



Published in final edited form as:

N Engl J Med. 2012 November ; 367(18): 1724–1731. doi:10.1056/NEJMsa1203980.

Comparing Local and Regional Variation in Healthcare Spending

Dr. Yuting Zhang, PhD, Dr. Seo Hyon Baik, PhD, Dr. A. Mark Fendrick, MD, and Dr. Katherine Baicker, PhD

From the Department of Health Policy and Management, Graduate School of Public Health, University of Pittsburgh, 130 De Soto St, Pittsburgh, PA 15261 (Drs. Zhang and Baik); from the Department of Internal Medicine and Department of Health Management and Policy, University of Michigan School of Public Health, 300 NIB, 7E14, Ann Arbor, Michigan, 48109-0429 (Dr. Fendrick); and from the Department of Health Policy & Management, Harvard School of Public Health, 677 Huntington Ave, Boston, MA 02115 (Dr. Baicker)

Abstract

Background—Wide geographic variation in healthcare spending has generated concern about inefficiency and policy debate about geographic-based payment reform. Evidence on variation has focused on hospital referral regions (HRRs), which incorporate numerous local hospital service areas (HSAs). If there is substantial variation across local areas within HRRs, then policies focusing on HRRs may be poorly-targeted.

Methods—Using pharmacy and medical claims data from a 5% random sample of Medicare beneficiaries in 2006–2009, we compared variation in health care spending and utilization in 306 HRRs and 3436 HSAs. We adjusted for beneficiary-level demographics, insurance status, and clinical characteristics to calculate adjusted use and spending.

Results—There is substantial local variation in drug and non-drug utilization and spending, and substantial dispersion of local areas within HRRs; many low-spending HSAs are located within the borders of high-spending HRRs and vice versa. Only about half of the HSAs located within the borders of the highest spending quintile of HRRs are in the highest spending quintile of HSAs; conversely, only about half of the highest spending HSAs are located within the borders of the highest-spending HRRs.

Conclusions—The effectiveness of payment reforms in reducing overutilization while maintaining access to high-quality care depends crucially on the effectiveness of targeting. Our analysis suggests that HRR-based policies may be too crudely targeted to promote the best use of healthcare resources.

A substantial body of evidence has emerged highlighting wide geographic variation in healthcare spending that is not driven by patient characteristics and not associated with the quality of care or patient outcomes.^{1–7} In light of this evidence, many policy proposals suggest targeting high spending areas for lower Medicare payments or other coverage constraints, focusing on areas such as Dartmouth hospital-referral regions (HRRs). These policies are predicated on the idea that there is a system-level component driving some areas to be high-utilization and others low. The effectiveness of these policies in reducing overutilization while maintaining access to high quality care depends crucially on the

Address reprint requests to Dr. Zhang (corresponding author) at the Department of Health Policy and Management, University of Pittsburgh, 130 De Soto Street, Crabtree Hall A664, Pittsburgh, PA 15261; Phone: 412-383-5340; Fax: 412-624-3146; yzhang@pitt.edu.

Conflicts of interest

Baicker is a Commissioner on the Medicare Payment Advisory Commission and a director of Eli Lilly.

effectiveness of targeting: if there is substantial variation across local areas within HRRs, then focusing on high-cost HRRs may leave many high-spending locales untouched while inadvertently penalizing some low-spending locales.

We compare variation in medical spending and prescribing patterns at the broader market level, the 306 Dartmouth HRRs, to variation within those markets, the 3436 Dartmouth hospital-service areas (HSAs).¹ HRRs represent the areas served by large tertiary hospitals where patients were referred for major cardiovascular surgical procedures and for neurosurgery. HSAs are contained within HRRs and represent areas whose residents receive most of their hospitalizations from the hospitals in the area. Thus, HSAs better capture the local health care markets where Medicare beneficiaries receive most of their care. We examine: (1) variation in prescription drug and medical care spending and use across HSAs versus across HRRs; and (2) the degree to which high-spending HSAs are clustered together within high-spending HRRs. This analysis can help evaluate the effectiveness of policies targeted at different levels of aggregation; if the intention of such policies is to capture variation in local markets, it is important to understand the local-market heterogeneity that larger units may mask.

