Variation in Quality and Costs of Care Across Physicians and Its Determinants

Variation in Quality and Costs of Care across Physicians and Its Determinants

A dissertation presented

by

Yusuke Tsugawa

to

The Committee on Higher Degrees in Health Policy

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Health Policy

Harvard University

Cambridge, Massachusetts

May 2016

Dissertation Advisors: Ashish Jha and Anupam Jena Yusuke Tsugawa

Variation in Quality and Costs of Care across Physicians and Its Determinants

ABSTRACT

This dissertation evaluates one of the key determinants of health care quality and costs – practice patterns of physicians. For decades, rapid health care spending growth and suboptimal quality of care have been fundamental issues of the U.S. health care system. A large body of literature has demonstrated substantial geographic variation in health care utilization without concomitant improvement in patient outcomes in higher spending regions. This literature has spurred debate about whether current levels of health care utilization are indeed socially wasteful – i.e., generate higher spending with no improvement in patient outcomes. While informative, however, this literature has not investigated variation due to individual physician behaviors, despite the central role of physicians as key decision makers in health care. In fact, surprisingly little is known about how individual physicians vary in their care, the determinants of that care, and the implications of that variation for patient outcomes.

This dissertation attempts to shed light on these questions by analyzing the productivity of health care spending at the physician level (Figure).

Chapter 1 investigates the proportion of service use variation that can be explained by variation in individual physician practice patterns, and examines the impact of that variation on patient outcomes. I analyze data on Medicare beneficiaries hospitalized with medical conditions treated by general internists. Using a cross-classified multilevel model, I find that variation in spending across physicians exceeds variation across hospitals (10.9% and 6.2% of overall spending, respectively). As for evaluating the impact of between-physician variation in spending on patient outcomes, I exploit a natural experiment of physicians who specialize in hospital-based care – hospitalist physicians. Hospitalists routinely work in shifts and therefore patients are plausibly quasi-randomized to these physicians within a hospital based on physician work schedule. Among 272,979 hospitalizations treated by 8,489 hospitalists, hospitalists in the highest-spending quartile had lower 30-day patient mortality than hospitalists in the lowest quartile within the same hospital, despite similar patient characteristics (adjusted mortality rate 10.7% vs. 11.2%; adjusted odds ratio 0.94, 95%CI: 0.90 to 0.98, p=0.002). I observe no relationship between physician spending and patients' readmission rates. Given larger variation in spending across physicians than across hospitals, our findings suggest that policies focused on individual physicians may be as or more effective than those targeted toward hospitals or regions.

Moreover, interventions targeted at high-spending physicians to reduce spending, without accounting for their quality of care, may have the unintended consequence of negatively impacting patients' health.

Chapter 2 begins my evaluation of the upstream determinants of variation in quality of care across physicians with a special focus on physicians' years in practice. Physicians with longer years in practice may accumulate, or conversely exhibit obsolescence of, knowledge and skills. However, the association between physicians' years in practice and patient outcomes is poorly understood. Using data on Medicare beneficiaries aged 65 years or older hospitalized during 2011-2013 with a medical condition, I investigate the association between hospitalist physicians' years since residency completion and patient outcomes, adjusting for patient and physician characteristics and hospital-specific fixed effects. I again rely on quasi-randomization of patients to hospitalists to circumvent the possibility that physicians with greater years in practice may treat patients that are sicker on unobserved dimensions. Of 386,159 hospitalizations treated by 14,650 hospitalists, hospitalists in practice longer had higher patient mortality than hospitalists in practice for fewer years, despite similar patient characteristics. Each additional 10 years in practice was associated with 0.5% increase (95% CI: 0.3% to 0.7%, $p<0.001$) in patients' mortality. Significant effects were present for low- and medium-volume physicians, but not high-volume physicians. Readmissions and costs of care were not meaningfully associated with physician years in practice. This study has implications for recent debates in the medical community regarding how best to ensure maintenance of clinical skills over a physician's career. Our findings suggest that evaluating patient outcomes, particularly among older physicians with low-to-medium patient volume, may be necessary to guarantee that quality care provided by physicians is high throughout their careers.

