

ORIGINAL ARTICLE

Variation in length of stay within and between hospitals

Thom Walsh^{1, 2}, Tracy Onega^{2, 3, 4}, Todd Mackenzie^{2, 3}

1. The Dartmouth Center for Health Care Delivery Science, Lebanon. 2. The Dartmouth Institute for Health Policy and Clinical Practice, Lebanon. 3. Geisel School of Medicine, Department of Community and Family Medicine, Lebanon. 4. Norris Cotton Cancer Center, Dartmouth Hitchcock Medical Center, Lebanon.

Correspondence: Thom Walsh. Address: The Dartmouth Institute for Health Policy and Clinical Practice, 35 Centerra, Lebanon, NH 03766. E-mail: thom.walsh@dartmouth.edu

Received: January 14, 2014

Accepted: February 19, 2014

Online Published: March 3, 2014

DOI: 10.5430/jha.v3n4p53

URL: <http://dx.doi.org/10.5430/jha.v3n4p53>

Abstract

Background and objective: Variation in the delivery of health care services and the lack of association between greater utilization and higher quality care signal inefficient, low value care. The extent to which patient and hospital variables can explain variation in hospital length of stay is unclear.

Methods: We examined hospital inpatient length of stay using data from 684 hospitals and 5.4 million discharges in the 2007 Healthcare Cost and Utilization Project's Nationwide Inpatient Sample. We used a mixed effects model with a random effect for hospitals to quantify variation in length of stay due to differences within and between hospitals.

Results: The interquartile range of hospital mean LOS was 3.4 days (3.3-6.7). Fifty-nine percent of the overall variation in length of stay remained unexplained after adjustment for discharge-level disease status, illness-severity, regional poverty, hospital-level contextual factors (*e.g.* proportion of patients from low-income ZIP-codes, proportion uninsured), and structural variables (*e.g.* teaching status, urban or rural location). Seventy-seven percent of the explainable variation was due to differences between hospitals.

Conclusion: These findings indicate that wide variability in length of stay persists after adjustment for patient and hospital variables, signaling an opportunity for improved productivity and efficiency in the delivery of health care.

Key words

Length of stay, Utilization, Variation, Multi-level models

1 Background

The wide variation in societal spending on healthcare is a concern for policy makers, providers, and administrators. The Patient Protection and Affordable Care Act required further examination of the extent of variation among regions, hospitals, and providers^[1]. As a result, the Institute of Medicine was commissioned by the Department of Health and Human Services to conduct a detailed assessment of variation in health care spending^[2]. Of particular concern was the low value care suggested by the lack of association between the quality of care received in hospitals and the dollar amount spent to provide the care^[3]. Spending is the product of prices paid for each service and the volume utilized. Variation is considered warranted when higher spending or utilization can be associated with care that addresses patients' needs or

wants and unwarranted when it is not associated with illness severity or patient preferences^[4]. Unwarranted variation in the utilization of services points to potential for improvement and can provide insight into opportunities for increasing the efficiency of healthcare delivery^[5].

Several decades of evidence have identified substantial variations in health care utilization; most notably the Dartmouth Atlas has found that 58% of the variation in price-adjusted Medicare expenditures per enrollee can be explained by differences in utilization^[6]. The McKinsey Global Institute has also found that 85% of the U.S. healthcare spending can be attributed to utilization of hospital and physician care beyond what would be expected following adjustment for a country of our size, productivity, and wealth^[7]. But critics have faulted prior variation studies for lacking variables pertaining to patient illness severity and regional poverty that could determine whether the variation was warranted, or not. Prior variation work had also been criticized for being difficult to act upon because the data were aggregated to a regional, rather than hospital level^[2].

2 Methods

2.1 Data sources

The Healthcare Cost and Utilization Project (HCUP) is funded by The Agency for Healthcare Research and Quality to allow hospital-level analysis of all-payer costs. The Nationwide Inpatient Sample (NIS) is a dataset within HCUP and is the largest collection of all-payer, discharge-level data in the US^[8]. The sampling strategy used to identify participating hospitals is meant to approximate a 20% sample of all US community hospitals. While 100% of the inpatient discharges from the sampled hospitals are collected, there are no unique patient identifiers in HCUP data. Each discharge is one discrete record; thus, a patient hospitalized multiple times during one year will be present multiple times.