METHODS

Data source and study sample

We obtained 2006–2009 enrollment, drug event and medical claims data from the Centers for Medicare & Medicaid Services (CMS) for a 5% random sample of Medicare beneficiaries. For each year between 2007 and 2009, we identified all beneficiaries having at least one month enrollment in Parts A, B, and stand-alone Part D (PDP) plans because CMS only has both medical and drug data for those enrolled in PDP plans. (2006 was used for calculation of prospective risk scores described below.) The resulting sample consisted of 1,013,477 beneficiaries in 2007, 1,024,183 in 2008, and 1,022,662 in 2009, for 3,060,322 total beneficiary-year observations. We assigned each beneficiary to 1 of 306 HRRs or 3436 HSAs based on the beneficiary's ZIP-Code of residence.¹ HSAs are nested within HRRs. The study was approved by the Institutional Review Board at the University of Pittsburgh.

We conducted sensitivity analysis for two subpopulations: (1) those aged ≥ 65 and (2) those enrolled for the full year or until they died, to ensure that our results are not substantially affected by a small proportion of disabled beneficiaries or those switching plans within the year (Online Supplement Table S1; coefficients of variation and interquartile ratios are virtually unchanged).

Outcomes

Our outcomes were utilization of and spending on medical services and prescription drugs. All outcome measures were calculated in per-person-per-year units (with spending of part-year enrollees annualized). For medication use, we defined two outcomes: (1) total gross drug spending including Part D plan payment before rebates, beneficiary out-of-pocket spending, and subsidy amount; and (2) number of monthly prescription drugs ($=1$ if days supply ≤ 30 ; $=\text{days supply}/30$ if days supply >30). For medical services, we defined four outcomes: (1) total non-drug medical spending, (2) number of inpatient admissions; (3) number of outpatient office visits; and (4) number of emergency room (ER) visits. Total non-drug medical spending included Medicare and beneficiary payments for all medical services (including inpatient, outpatient, physician, home health, hospice, skilled nursing home, and medical devices) and was adjusted for local price-level differences using county-level factor prices given to us by the Medicare Payment Advisory Commission

(MedPAC).^{7,8} We did not adjust drug spending for regional price differences because the variation in drug prices among regions was negligible.⁴

Adjustment variables

To account for differences in population characteristics across regions, we adjusted for three major categories of beneficiary-level variables: patient demographics; income and insurance status; and clinical characteristics. Demographics included age in 5-year bins (≤ 34 , 35–39, 40–44, 45–49, 50–54, 55–59, 60–64, 65–69, 70–74, 75–79, 80–84, 85–89, 90–94, 95+), gender, race/ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, Asian, and other). Part D data have an enhanced Research Triangle Institute Race Code verified by first and last name algorithms, with much improved sensitivity ($>77\%$; Kappa coefficient 0.79).⁹

We adjusted for individual-level insurance status and a proxy for income. We used variables indicating Medicaid coverage (available to those under about 75% of the Federal poverty level (FPL), but with some state variation) and non-dual federal low-income subsidies (which vary based on FPL cut-points) to create income bins: $<75\%$ FPL, 75%–135% FPL, 135–150% FPL, and $>150\%$ FPL. We also included two indicators for supplemental drug coverage using Part D data: those with generic-coverage in the “donut hole” gap and those with both generic and brand-name drug in the gap. We also controlled for share of the year the beneficiary was enrolled. In addition, we adjusted for an indicator of being disabled (<65 years old) for beneficiaries who were eligible for Medicare because of disability. Last, we controlled for ZIP-Code level income (logarithms of the median household income within the ZIP-Code in which the individual lived) and educational attainment (share with less than high school, completing high school only, and completing part of college and above in the individual’s ZIP-Code).