Chapter 3 assesses another upstream determinant of between-physician variation in quality of care – physician sex. Studies have found differences in practice patterns between male and female physicians, with female physicians more likely to adhere to clinical guidelines and evidence-based practice. However, whether patient outcomes differ between male and female physicians is largely unknown. While physician sex is not a modifiable factor, understanding whether quality of care differs between male and female physicians is critically important, as it allows us to further investigate which aspects of practice patterns that vary between male and female physicians lead to better patient outcomes. Using nationally representative data on Medicare beneficiaries in 2012-2013, I examine the association between physician sex and patient outcomes among general internists. Despite similar observed illness severity of patients, female physicians have lower 30-day patient mortality (adjusted mortality rate 10.9% vs 11.4%; adjusted risk difference -0.5%, 95%CI: -0.7% to -0.4%, $p<0.001$) and lower 30-day readmissions (adjusted readmission rate 15.1% vs 15.8%; adjusted risk difference -0.7%, 95%CI: -0.8% to -0.5, $p<0.001$) within same hospital. These findings are unaffected when restricting analyses to hospitalist physicians for whom patients are plausibly randomized. Although the exact mechanism underlying these differences remains unclear, understanding why these differences in care quality exist, and what we might do to alleviate them, is critical to ensuring that all patients get high quality care.

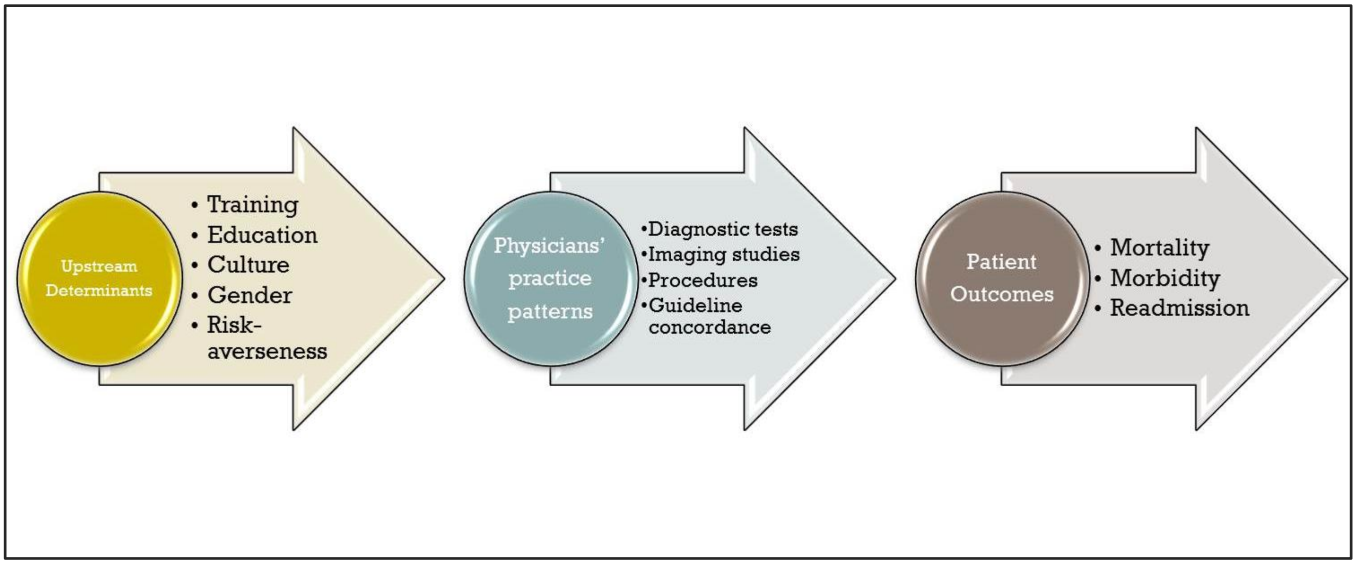

Figure | Conceptual framework. Variation in quality and costs of care across physicians and its determinants.