2.2 Discharge variables

The core NIS file contains demographic variables for each discharge's age, gender, race (white, black, Hispanic, or other) and ZIP-code based income quartile (< \$39,000; \$39,000 - \$47,999; \$48,000 - \$62,999; \$63,000 or higher). The expected payer for hospital care was recorded as Medicare, Medicaid, Private, Self-pay, or no-charge.

Each discharge record contains a diagnosis-related group (DRG) code. To facilitate risk-adjustment, the NIS includes variables for disease severity and risk of mortality. Severity and risk are reported for each discharge on an ordinal scale from one (minor) to four (extreme)^[9, 10]. In this manner, disease severity and risk of mortality allow fine tuning of a DRG. For example, a discharged patient with a DRG code for chest pain and also given a severity rating of "moderate" and mortality risk of "minor" can be distinguished from another patient with the same DRG code, but having severity and mortality risk ratings of "extreme". The number of diagnoses is also recorded for each encounter. Length of stay is calculated by counting the days between admission and discharge. A person admitted and discharged on the same day would have a LOS equal to one.

2.3 Hospital context variables

Hospital-level contextual variables were generated from NIS data by calculating the proportion of discharges at each hospital that were female, that had Medicaid, Medicare, or private insurance, were self-pay, or had no charge for the encounter. Self-pay and no-charge encounters were collapsed to form an uninsured category at each hospital. The proportion of discharges occurring for individuals from the lowest two ZIP-code income quartiles was also calculated. The data file also includes an area wage index that reflects the hospital wage level relative to the national average.

2.4 Hospital structure variables

The data file includes five structural characteristics for each hospital: bed size (small, medium, large), control (government, for-profit, not-for-profit, private), location (urban, rural), geographic region (Northeast, Midwest, West, and South), and teaching status (teaching or non-teaching).

2.5 Statistical analyses

We used a mixed effects model with a random effect for hospitals to quantify variation in LOS due to hospitals. We measured the extent to which variation is reduced when adjusted for the covariates described above. Discharge-level variables were modeled first; hospital-level contextual variables were then added, and finally hospital structural variables. We then determined the interquartile range for mean LOS, parsed the amount of variation explained by the addition of each variable set, and determine how much of the variation was due to differences between hospitals.

The LOS variable was right skewed. We used a logarithmic transformation for the analyses and back-transformed the output to natural units. Stata 11 software was used for the analysis.

3 Results

After adjusting for disease status, illness-severity, and regional poverty, contextual factors, and structural variables the interquartile range for LOS was 3.4 days (3.3 – 6.7). The fully adjusted model's R^2 was 0.41; meaning 59% of the variation in mean adjusted LOS per discharge within hospitals remained unexplained.

Table 1 shows the summary of 5.6 million discharges in 684 hospitals from 23 states stratified by hospital level variables. The average proportion of female discharges across various hospital sizes, locations, teaching status, geographic region, and differing ownership status is roughly 60%; the range across all hospitals is from 33% to 89%.

The proportion of discharges from low income ZIP-codes is highest, 66%, in rural hospitals and publically owned hospitals with 62%. Hospitals in the western region of the US have the smallest proportion of discharges from low income ZIP-codes, 26%. Across all structural strata, the range was from 0 to at least 95% of discharges from low income ZIP-codes.

Hospitals in the northeastern region have the highest proportion of discharges from high income ZIP-codes, 29%, followed by teaching hospitals with 26%. The range for rural hospitals is 0-39% and 0-74% in the Midwest.

The proportion of discharges with Medicare as the expected primary payer is greatest (55%) in small hospitals. The range across all hospital types is 0-99%. Discharges with Medicaid as the expected primary payer were most numerous in teaching hospitals, 23%, and in the western region, 22%, with a range of 0-66%. Private-pay discharges were also most numerous in teaching hospitals (35%), with a range of 0-94% across all hospitals. The average proportion of uninsured discharges is consistent across all strata at roughly 6%, the range within each strata, like all contextual variables, is wide from 0-61%.