Clinical characteristics included risk scores, indicators for institutionalization (defined as having 90 days in a nursing home), and death during the year. We calculated the two prospective risk scores using prior-year diagnosis and spending: CMS Hierarchical ConditionCategory scores (CMS-HCC) for non-drug medical care services and the analogous prescription drug hierarchical condition category (RxHCC) scores.¹⁰ Risk scores represent a proxy for health status, with higher scores indicating greater severity of illness and higher expected health care utilization (in our study sample, RxHCC ranges from 0 to 6.6 and CMS-HCC ranges from 0.1 to 12.4). We used prior-year instead of current-year diagnosis and spending to calculate risk score (except for the 4% of the sample who are new enrollees, for whom we use concurrent risk scores based on age and gender). Even with the use of the prior year’s risk score, physician coding may be endogenous; for example, physicians in higher-spending regions may code patients as sicker than physicians in lower-spending regions code similar patients.¹¹ We conducted sensitivity analysis to exclude risk scores (Online Supplement Tables S1–S3 and Figure S1; results are robust to the exclusion of risk scores).

Statistical analysis

We used these data to generate an adjusted average value for each outcome for each HSA (or HRR). We pooled three years (2007–2009) and conducted an individual-level linear regression for each outcome.¹² Each regression included HSA (or HRR) indicator variables, year indicators, and the adjustment variables described above. Regressions were weighted by the percent of year enrolled so those who only had partial-year enrollment would contribute less to the model. We then calculated the predicted value for each HSA (or HRR) using the estimating equation evaluated at national averages for the covariates, thus capturing variation at the HSA (or HRR) level that was purged of variation in population characteristics that we were able to control for with our observed covariates.

We used these adjusted HSA (or HRR) outcomes to perform two sets of analysis. First we described the degree of variation between HSAs (or HRRs), calculating statistics such as ratios of 75th to 25th percentiles and coefficients of variation (CV). CV is defined as the ratio of the standard deviation to the mean, and represents a normalized measure of dispersion. We included only 2908 out of 3436 HSAs with 50 enrollees or more to avoid the introduction of noise driven by small cell sizes and used the same sample to conduct HRR level analysis.

Second, we evaluated the degree to which HSAs with similar spending levels clustered together within HRRs. Clearly there will be a correlation between spending at the HSA level and spending at the HRR level – HRRs are just an aggregation of HSAs – but we assessed the degree of dispersion of HSA spending within and between HRRs. Specifically, we divided HRRs into quintiles based on their adjusted spending and also divided HSAs into quintiles based on their adjusted spending. We then tabulated the share of high- and low-spending HSAs located within the borders of high- and low-spending HRRs to gauge the variation of HSAs within and between HRRs

RESULTS

Regional variations in adjusted outcomes

The Table presents the variation in pharmacy and non-drug medical spending, counts of monthly prescriptions filled, inpatient admissions, outpatient office visits and ER visits per person per year in different regions. Panel A shows the variation at the HSA level. Beneficiaries in the median HSA filled 53 monthly prescription drugs per year, or 4.4 prescriptions per month, corresponding to \$2912 in annual gross drug spending. Medical spending and drug spending are comparably variable, with coefficients of variation of .15. The ratio of drug spending at the 75th percentile to that at the 25th percentile is 1.21, whereas the corresponding ratio for drug counts is only 1.13. This suggests that variation in the mix of drugs prescribed is larger than variation in number of drugs prescribed. Of course, part of this variation may be due to other unmeasured patient characteristics or illness severity.

Panel B reports the analogous variation across HRRs. The pattern of variation across categories is quite similar, although the overall degree of variation is somewhat lower. We explore below whether high-spending HSAs are located primarily within high spending HRRs, or whether there is substantial variation between HSAs within HRRs.

Variation across HSAs within HRRs

We gauge the degree to which HSAs with high spending are concentrated together in HRRs with high spending. A formal test of whether there is variation across HSAs nested within HRRs can (unsurprisingly) reject the null hypothesis of no systematic HSA variation within HRRs (joint $F=4.86$, $p<0.01$). About 41% of the variation in adjusted HSA drug spending is between HRRs, and 59% is within HRRs. About 43% of the variation in adjusted HSA non-drug medical spending is between HRRs, 57% within. (These results are virtually unchanged if we exclude the top and bottom 1% of the HSA values.) There is substantial variation of HSAs within HRRs. For example, Manhattan is one of the HRRs with the highest adjusted drug spending and Albuquerque is one of the lowest – but there is substantial dispersion in spending across the HSAs within those HRRs: the lowest-spending HSA in Manhattan has lower spending than about 25% of the HSAs within Albuquerque. For comparison, about 66% of the variation in HSA-level income is between HRRs, while 34% is within; about 55% of the variation in HSA-level education is between HRRs, 45% within; about 70% of the variation in HSA-level proportion of Whites is between HRRs,