Table of Contents

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my dissertation committee. I have had the distinct pleasure and honor to be mentored by Ashish Jha, Anupam Bapu Jena, Joseph Newhouse, and Alan Zaslavsky. This dissertation could not have been completed without dedicated support that I have received from them over the years.

Ashish Jha has been a great mentor, a role model, and a friend. Throughout my time as a graduate student, he has given me excellent guidance, and taught me how to address important health policy questions through rigorous, well-timed research. Ashish taught me to be ambitious and creative about tackling big questions that interest me most, by decomposing a big question into smaller analyzable pieces. Ashish always has been supportive and caring, and took it serious about advancing my career to become an independent health policy researcher.

I owe a great debt of gratitude to Anupam Bapu Jena for teaching me how to think "outside the box" and develop novel research ideas. Bapu taught me how to understand health care through the eyes of health economist and physician. I immensely appreciated his generosity throughout my time as a student. Bapu and I had weekly meetings, even when I was in Japan, through which I learned a lot about how to conduct high-quality research. It has been a joy to get to know Bapu as a mentor, collaborator, and friend.

I would like to express my special appreciation and thanks to Joseph Newhouse, who has been extraordinarily supportive and caring throughout my time at Harvard. Joe taught me to be accurate and rigorous about numbers. I worked with Joe closely when I was working on standardized costs in Medicare data. Joe has always been patient and supported me cheerfully. Joe's insights and suggestions improved every part of this dissertation, and it has been an incredible honor to serve as his student.

I am deeply indebted to Alan Zaslavsky, who has been my academic advisor ever since I entered the Ph.D. program. Alan taught me how best to disentangle complex statistical questions into a series of surprisingly simple yet comprehensive components. I always enjoyed discussing about statistics with Alan. Alan challenged and pushed the boundaries of my understanding about statistics, for which I am extremely grateful.

I thank my collaborators who are not named above, Jose F. Figueroa and E. John Orav. Jose has been a great friend, and I have always enjoyed discussing random, yet relevant, research ideas with him. My friends and colleagues at 42 Church Street (Laura Burke, Maggie Lamond, Liana Woskie, Jie Zheng, Xiner Zhou, Garret Johnson, Zoe Lyon, Anthony Moccia, Irene Papanicolas, Winta Mehtsun, Igna Bonfrer, and Haley Abing) have all extended their support, and provided an intellectually exciting environment to do research.

Debbie Whitney and Jessica Livingston at Harvard Ph.D. program in Health Policy contributed immensely to my personal and professional time at Harvard. For me, their support extends far beyond professional administrative support.

I gratefully acknowledge the funding sources that made my Ph.D. work possible. I was funded by the World Bank and Honjo International Scholarship Foundation. I am also very thankful to my mentor Fukui Tsuguya, and my colleagues at St. Luke's International University in Japan, for being extremely supportive throughout my Ph.D. program.

Finally, I wish to express my unqualified thanks to my wife, Erina Tsugawa. She has always stood by me cheerfully through the good times and bad times. Without her love and support, I would not have been able to accomplish my work. I am also immensely grateful to my father Motomu Tsugawa, and my mother Seiko Tsugawa, who raised me and taught me the importance of diligence and perseverance, and to give priority in my life to the quest for knowledge.

Chapter 1: Variation in Physician Spending and Association with Patient Outcomes[1]

[1] A version of this chapter is submitted to a journal for publication. Please do not circulate.