Table 2 illustrates the coefficients and z-scores for the fully adjusted multi-level LOS model with random effects for hospitals. After adjusting for all other variables, an added year of age has a significant, but clinically insubstantial, decrease in LOS of 0.3%. Female gender was associated with a 6% increase in LOS over male gender. Black patients remain in the hospital 4% longer than whites do, while Hispanic patients remain 3% longer. Discharges from the highest ZIP-code income quartile had LOS that was 3% less compared to those from the lowest income quartile. A severe mortality risk (3rd category of 4) increased average adjusted LOS, but an extreme risk shortened it by 9%. As illness severity rating increased, LOS increased in a hierarchical manner with an extreme rating having a 47% increase in LOS compared to the mild referent category. Each additional diagnosis was associated with a 3% increase in LOS.

At the contextual level, a greater proportion of female patients was associated with a decrease in mean adjusted LOS. No other contextual variables had significant associations with LOS. At the structural level, rural location had a 9% LOS.

Table 1. The context of US hospitals (23 states 684 hospitals 5.6 million discharges)

	Mean Age (Range)	Mean Illness Severity	Female (%)	ZIP-code Income			Expected Payer			
				Low (%)	Mid (%)	High (%)	Medicare (%)	Medicaid (%)	Private (%)	Uninsured (%)
All Hospitals										
	47 (2-81)	1.88 (1.00-3.58)	59 (5-89)	30 (0-100)	46 (0-100)	23 (0-100)	36 (0-99)	21 (0-66)	34 (0-94)	6 (0-61)
Bed Size										
Small (n = 307)	47 (3-81)	1.84 (1.00-3.58)	59 (33-89)	44 (0-100)	44 (0-100)	12 (0-90)	55 (0-98)	13 (0-66)	25 (0-94)	6 (0-61)
Medium (171)	46 (6-72)	1.99 (1.42-3.26)	59 (42-71)	38 (0-100)	43 (0-98)	19 (0-94)	44 (10-89)	19 (0-63)	26 (3-65)	6 (0-35)
Large (206)	46 (6-72)	1.88 (1.54-3.21)	59 (41-72)	37 (0-100)	46 (0-99)	18 (0-94)	39 (0-93)	12 (0-61)	32 (2-73)	6 (0-36)
Hospital Location										
Rural (233)	50 (26-81)	1.84 (1.15-3.43)	61 (41-80)	66 (0-100)	32 (0-99)	2 (0-39)	54 (0-95)	17 (0-53)	21 (3-73)	6 (0-61)
Urban (451)	46 (2-80)	1.88 (1.00-3.58)	58 (33-89)	27 (0-99)	50 (1-100)	23 (0-95)	43 (0-99)	16 (0-66)	32 (0-94)	5 (0-36)
Teaching Status										
Non-teaching (557)	49 (2-81)	1.88 (1.00-3.58)	60 (33-89)	43 (0-100)	44 (0-100)	13 (0-95)	50 (0-99)	15 (0-63)	26 (0-94)	5 (0-61)
Teaching (127)	44 (5-71)	1.89 (1.36-3.21)	57 (36-78)	30 (0-98)	43 (2-95)	26 (0-94)	32 (0-94)	23 (0-66)	35 (2-75)	6 (0-36)
Hospital Region										
Northeast (125)	49 (19-75)	1.85 (1.36-3.31)	58 (41-78)	20 (0-93)	51 (3-99)	29 (0-95)	43 (3-93)	16 (0-52)	33 (5-62)	6 (0-36)
Midwest (110)	48 (2-80)	1.92 (1.15-3.58)	60 (45-75)	36 (0-100)	56 (0-100)	8 (0-74)	52 (0-93)	15 (0-61)	27 (4-84)	4 (0-17)
South (324)	46 (5-81)	1.90 (1.00-3.44)	60 (33-89)	55 (0-100)	36 (0-100)	9 (0-88)	50 (0-97)	16 (0-66)	24 (2-94)	7 (0-46)
West (126)	45 (5-79)	1.85 (1.24-3.30)	59 (36-78)	26 (0-98)	49 (2-98)	25 (0-95)	38 (0-99)	22 (0-63)	32 (0-75)	5 (0-61)
Hospital Control										
Public (116)	48 (33-81)	1.84 (1.50-2.78)	61 (47-80)	62 (0-100)	33 (0-100)	5 (0-95)	54 (0-94)	16 (0-50)	21 (4-73)	7 (0-42)
Private/ Non-Profit (131)	46 (4-79)	1.86 (1.32-3.15)	60 (41-78)	42 (0-99)	41 (0-100)	17 (0-89)	46 (0-95)	18 (0-63)	28 (2-70)	6 (0-17)
Private/ Profit (141)	51 (28-76)	1.90 (1.00-3.44)	59 (33-89)	46 (0-100)	42 (0-100)	12 (0-89)	53 (0-96)	13 (0-63)	26 (0-94)	5 (0-46)