30% within; about 66% of the variation in HSA-level CMS-HCC risk score is between HRRs, 34% within. Analysis at the HRR level thus masks more within-HRR local area variation in healthcare spending than it does in several (although not all) covariates.

Figure Panels A–B show the degree of clustering of high-spending HSAs within high-spending HRRs (See Online Supplement Table S2 for a complete set of conditional probabilities). For HRRs in each quintile of adjusted HRR spending, Panel A shows the share of the HSAs within that HRR that are high- or low-spending HSAs. For example, for adjusted pharmacy spending, 50.7% of the HSAs located within the highest drug-spending HRR quintile are in the highest drug-spending quintile of HSAs; 50.3% of the HSAs in the lowest drug-spending HRR quintile are in the lowest HSA quintile. Similar patterns are observed for adjusted medical spending: 57.4% of the HSAs in the highest HRR quintile are in the highest quintile of HSAs; 54.0% of the HSAs in the lowest HRR quintile are in the lowest HSA quintile.

For HSAs in each quintile of HSA spending, Panel B shows the share of the HSAs that are located in high- and low-spending HRRs. For adjusted pharmacy spending, 51.5% of the HSAs in the highest-spending quintile are located within the highest quintile HRR; 49.6% of the lowest drug-spending HSAs are located within the lowest-spending HRR quintile. For adjusted medical spending: 52.0% of the highest-spending HSAs are located within the highest-spending HRR; 46.0% of the lowest-spending HSAs are located within the lowest-spending HRR.

In sum, there is substantial misalignment of high-spending HSAs and HRRs, with many low-spending HSAs located within the borders of high-spending HRRs and many high-spending HSAs located within the borders of low-spending HRRs.

DISCUSSION

Much policy attention has been drawn to the large and persistent geographic variation in healthcare spending – for good reason. The presence of such variation (in the absence of commensurate variation in patient needs or even in health outcomes) suggests that high-intensity practice patterns in some areas signal inefficient resource use. This has led to discussion of policy levers to rein in spending in high-utilization, high-cost areas, such as lowering Medicare payments to providers in those areas. Such policy levers aim to focus on a level of aggregation that captures the local healthcare delivery system – too low a level (such as individual physician payments) could miss the system-level factors that drive some areas to be systematically high-utilization, while too high a level (incorporating multiple systems and markets) could reward or punish utilization well beyond providers' control and mask substantial heterogeneity.

Analysis has primarily been focused on variation between HRRs – areas defined based on large tertiary facilities and incorporating numerous HSAs. There are advantages to looking at such large areas (they may be large enough to capture more homogeneous patient pools), but the disadvantage of such an exclusive focus is that it can mask substantial heterogeneity at the more local level. This is particularly important when considering the effects of policy levers that aim to act on local practice patterns for primary care or avoidable hospitalizations, for example.

We examined the degree of heterogeneity within HRRs. We found that there is substantial local variation in utilization and spending for both drug and non-drug medical spending and that there is substantial dispersion of local spending within HRRs.

These findings are of course subject to several limitations. First, our analysis is based on the Medicare population. Patterns among the commercially insured may differ. Medicare does, however account for 20% of all national health care spending as of 2010,¹³ and many of the policy levers discussed apply to Medicare payment rates. Second, our risk-adjusters are imperfect and we do not capture patient preferences.¹⁴ To the extent that these vary across localities, they could drive some of the observed patterns of heterogeneity. It is somewhat reassuring on this front that patterns of unadjusted outcomes are quite similar to those with adjusted outcomes. Third, causal connections are inherently difficult to draw from ecological data. While the variation described here (and elsewhere in the literature) is strongly suggestive of inefficient use of resources, it is difficult to use these data to forecast what the effect of different policy levers might be on spending patterns.