1.1 INTRODUCTION

The substantial variation in spending and health care services across U.S. regions as well as hospitals within the same community is well known. (Wennberg and Gittelsohn 1973, Barnato, Herndon et al. 2007, Sutherland, Fisher et al. 2009, Zhang, Wright et al. 2009, Zuckerman, Waidmann et al. 2010, Zhang, Baik et al. 2012, Institute of Medicine 2013, Newhouse and Garber 2013, Newhouse and Garber 2013) Findings on the clinical implications of this variation are mixed, but most suggest that greater health care spending does not translate into better patient outcomes (Table 1.1). (Fisher, Wennberg et al. 2003, Fisher, Wennberg et al. 2003, Jha, Orav et al. 2009, Yasaitis, Fisher et al. 2009, Chen, Jha et al. 2010, Lagu, Rothberg et al. 2011, Romley, Jena et al. 2011, Stukel, Fisher et al. 2012, Hussey, Wertheimer et al. 2013) However, decisions on utilization are ultimately made by front-line clinicians, and the degree to which their individual utilization decisions vary, and the implications of that variation, are not well understood.

Table 1.1 Major studies linking spending level and patient outcomes

Level	Authors (year)	Data	Conditions	Treatment	Outcome	Main findings
Region	Wennberg and Gittelsohn (1973)	HSAs (n=13) in Vermont state				A wide variation in utilization of care across HSAs
	Fisher et al. (2003)	HRR (n=306) in Medicare Current Beneficiary Survey (MCBS)	Hip fracture, colorectal cancer, or AMI	End-of-life spending	5-year mortality rate, change in functional status, and satisfaction	Higher spending regions had slightly higher mortality, and no differences in functional status and satisfaction
	Baicker and Chandra (2004)	State level	6 common medical conditions	Adjusted Medicare reimbursement per capita	Process measures	States with higher Medicare spending have lower-quality care
Hospital	Jha et al. (2009)	Medicare data	AMI, CHF, PNA	Hospital risk-adjusted costs	30-day mortality rate	No association between hospital spending and patient mortality
	Chen et al. (2010)	Medicare data	CHF, PNA	Hospital risk-adjusted costs	30-day mortality rate	Higher-spending hospitals had lower mortality for CHF, and higher mortality for PNA
	Romley, Jena, and Goldman (2011)	California hospitals	6 major medical conditions	End-of-life hospital spending	Inpatient mortality	High-spending hospitals had lower inpatient mortality
	Stukel et al. (2012)	Hospitals in Ontario, Canada	AMI, CHF, hip fracture, or colon cancer	End-of-life hospital spending	30-day and 1-year mortality, readmissions, major cardiac events	Higher spending intensity was associated with lower mortality, readmissions, and cardiac event rates

HAS denotes hospital service area, HRR denotes hospital referral regions, AMI denotes acute myocardial infarction, CHF denotes congestive heart failure, and PNA denotes pneumonia.

While regional culture and hospital capabilities obviously influence a patient's treatment, individual physicians are critical contributors to the observed variation in health care services. They make key decisions about admitting and discharging patients, ordering imaging and diagnostic tests, engaging consultants, and referring for procedures. Despite the central importance of physicians in medical decision making, we know surprisingly little about how much utilization behavior varies across physicians. Even more importantly, we have almost no data on whether variation in spending at the individual physician level matters for patient outcomes. While a few studies have profiled physician spending for specific conditions (Adams, McGlynn et al. 2010, Adams, Mehrotra et al. 2010, Mehrotra, Adams et al. 2010, Mehrotra, Adams et al. 2010), they have neither analyzed nationally representative data nor examined its association with patient outcomes.

To address this gap, we used a national sample of Medicare beneficiaries hospitalized with a medical condition in 2011-2013 to: (1) compare variation in spending across individual internists (both hospitalists and non-hospitalist internists) to variation in spending across hospitals; (2) investigate the characteristics of physicians associated with higher spending; and (3) analyze the association between physician spending and 30-day patient mortality and readmission rates. In examining the relationship between physician spending and patient outcomes, we focused on hospitalist physicians, hypothesizing that within the same hospital patients are plausibly quasi-randomized to a given hospitalist based on that physician's particular work schedule.

