in survival. From a policy perspective, wide variations in intensity raises further concerns about the efficiency of our health care system.

Future studies should explore the relationship of these measures to outcomes of real interest -- such as survival, quality-of-life, and patient and family satisfaction--in order to discern whether one or another of these correlated measures (or an amalgam of measures) is more apt at identifying hospital inefficiency.

Acknowledgments

Source of support: This study was funded by NIH grant K08 AG021921 (Barnato PI), with additional support from P01 AG019783 (Skinner PI) and 1UL1 RR024153 (Reis PI).

The authors acknowledge the Dartmouth Atlas (Elliott S. Fisher, Principal Investigator) for provision of data and Galen Switzer (University of Pittsburgh) and Douglas Staiger (Dartmouth College) for contributions to the analysis.

References

1. Earle CC, Park ER, Lai B, Weeks JC, Ayanian JZ, Block S. Identifying potential indicators of the quality of end-of-life cancer care from administrative data. *J Clin Oncol* 2003;21(6):1133–8. [PubMed: 12637481]
2. 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(7440):607. [PubMed: 15016692]

3. Wennberg JE, Fisher ES, Baker L, Sharp SM, Bronner KK. Evaluating the efficiency of california providers in caring for patients with chronic illnesses. *Health Aff (Millwood)*. 2005Suppl Web Exclusives:W5-526-43
4. Wennberg JE, Fisher ES, Stukel TA, Sharp SM. Use of Medicare claims data to monitor provider-specific performance among patients with severe chronic illness. *Health Aff (Millwood)*. 2004Suppl Web Exclusives:VAR5-18
5. Wennberg JE, Fisher ES, Sharp SM, et al. Center for Evaluative Clinical Sciences, Dartmouth Medical School. The Care of Patients with Severe Chronic Illness: An Online Report on the Medicare Program By the Dartmouth Atlas Project. 2006
6. 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. *Ann Intern Med* 2003;138(4):288-98. [PubMed: 12585826]
7. 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. *Ann Intern Med* 2003;138(4):273-87. [PubMed: 12585825]
8. Field, MJ.; Cassel, CK., editors. Committee on Care at the End of Life, Institute of Medicine. *Approaching Death: Improving Care at the End of Life*. Washington, D.C: National Academy Press; 1997.
9. Barnato AE, Herndon MB, Anthony DL, et al. Are regional variations in end-of-life care intensity explained by patient preferences? A Study of the US Medicare Population. *Med Care* 2007;45(5):386-93. [PubMed: 17446824]
10. Bach PB, Schrag D, Begg CB. Resurrecting treatment histories of dead patients: a study design that should be laid to rest. *Jama* 2004;292(22):2765-70. [PubMed: 15585737]
11. Pine M, Jordan HS, Elixhauser A, et al. Enhancement of claims data to improve risk adjustment of hospital mortality. *Jama* 2007;297(1):71-6. [PubMed: 17200477]
12. Rodriguez KL, Barnato AE, Arnold RM. Perceptions and utilization of palliative care services in acute care hospitals. *J Palliat Med* 2007;10(1):99-110. [PubMed: 17298258]
13. Ash, AS.; Schwartz, M.; Pekoz, EA. Chapter 12: Comparing outcomes across providers. In: Iezzoni, LI., editor. *Risk Adjustment for Measuring Health Outcomes*. Vol. 3. Chicago, IL: Academy Health; 2005. p. 297-333.
14. Winslow R. Care Varies at Top Medical Centers. *Wall Street Journal*. 2006 May 16;
15. Iezzoni LI, Heeren T, Foley SM, Daley J, Hughes J, Coffman GA. Chronic conditions and risk of in-hospital death. *Health Serv Res* 1994;29(4):435-60. [PubMed: 7928371]
16. Earle CC, Neville BA, Landrum MB, et al. Evaluating claims-based indicators of the intensity of end-of-life cancer care. *Int J Qual Health Care* 2005;17(6):505-9. [PubMed: 15985505]
17. Knaus WA, Harrell FE Jr, Lynn J, et al. The SUPPORT prognostic model. Objective estimates of survival for seriously ill hospitalized adults. Study to understand prognoses and preferences for outcomes and risks of treatments. *Annals of Internal Medicine* 1995;122(3):191-203. [PubMed: 7810938]
18. Cook D, Rocker G, Marshall J, et al. Withdrawal of mechanical ventilation in anticipation of death in the intensive care unit. *N Engl J Med* 2003;349(12):1123-32. [PubMed: 13679526]