Nevertheless these findings do have policy implications. Policies that aim to reduce the spending in high-cost areas by targeting high-spending HRRs may fail on both sensitivity and specificity: about half of the HSAs in the highest spending HRR quintile are not in the highest spending quintile of HSAs; and about half of the HSAs in the lowest spending HRR quintile are not in the lowest HSA quintile. That said, higher spending HSAs are generally in higher spending HRRs. Whether or not this degree of concordance is sufficient to achieving policy aims depends on policy-makers' tolerance for the ramifications of imperfect targeting.

This does not, however, tell us what the "right" level of aggregation for policy is. There is clearly variation in spending within HSAs – should policy focus on an even more local level? The movement towards Accountable Care Organizations (ACOs) aims to tie payments to the care delivered by provider groups that are large enough to pool risk and abstract from individual-level variation in needs and idiosyncratic outcomes, but small enough to hold the group accountable for the use of resources. In the absence of formal ACOs, payments tied to local area practice patterns aim to accomplish similar goals. This analysis suggests that policies focused exclusively on the hospital referral region may be too blunt to promote the best use of health care resources.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding source:

Institute of Medicine (HHSP22320042509XI), National Institute of Mental Health (No. RC1 MH088510) and Agency for Healthcare Research and Quality (No. R01 HS018657) to Dr. Zhang.

References

1. The Dartmouth Atlas Project. The Dartmouth Atlas of Health Care. Lebanon, NH: 2010.
2. 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; 138:288–98. [PubMed: 12585826]
3. 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; 138:273–87. [PubMed: 12585825]
4. Zhang Y, Baicker K, Newhouse JP. Geographic variation in Medicare drug spending. *N Engl J Med.* 2010; 363:405–9. [PubMed: 20538621]

5. Zhang Y, Baicker K, Newhouse JP. Geographic variation in the quality of prescribing. *N Engl J Med*. 2010; 363:1985–8. [PubMed: 21047217]
6. Donohue JM, Morden NE, Gellad WF, et al. Sources of Regional Variation in Medicare Part D Drug Spending. *N Engl J Med*. 2012; 366:530–8. [PubMed: 22316446]
7. Report to the Congress: Regional Variation in Medicare Service Use. Medicare Payment Advisory Commission; 2011. at http://www.medpac.gov/documents/Jan11_RegionalVariation_report.pdf
8. Report to the Congress: Measuring Regional Variation in Service Use. Medicare Payment Advisory Commission; 2009. at http://www.medpac.gov/documents/Dec09_RegionalVariation_report.pdf
9. Final Report. Sub-Task 2. (Prepared by RTI International for the Centers for Medicare and Medicaid Services through an interagency agreement with the Agency for Healthcare Research and Policy, under Contract No. 500-00-0024, Task No. 21). Agency for Healthcare Research and Quality; 2008. Creation of New Race-Ethnicity Codes and Socioeconomic Status (SES) Indicators for Medicare Beneficiaries. at <http://www.ahrq.gov/qual/medicareindicators/medicareindicators.pdf>
10. 2011 RxHCC Model Software. Centers for Medicare & Medicaid Services; 2011. at https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk_adjustment.html
11. Song Y, Skinner J, Bynum J, Sutherland J, Wennberg JE, Fisher ES. Regional Variations in Diagnostic Practices. *N Engl J Med*. 2010; 363:45–53. [PubMed: 20463332]
12. Buntin MB, Zaslavsky AM. Too much ado about two-part models and transformation? Comparing methods of modeling Medicare expenditures. *J Health Econ*. 2004; 23:525–42. [PubMed: 15120469]
13. Martin AB, Lassman D, Washington B, Catlin A. Team tNHEA. Growth In US Health Spending Remained Slow In 2010; Health Share Of Gross Domestic Product Was Unchanged From 2009. *Health Aff (Millwood)*. 2012; 31:208–19. [PubMed: 22232112]
14. Doyle JJ. Returns to local-area health care spending: using health shocks to patients far from home. *American Economic Journal: Applied Economics*. 2011; 3:221–43.