Table 2. Fully adjusted hospital length of stay coefficients (mixed level model of log-transformed LOS with random effects for hospitals, 684 hospitals, 5.4million discharges)

		All-Payers		
		Coefficient	z-score	p-value
Discharge Covariates				
Age		-0.0254	-34.78	.000
Female		0.0620	20.49	.000
Race	White	Referent category		
	Black	0.0391	7.58	.000
	Hispanic	0.0265	5.32	.000
	Other	0.0474	7.51	.000
ZIP-code Income Quartile	Lowest	Referent category		
	Low	-0.0004	-0.09	.927
	High	-0.0065	-1.28	.199
	Highest	-0.0254	-4.65	.000
Mortality Risk	Mild	Referent category		
	Moderate	-0.0274	-5.77	.000
	Severe	0.0114	1.60	.109
	Extreme	-0.0879	-7.06	.000
Illness Severity	Mild	Referent category		
	Moderate	0.0425	11.05	.000
	Severe	0.1941	32.70	.000
	Extreme	0.4706	40.76	.000
No. of Diagnoses		0.0336	65.93	.000
Hospital Context Variables				
% Female		-0.4685	-2.91	.004
% Medicare		0.3778	1.58	.114
% Medicaid		0.3266	1.26	.209
% Private Pay		0.0628	0.27	.790
% Uninsured		-0.0431	-0.15	.880
% Lowest Income		-0.0234	-0.41	.679
% Low Income		-0.0950	-1.86	.063
% High Income		-0.1260	-2.26	.226
% Highest Income		Omitted due to collinearity		
Wage Index		-0.0872	-1.06	.288
Hospital Structural Variables				
Bed Size	Small	Referent category		
	Medium	-0.0111	-0.57	.570
	Large	0.0234	1.25	.211
Rural		0.0968	3.30	.001
Teaching Status		-0.0040	-0.17	.86

Note. The model is further adjusted for discharge-level DRGs, hospital ownership, and US region

The figure demonstrates that 33% of the overall variation in the LOS per discharge within a hospital can be explained by DRG coding. Additional discharge variables such as age, gender, ZIP-income, and measures of illness severity add an additional 3% explanatory power. Contextual variables explain very little, just 0.66%, while structural variables add just 0.79%.

Figure. Variation in mean discharge-level LOS explained by vectors of variables and hospital random effects