Figure 1. Frequency of Predicted Probability of Death among all Admissions (Panel A) and among Terminal Admissions (Panel B) Age 65 and Older, Pennsylvania 2001–2005

The histogram on the left (Panel A) represents all admissions among Pennsylvanians over age 65 over the 4 study years that fall within particular bands of predicted probability of death upon admission, based upon the MediQual Atlas proprietary risk model. Admissions in the top 5% of risk of death of admission, corresponding to a predicted risk of 21% or higher, make up the high probability of dying admission group. The histogram on the right (Panel B) represents the total number of terminal admissions among Pennsylvanians over age 65 that fall within particular bands of predicted probability of death upon admission. Many terminal admissions occur among persons who had relatively low predicted probability of death upon admission. Thirty-nine percent of terminal admissions were among patients in the high probability of dying group.

Hospital	Beds	ICU O:E (decedents)	ICU O:E (high risk)
Pittsburgh hospital A (higher intensity)	603	1.21	1.28
Pittsburgh hospital B (lower intensity)	524	1.05	1.00
Philadelphia hospital C (higher intensity)	703	1.22	1.22
Philadelphia hospital D (lower intensity)	541	0.94	0.95

Figure 2. Crude ICU Admission Rate by Deciles of Admission Risk of Death, 4 Pennsylvania Hospitals and the State Average, 2001–2005

This graph depicts the crude ICU admission rate among patients over age 65 along the y-axis, against the patient's predicted probability of death upon admission, categorized in deciles of risk along the x-axis, for two Pittsburgh hospitals and two Philadelphia hospitals, compared to the state average (dotted line).

Figure 3. Correlation between Retrospective and “Prospective” PHC4 Indices of Hospital End-of-life Intensity, Pennsylvania 2001–2004

The figure is a plot of 172 hospitals. Each bubble represents one hospital and the size of the bubble is proportional to the number of admissions among patients over age 65 during the 4 years of observation. The hospital's index of intensity among high probability of dying (prospective “end of life”) admissions is represented along the y-axis and its factor-based index of intensity among decedent (retrospective “end of life”) admissions is represented along the x-axis. The index is a factor score of 6 underlying standardized (observed-to-expected) ratios: ICU admission, ICU length of stay, intubation/mechanical ventilation, tracheostomy, hemodialysis, feeding tube placement. Higher values reflect higher intensity. The cross-hatches represent 95% confidence intervals for each index value.

Appendix Figure. Correlation between Retrospective and “Prospective” PHC4 Measures of Hospital End-of-life Intensity, Pennsylvania 2001–2004

Each of the 8 panels is a plot of individual hospitals with each dot representing one hospital. The hospital’s standardized (observed-to-expected) ratio among high probability of death (prospective “end of life”) admissions is represented along the y-axis and its standardized ratio among decedent (retrospective “end of life”) admissions is represented along the x-axis.