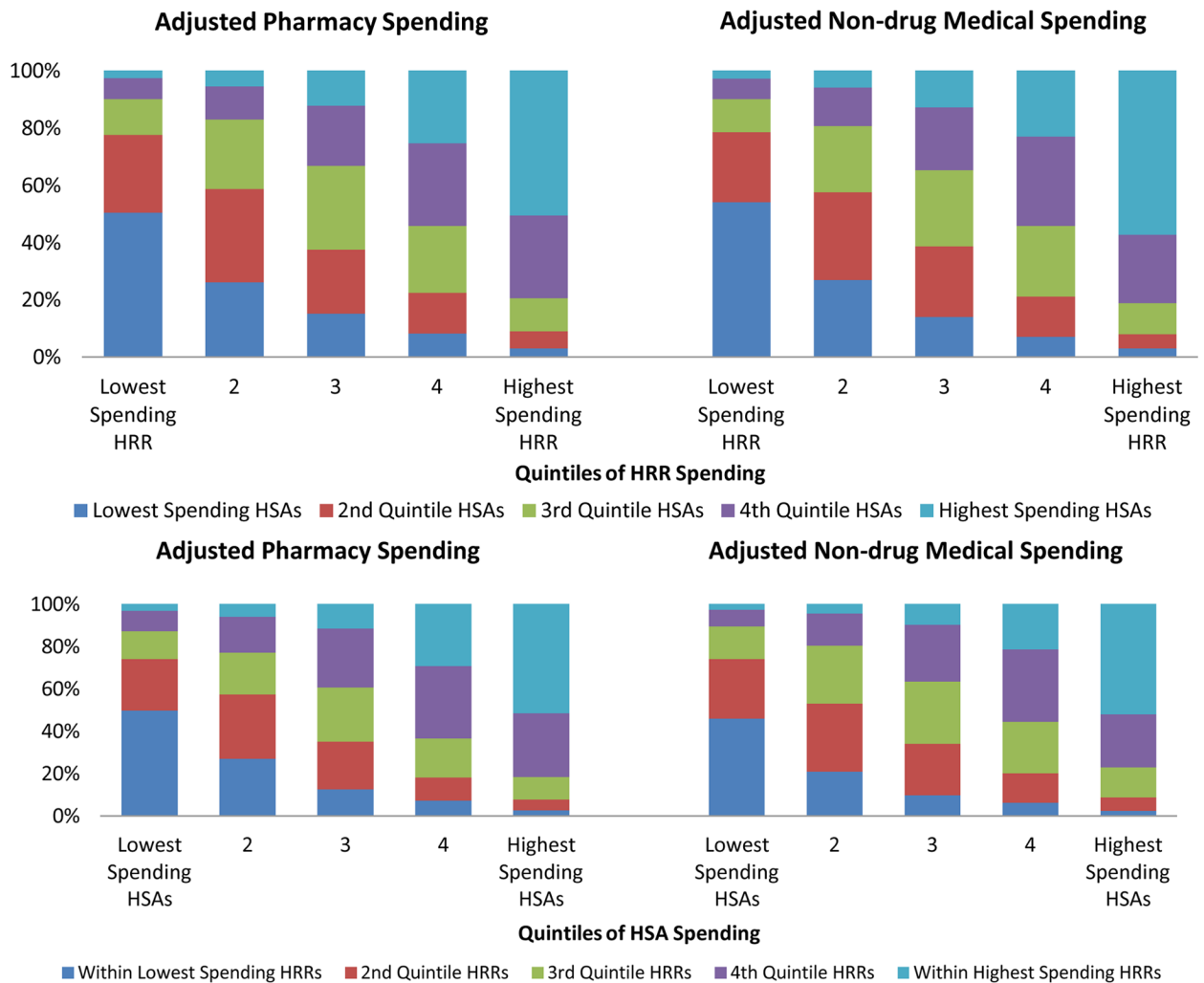


Figure.
 Quintile Analysis for Adjusted Pharmacy Spending and Non-drug Medical Spending Between Hospital-referral Regions and Hospital-service Areas
 Panel A. The Type of HSAs Located within High and Low-Spending HRRs
 For each quintile of HRR based on HRR spending, this chart shows what shares of the HSAs located within that type of HRR are high or low spending HSAs. For example, 50.7% of the HSAs in the highest drug-spending HRR quintile are in the highest drug-spending quintile of HSAs; 50.3% of the HSAs in the lowest drug-spending HRR quintile are in the lowest HSA quintile.
 Panel B. The Type of HRRs where High and Low Spending HSAs are Located
 For each quintile of HSA based on HSA spending, this chart shows what share of that type of HSA are located in high and low spending HRRs. For example, for adjusted pharmacy spending, 51.5% of the highest spending quintile HSAs are in the highest-spending quintile HRR; 49.6% of the lowest drug-spending HSAs are in the lowest drug-spending HRR quintile.

\$watermark-text

\$watermark-text

\$watermark-text

Table 1

Variation in Adjusted Medicare Drug and Non-Drug Medical Spending, Counts of Monthly Prescriptions Filled, Inpatient Admissions, Outpatient Office Visits, and Emergency Room Visits, All in Per-Person-Per-Year Units in Different Regions (among Hospital-Service Areas with Greater Than 50 Beneficiaries)

Panel A. Hospital-Service Areas (2908 out of 3436 HSAs)											
	Min	10 th	25 th	Median	75 th	90 th	Max	Mean	S.D.	75 th /25 th	CV
Drug spending	1001	2398	2645	2912	3198	3445	5261	2931	433	1.21	0.15
Drug counts	28.26	46.86	49.48	52.60	55.69	58.78	75.39	52.72	4.89	1.13	0.09
Medical spending	6487	9858	10791	11812	12964	14144	23888	11961	1798	1.20	0.15
IP counts	0.15	0.34	0.38	0.43	0.49	0.55	0.88	0.44	0.09	1.28	0.20
OP counts	2.02	5.25	5.92	6.59	7.30	8.03	12.86	6.62	1.18	1.23	0.18
ER counts	0.00	0.55	0.62	0.71	0.80	0.91	2.40	0.72	0.16	1.28	0.22