1.2 METHODS

Data

We used the Medicare 20% Carrier and Inpatient Files to identify patients hospitalized in acute care hospitals between January 1, 2011 and November 30, 2013 among fee-for-service (FFS) beneficiaries age 65 years or older with a medical diagnosis related group (DRG). We excluded patients admitted in December 2013 from mortality analyses, and those discharged in December 2013 from readmission analyses to allow sufficient follow-up periods. For readmission analyses, we excluded patients who were out of hospital for less than 30 days prior to the index admissions to make sure their admissions were not related to prior episodes. Finally, we excluded elective hospitalizations to ensure that spending related to the reason for hospitalization was unlikely to occur in the outpatient setting prior to hospitalization, e.g., imaging in advance of planned inpatient chemotherapy administration, and hospitalizations in which a patient left against medical advice.

We assigned a physician to each hospitalization based on the physician's National Provider Identifier (NPI) in the Carrier File who accounted for the most Part B spending (visits, tests, procedures) during that hospitalization. (McWilliams, Landon et al. 2014) On average, these physicians were responsible for 52.1% of total Part B spending for a given hospitalization. We restricted our analysis to hospitalizations in which the assigned physician was a general internist in order to avoid comparing behavior of physicians in different specialties. We extracted physician characteristics (specialty, age, sex, and allopathic versus osteopathic training) from a comprehensive physician database obtained from Doximity, a company that provides online professional networking services for U.S. physicians. The database includes information on majority of U.S. physicians, obtained from multiple sources and data partnerships including the National Plan and Provider Enumeration System, U.S. Department of Health and Human Services NPI Registry, state medical boards, and specialty boards (e.g., the American Board of Medical Specialties and the American Board of Surgery). (Jena, Khullar et al. 2015, Olmsted, Geisen et al. 2015) The database contained information on 93% of all physicians treating at least one Medicare beneficiary in the 2011 Medicare Inpatient file. Details of the database and its validation have been published elsewhere. (Jena, Khullar et al. 2015, Olmsted, Geisen et al. 2015)

Physician-level spending

To measure a physician's health care spending level, we used standardized Part B spending, adjusted for patient characteristics and hospital-specific fixed effects. We focused on Part B spending for several reasons. First, it encompasses professional and other services at the discretion of physicians, e.g., consultations, radiologist interpretation of diagnostic imaging, and procedures (some spending that the physician controls falls under Part A [e.g., medication use] and is omitted from our calculations). Second, conditional on admission, Part A spending is largely invariant to physician decisions on an individual patient's treatment because of the fixed DRG payment (only 1.0% of hospitalizations in our data received an outlier payment). Third, Part B spending is a proxy for the intensity of resource use of physicians, and thus likely to correlate with resource use in other components.

To calculate standardized spending per hospitalization for each physician, we first standardized raw spending using methods established by the Center for Medicare and Medicaid Services. (Centers for Medicare & Medicaid Services 2013) Doing so accounted for variation in factor prices and malpractice premiums across areas and thus effectively yielded a measure of aggregate health care spending. We then risk-adjusted spending by regressing standardized spending on patient characteristics (described below) and hospital-specific fixed effects, using ordinary least squares.

$$
\begin{aligned}
PartBSpending_i \\
&= \beta_0 + \beta_1 Age_i + \beta_2 Female_i + \beta_3 Race_i \\
&+ \beta_4 MedicaidStatus_i +_i \beta_5 MSDRG_i + \beta_6 ElixhauserIndex_i \\
&+ \beta_7 HospitalIndicator_i + \varepsilon_i
\end{aligned}
$$

Where $PartBSpending_i$ denotes the total Part B service use for a given hospitalization, $Age_i$is age as categorical variable (as shown above), $Race_i$ is race as categorical variable, $MSDRG_i$ represents dummy-coded DRGs, $ElixhauserIndex_i$is dummy-coded Elixhauser comorbidity index (i.e., 27 coexisting conditions), and $HospitalIndicator_i$ is hospital indicators.

Diagnostics confirmed good fit of the model including linearity, homoscedasticity, and normality. Hospital-specific fixed effects account for differences between hospitals in mean costs of care, including those arising from unobserved differences in patient populations.