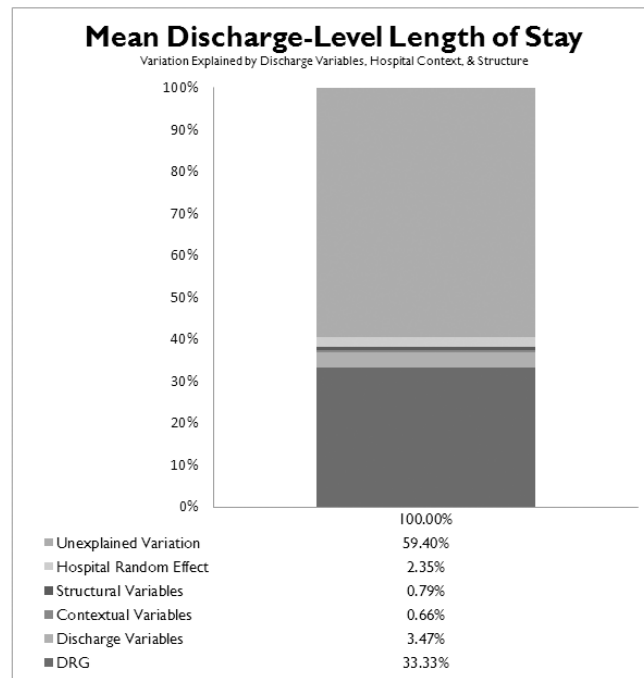


Table 3 details how the differences across hospitals affect overall variation in LOS. Recall, 41% of the overall variation in discharge-level LOS is explained by our model's covariates. Differences across hospitals in diagnostic coding explained 54% of the overall variation in LOS. The addition of discharge-level demographics and illness severity raised the explained proportion to 58%. Additional adjustment for differences in hospital context increased the explanatory power to 64% and hospital structure variables raised it further to 77%.

Table 3. Proportion of variation in LOS between hospitals explained by model covariates

Mixed Model with Random Effects for Hospitals	
Null Model	n/a
DRG-only Adjusted Model	54.40%
Add Discharge Variables to model	58.36%
Add Contextual Variables	64.28%
Add Structural Variables	76.63%

Note. Discharge variables: Age, Gender, Income, Number of Diagnoses, Mortality Risk, & Illness Severity; Hospital Contextual Variables: proportion of female patients, patients with Medicaid, uninsured patients, & patients from low income ZIP-code; Hospital Structure Variables: hospital wage differential, bedsize, ownership, urban location, US region, & teaching status.

4 Discussion

The majority of overall variation in mean LOS in hospitals could not be explained statistically using a multi-level model of discharge-level and hospital-level variables. Seventy-seven percent of that overall variation is due to differences across hospitals and over half of the differences across hospitals are due to diagnostic coding. A large proportion of unwarranted variation in utilization, as found here, is indicative of inefficient, low value care processes.

Moreover, our results are likely to overestimate the proportion of variation that can be considered warranted because the largest drivers of variation were related to diagnosis and severity ratings. These variables had the largest explanatory effects on variation in our analysis, but caution is advised regarding the interpretation of illness severity ratings.

In 2010, Song *et al.* took advantage of a natural experiment occurring when Medicare beneficiaries moved from lower to higher utilization regions and found a significant increase in severity rating for the same diagnosis ^[11]. Those patients moving from the lowest intensity region to the highest intensity region were coded as “sicker” than other patients, even though there were no substantial differences before the move. Patients who moved from high utilization regions to the lowest had the least increase in disease severity over the 3-yr follow-up period.

In 2011, Welch and colleagues performed a cross sectional analysis of variation in diagnoses for chronic conditions and the case-fatality rate using Medicare data ^[12]. As expected among patients, the case fatality rate rose in sync with the mean number of diagnoses. In contrast at the regional level, there was an inverse relationship between diagnostic frequency and risk of death; as the mean number of diagnoses increased, the case fatality rate declined. In other words, a patient with multiple diagnoses living in a region with a high mean number of diagnoses had a lower fatality rate than a similarly diagnosed patient in a region with less diagnostic intensity.

Taken together, the findings from Song and Welch suggest high utilization hospitals code patients as more severely diseased when compared to low utilization hospitals even though the patients’ true health status are similar. Our findings suggest that differences between hospitals’ LOS are susceptible to this bias because of differences in the way hospitals use DRG codes with ratings of illness severity and mortality risk.

Patient-reported measures of health status and illness severity could help diminish the potential for bias inherent in ratings that rely on diagnostic groups, particularly given that patients do not benefit financially from rating themselves sicker. Wennberg and colleagues have described an alternative method to diminish observational intensity bias that involves correcting illness severity ratings for the frequency of physician visits ^[13]. More research is needed in this area given the importance of baseline health status and illness severity in risk-adjustment models required for comparative effectiveness research and valid public reporting.