Total spending	8871	12614	13601	14767	15988	17245	28178	14892	1931	1.18	0.13

Panel B. Hospital Referral Regions (306 out of 306 HRRs)											
	Min	10 th	25 th	Median	75 th	90 th	Max	Mean	S.D.	75 th /25 th	CV
Drug spending	2347	2665	2787	2952	3136	3275	4211	2972	268	1.13	0.09
Drug counts	44.37	48.12	50.45	52.42	54.18	55.99	64.04	52.32	3.05	1.07	0.06
Medical spending	8123	10403	11081	11766	12624	13580	19478	11912	1369	1.14	0.11
IP counts	0.30	0.36	0.39	0.43	0.47	0.50	0.61	0.43	0.05	1.20	0.12
OP counts	3.64	5.62	6.11	6.62	7.10	7.78	10.74	6.66	0.91	1.16	0.14
ER Counts	0.35	0.60	0.65	0.70	0.75	0.80	1.01	0.70	0.08	1.16	0.12
Total spending	10656	13255	13849	14782	15680	16586	23689	14884	1477	1.13	0.10

Total drug spending includes Part D plan payment before rebates, beneficiary out-of-pocket spending, and subsidy amount. Drug counts are number of monthly prescription drugs (≤ 1 if days supply ≤ 30 ; ≥ 1 if days supply > 30). Total medical spending included Medicare and beneficiary payment for all medical services (including inpatient, outpatient, physician, home health, hospice, skilled nursing home, and medical devices) and was adjusted for local price-level differences using county-level factor prices given to us by the Medicare Payment Advisory Commission. We did not adjust drug spending for regional price differences because the variation in drug prices among regions was negligible. To account for the difference in population mix across regions, we adjusted for three major categories of beneficiary-level variables: patient demographics, insurance status, and clinical characteristics in the individual-level regressions for each outcome. All outcomes in the table are after adjustment.

Abbreviations: IP = inpatient, OP = outpatient, ER = emergency room, CV = coefficient of variation, which is defined as the ratio of the standard deviation to the mean, and represents a normalized measure of dispersion.