We addressed outliers by excluding hospitalizations in the bottom and top 5 percent of residuals.

We then refit the regression model and estimated expected spending for each hospitalization. After aggregating expected and observed spending over all a physician's hospitalizations, we calculated an observed-to-expected ratio for each physician. When a given physician is responsible for the care of the patient in hospitalization *i = 1, 2, 3, ….., j*, the observed-to-expected ratio for that physician was calculated as following:

$$Observed\ to\ expected\ cost\ ratio = \frac{\sum_{i=1}^{j} Observed\ total\ Part\ B\ spending_i}{\sum_{i=1}^{j} Expected\ total\ Part\ B\ spending_i}$$

Multiplying these ratios by the grand mean of spending per hospitalization yielded a standardized spending level which we classified into quartiles to allow for a non-linear relationship between physician-level spending and patient outcomes. Similar methods have been used to estimate physician-level measures in other domains. (Hannan, Kilburn et al. 1990, O'Connor, Plume et al. 1991, Grover, Johnson et al. 1994, Hannan, Kilburn et al. 1994, O'Connor, Plume et al. 1996, Ferguson, Peterson et al. 2003, Tsai, Joynt et al. 2013)

To avoid unstable estimates of spending, we restricted our sample to physicians with at least 10 observed hospitalizations in 2011. Because our data were based on a 20% random sample of

Medicare beneficiaries, and as Medicare FFS beneficiaries account for approximately 35% of U.S. hospitalizations (Argency for Healthcare Research and Quality), our analysis implicitly reflects physicians treating about 71 or more inpatients per year. We restricted our sample to patients with both admission and discharge dates during 2011, and excluded hospitalizations in which a patient was transferred to another hospital, electively admitted, or left against medical advice.

Adjustment variables

Adjustment variables included patient characteristics, physician characteristics, and hospital-specific fixed effects. Patient characteristics included age in 5-year increments (from 65-69, 70-74, and so on through 90-94, and ≥ 95), sex, race or ethnic group (non-Hispanic white, non-Hispanic black, Hispanic, and other), MS-DRG, Elixhauser comorbidity index (27 coexisting conditions), (Elixhauser, Steiner et al. 1998) and an indicator for dual Medicaid coverage. Physician characteristics included age in 5-year increments (from <35, 35-39, and so on through 65-69, and ≥70), sex, and allopathic (MD) versus osteopathic (DO) training.

Statistical analysis

First, using data on hospitalizations treated by general internists (including both hospitalists and non-hospitalist internists), we analyzed how variation in spending across individual physicians compared with variation across hospitals. To do so, we fit a cross-classified multilevel model and partitioned the variance in spending into hospital-, physician-, and patient-level variation. (Diggle, Heagerty et al. 2002) We reported variance partition coefficients (VPCs), which describe the proportion of the variance attributable to each level of the model hierarchy (Goldstein, Browne et al. 2002), as well as inter-unit reliability (IUR), which assesses the degree to which the observed variation among physicians is due to true differences among physicians rather than random noise (Zaslavsky 2001).

$$IUR = \frac{\sigma^2_{Between\ physicians}}{\sigma^2_{Between\ physicians} + \left(\frac{\sigma^2_{Between\ patients}}{n}\right)}$$

Second, we estimated the association between average physician spending and 30-day mortality using multivariable logistic regression models adjusting for patient and physician characteristics, as well as hospital-specific fixed effects.

$$logit(E[Outcome_{ij}]) = \beta_0 + \beta_1 PhysicianSpending_j + \beta_2 Age_i + \beta_3 Female_i + \beta_4 Race_i + \beta_5 MedicaidStatus_i + \beta_6 MSDRG_i + \beta_7 ElixhauserIndex_i + \beta_8 HospitalIndicator_i$$

Where $Outcome_{ij}$ is either mortality or readmission (binary variables) for patient i treated by physician j in 2012-2013. $PhysicianSpending_{ij}$ denotes physicians' spending quantiles defined using 2011 data.