Our independent variables are extensive, not exhaustive. Other difficult to define and measure concepts, like leadership, communication, “practice patterns” and “culture”, may influence variation in utilization. While it is unlikely that administrative data will contain reliable and valid measures of these constructs, the addition of qualitative research techniques including ethnography and interviews with patients, providers, and administrators could help build a more robust understanding of differences in the delivery of health care services across hospitals.

5 Conclusion

Achieving high value health care is a concern for citizens, policy makers, providers, and administrators. Clinicians and scientists interested in the implementation of health care reform must move from identifying variation at regional levels to examining differences in utilization within and between hospitals. Reform efforts must include testable changes designed to increase the value of care delivered. To do so, the science of health care delivery will need to advance our understanding of illness severity measurement, patient-reported outcomes, and incorporate qualitative methodology to more fully explore variation in utilization.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements and funding

The manuscript was completed as partial fulfillment of the requirement for Dr. Walsh’s doctoral degree. Support for the work was provided by the Dartmouth Institute and the Dartmouth Center.

References

- [1] H.R. 3590 -- 111th Congress. (2010). Patient Protection and Affordable Care Act. September 03, 2013, Available from: <http://www.govtrack.us/congress/bills/111/hr3590>
- [2] Newhouse, J. P., Garber, A., Graham, R., McCoy, M., Mancher, M., Kibria, A. Variation in health care spending: Target decision making, not geography. Washington, DC. 2013.
- [3] Yasaitis, L., Fisher, E. S., Skinner, J. S., Chandra, A. Hospital quality and intensity of spending: is there an association? Health affairs (Project Hope). 2009; 28(4): w566-72. <http://dx.doi.org/10.1377/hlthaff.28.4.w566>
- [4] Mulley, A., Trimble, C., Elwyn, G. Patients' preferences matter: Stop the silent misdiagnosis. The Kings Fund, London, UK. 2012.
- [5] Wennberg, J. E. Unwarranted variations in healthcare delivery: implications for academic medical centres. BMJ (Clinical research ed.). 2002; 325(7370): 961-4. PMID: 12399352. <http://dx.doi.org/10.1136/bmj.325.7370.961>
- [6] Dartmouth Atlas. June 17, 2011. Available from: <http://www.dartmouthatlas.org/>
- [7] Farrell, D., Jensen, E., Kocher, B., Lovegrove, N., Melhem, F., Mendonca, L., *et al.* Accounting for the cost of US health care: A new look at why Americans spend more. Washington, DC. 2008.
- [8] Steiner, C., Elixhauser, A., Schnaier, J. The healthcare cost and utilization project: an overview. Effective clinical practice : ECP. 2002; 5(3): 143-51. PMID: 12088294.
- [9] Edwards, N., Honemann, D., Burley, D., Navarro, M. Refinement of the Medicare diagnosis-related groups to incorporate a measure of severity. Health care financing review. 1994; 16(2): 45-64. PMID: 10142374.
- [10] Leary, R., Johantgen, M., Farley, D., Forthman, M., Wooster, L. All-Payer severity-adjusted diagnosis-related groups: A uniform method to severity-adjust discharge data. Health Information and Management. 1997; 17(3): 60-71.
- [11] Song, Y., Skinner, J., Bynum, J., Sutherland, J., Wennberg, J. E., Fisher, E. S. Regional variations in diagnostic practices. The New England Journal of Medicine. 2010; 363(1): 45-53. <http://dx.doi.org/10.1056/NEJMsa0910881>
- [12] Welch, H. G., Sharp, S. M., Gottlieb, D. J., Skinner, J. S., Wennberg, J. E. Geographic variation in diagnosis frequency and risk of death among Medicare beneficiaries. JAMA. 2011; 305(11): 1113-8. <http://dx.doi.org/10.1001/jama.2011.307>
- [13] Wennberg, J. E., Staiger, D. O., Gottlieb D. J., Bevan, G., McPherson, K., Welch, G. H. Observational intensity bias associated with illness adjustment; cross sectional analysis of insurance claims. BMJ. 2013; 346: f549. PMID: 23430282. <http://dx.doi.org/10.1136/bmj.f549>