Table 1

Characteristics of the Hospital Sample

Hospital Characteristic	N=174
Beds, mean (range)	253 (40–1064)
Teaching status	
Resident-to-bed ratio, mean (range)	0.1 (0–1)
Major teaching [*] , N (%)	25 (14.6)
Minor teaching [†] , N (%)	50 (29.2)
Non-teaching, N (%)	95 (56.1)
Percent of patients who have Medicaid, mean (range)	0.7 (0.03–9.2)
Mean admission risk of death among admissions 65+, % (range)	4.7 (2.7–7.3)

^{*} >0.25 residents/bed

[†] 0 < residents/bed < 0.25

Table 2

PHC4 and Dartmouth Atlas Measures of Hospital-level Variation in End-of-life Intensity

Intensity Measure	Median	Mean (SD)	Number of hospitals	
			95% CI of O-to-E < 1	95% CI of O-to-E > 1
PHC4 Decedent Admission Intensity Measures, 2001–2005 (N=174 hospitals)				
Terminal admissions at each hospital, N		589 (490)		
Standardized (observed-to-expected) utilization ratios among terminal admissions				
ICU admission	0.96	0.93 (0.28)	102	85
ICU LOS (N=172 hospitals)	1.32	1.40 (0.45)	21	153
Intubation/mechanical ventilation	0.89	0.89 (0.28)	111	75
Tracheostomy	0.50	0.71 (0.73)	128	55
Gastrostomy tube placement	0.84	0.89 (0.45)	96	78
Hemodialysis	0.86	0.78 (0.59)	108	77
Enteral/parenteral nutrition	0.51	1.01 (1.26)	112	76
CPR	0.59	0.91 (1.06)	119	66
PHC4 High Probability of Dying Admission Intensity Measures, 2001–2005 (N=174 hospitals)				
Admissions with a high probability of dying at each hospital, N		684 (585)		
Standardized (observed-to-expected) utilization ratios among high probability of dying admissions				
ICU admission	0.99	0.98 (0.29)	89	98
ICU LOS (N=172 hospitals)	1.24	1.33 (0.43)	30	146
Intubation/mechanical ventilation	0.92	0.93 (0.31)	100	82
Tracheostomy	0.60	0.76 (0.65)	119	59
Gastrostomy tube placement	0.88	0.89 (0.38)	102	78
Hemodialysis	0.87	0.82 (0.32)	104	79
Enteral/parenteral nutrition	0.43	1.07 (1.54)	113	69
CPR	0.55	0.91 (1.15)	110	70
Dartmouth Atlas Decedent Cohort Intensity Measures, 1999–2003 (N=169 hospitals)				
Decedents assigned to each hospital, N		1550 (997)		
Last 6 months' Medicare utilization				
Average inpatient reimbursement per decedent, \$	12861	15074 (6844)	-	
Average hospital days per decedent, days	14.3	14.4 (2.80)	-	
Average ICU days per decedent, days	3.1	3.8 (2.2)	-	
Average MD visits per decedent (2000–2003, N=159 hospitals), visits	36.7	38.1 (10.9)	-	

APPENDIX Table 1

Correlation Matrix among 20 Measures of Hospital End-of-life Treatment Intensity

		PHC4 Terminal Admission							PHC4 High Risk Admission							Dartmouth Atlas							
		ICU	LOS	MV	Trach	Gtube	HD	Nutr	CPR	ICU	LOS	MV	Trach	Gtube	HD	Nutr	CPR	\$	Hosp	ICU	MD		
PHC4 Terminal Admission	ICU	1.0																					
	LOS	0.4927	1.0																				
	MV	0.5453	0.5364	1.0																			
	Trach	0.4163	0.6563	0.5626	1.0																		
	Gtube	0.2428	0.3192	0.4554	0.3561	1.0																	
	HD	0.3262	0.3713	0.5279	0.4359	0.3301	1.0																
	Nutr	0.0490	0.1647	0.1103	0.0758	0.0067	0.0132	1.0															
PHC4 High Risk Admission	CPR	0.0360	0.0153	0.1755	0.0003	0.0540	0.1050	0.3453	1.0														
	ICU	0.9528	0.4624	0.5001	0.3548	0.2127	0.2730	0.0041	0.0175	1.0													
	LOS	0.6045	0.9265	0.5565	0.3548	0.2944	0.3857	0.1151	0.0173	0.6055	1.0												
	MV	0.4353	0.4658	0.8761	0.5570	0.3993	0.4989	0.0570	0.0938	0.4372	0.5149	1.0											
	Trach	0.3857	0.5553	0.6537	0.8054	0.3781	0.4761	0.1386	0.0251	0.3638	0.6192	0.6888	1.0										
	Gtube	0.2809	0.4123	0.5164	0.3926	0.6722	0.3838	0.1177	0.0091	0.2618	0.4168	0.4590	0.5305	1.0									
	HD	0.2697	0.2807	0.4752	0.2962	0.3097	0.8936	0.0091	0.1541	0.2635	0.3133	0.4551	0.3767	0.3430	1.0								
Dartmouth Atlas	Nutr	0.0945	0.1392	0.1797	0.0777	0.0115	0.0562	0.9177	0.3296	0.0573	0.1099	0.1393	0.1683	0.1683	0.0193	1.0							
	CPR	0.0627	0.1016	0.1883	0.0369	0.0371	0.0866	0.3958	0.9196	0.0593	0.1143	0.1071	0.0741	0.0811	0.1270	0.3836	1.0						
	\$	0.3740	0.5285	0.5017	0.7805	0.3020	0.3503	0.1727	0.0041	0.7319	0.5399	0.5317	0.6705	0.2778	0.2819	0.1909	0.0433	1.0					
	Hosp	0.4132	0.6154	0.5406	0.6019	0.3525	0.3836	0.0913	0.0083	0.4032	0.6121	0.4985	0.5947	0.4541	0.3686	0.1411	0.042	0.7263	1.0				
	ICU	0.6670	0.7399	0.5235	0.6251	0.2253	0.3204	0.0786	0.0443	0.6716	0.8158	0.4770	0.5533	0.3297	0.2710	0.0713	0.1293	0.6450	0.6725	1.0			
	MD	0.4831	0.5845	0.6282	0.6017	0.3094	0.4821	0.0256	0.0511	0.4653	0.6368	0.6206	0.6373	0.4356	0.4036	0.0072	0.0173	0.6023	0.7775	0.7019	1.0		

p-value ≤ 0.0002; p-value > 0.0002 (often <0.05)

Appendix Table 2

Factor Loadings for 20 Measures of Hospital End-of-life Treatment Intensity*

Intensity Measure	Factor 1	Factor 2
<i>PHC4-based decedent measures</i>		
Standardized ICU admission ratio, terminal admissions	0.6900	-0.0485
Standardized ICU LOS ratio, terminal admissions	0.7671	0.0123
Standardized Intubation/mechanical ventilation ratio, terminal admissions	0.7913	0.0085
Standardized Tracheostomy ratio, terminal admissions	0.7938	-0.0995
Standardized gastrostomy tube ratio, terminal admissions	0.4555	-0.1064
Standardized hemodialysis ratio, terminal admissions	0.5861	0.0947
Standardized enteral/parenteral nutrition ratio, terminal admissions	0.1330	0.7848
Standardized CPR ratio, terminal admissions	0.0850	0.7972
<i>PHC4-based high risk measures</i>		
Standardized ICU admission ratio, high risk admissions	0.6533	-0.0750
Standardized ICU LOS ratio, high risk admissions	0.8255	-0.0063
Standardized Intubation/mechanical ventilation ratio, high risk admissions	0.7875	-0.0646
Standardized Tracheostomy ratio, high risk admissions	0.8131	-0.1093
Standardized gastrostomy tube ratio, high risk admissions	0.5495	-0.0352
Standardized hemodialysis ratio, high risk admissions	0.5495	0.1759
Standardized enteral/parenteral nutrition ratio, high risk admissions	0.1636	0.7509
Standardized CPR ratio, high risk admissions	0.1580	0.8030
<i>Dartmouth Atlas measures</i>		
Average inpatient reimbursement per decedent, last 6 months of life	0.7506	0.0042
Average hospital days per decedent, last 6 months of life	0.7927	0.7927
Average ICU days per decedent, last 6 months of life	0.8355	-0.0516
Average MD visits per decedent, last 6 months of life	0.8142	-0.1619

* Among 154 hospitals for which all 20 measures were available