Given that unobservable patient characteristics that influence outcomes may vary systematically across physicians with different spending levels even within the same hospital, we specifically focused on hospitalist physicians, hypothesizing that within the same hospital patients are plausibly quasi-randomized to hospitalists based on work schedules. We defined hospitalists using a validated approach; general internists with at least 20 evaluation-and-management (E&M) billings in a given year who filed at least 90% of their total E&M billings in inpatient setting as defined by Current Procedural Terminology [CPT] codes 99221-99223, 99231-99233, and 99251-99255). (Kuo, Sharma et al. 2009) Our data showed that 56% of all Medicare hospitalizations treated by general internists were cared for by hospitalists in 2011. As a sensitivity analysis, we assessed the impact of defining hospitalists as general internists filing at least 80% of their total E&M billings in the inpatient setting. We assessed the validity of our quasi-randomization assumption by testing the balance of patient characteristics across quartiles of hospitalist physician spending.

Because our regression models included hospital-specific fixed effects, we effectively compared hospitalist physicians with different spending levels within the same hospital, thereby obviating concerns that higher-spending physicians may tend to practice in hospitals with both greater unobserved illness severity and costs. Furthermore, to ensure that patient characteristics do not directly affect a physician's spending level, we classified physicians into spending levels using 2011 data and examined their patients' outcomes in 2012-2013.

We calculated risk-adjusted patient mortality and readmission rates for each quartile of physician spending using predicted probabilities of outcomes for each hospitalization with the distribution of covariates in the national sample, fixing physician spending quartile at a specific level (known as the marginal standardization form of predictive margins). (Williams 2012, Zhang, Baik et al. 2012) To overcome failure of the likelihood maximization algorithm to converge for logistic regression models, we combined DRG codes that had no death or readmission into clinically similar categories. (Allison 2008) Standard errors were clustered at the physician level.

Sensitivity analyses

We conducted several sensitivity analyses. First, we evaluated two alternative approaches to addressing outlier cases: (1) trimming the bottom and top 3 percent of residuals, and (2) trimming the bottom and top 7 percent of residuals. Second, to address the possibility that physicians who consistently transfer their sickest patients may be misclassified as low spending physicians, we included all patients who were transferred to other acute care hospitals and allocated all subsequent spending to the initial hospitalization. Third, it is possible that the association between physician spending and mortality could be confounded by unobserved care preferences of patients, such as Do-not-resuscitate (DNR) directives. For example, if certain physicians consistently treat patients with DNR directives, and patients with DNR directives have both lower costs and higher risk of mortality, those physicians may have lower spending and higher mortality. Although our assumed quasi-randomization of patients across hospitalist physicians should address this, we nonetheless conducted an additional analysis that excluded patients with cancer, assuming that rates of DNR directives would be disproportionately higher among these patients. Lastly, to assess the generalizability of our findings, we repeated the analyses among general internists overall (i.e., hospitalists and non-hospitalist internists) as opposed to hospitalists alone. This study was approved by the institutional review board at

Harvard Medical School.

1.3 RESULTS

Variation in spending across physicians

We examined 434,616 hospitalizations treated by 43,632 physicians at 3,145 acute care hospitals for the analysis of general internists (hospitalists and non-hospitalist internists); and 248,283 hospitalizations treated by 20,542 physicians at 2,751 acute care hospitals for the analysis of hospitalists. We observed wide variation in standardized and adjusted average Part B spending per hospitalization across individual physicians (i.e., after adjusting for hospital effects) (Figure 1.1). This variation in spending across physicians was larger in magnitude than variation across hospitals (10.9% across general internists vs 6.2% across hospitals; 8.2% across hospitalist physicians vs 7.9% across hospitals) (Table 1.2). Using an average number of hospitalizations per general internist of 31, the IUR was 0.80, indicating that the measured variation between physicians was mostly due to true differences rather than random noise. (Zaslavsky 2001)

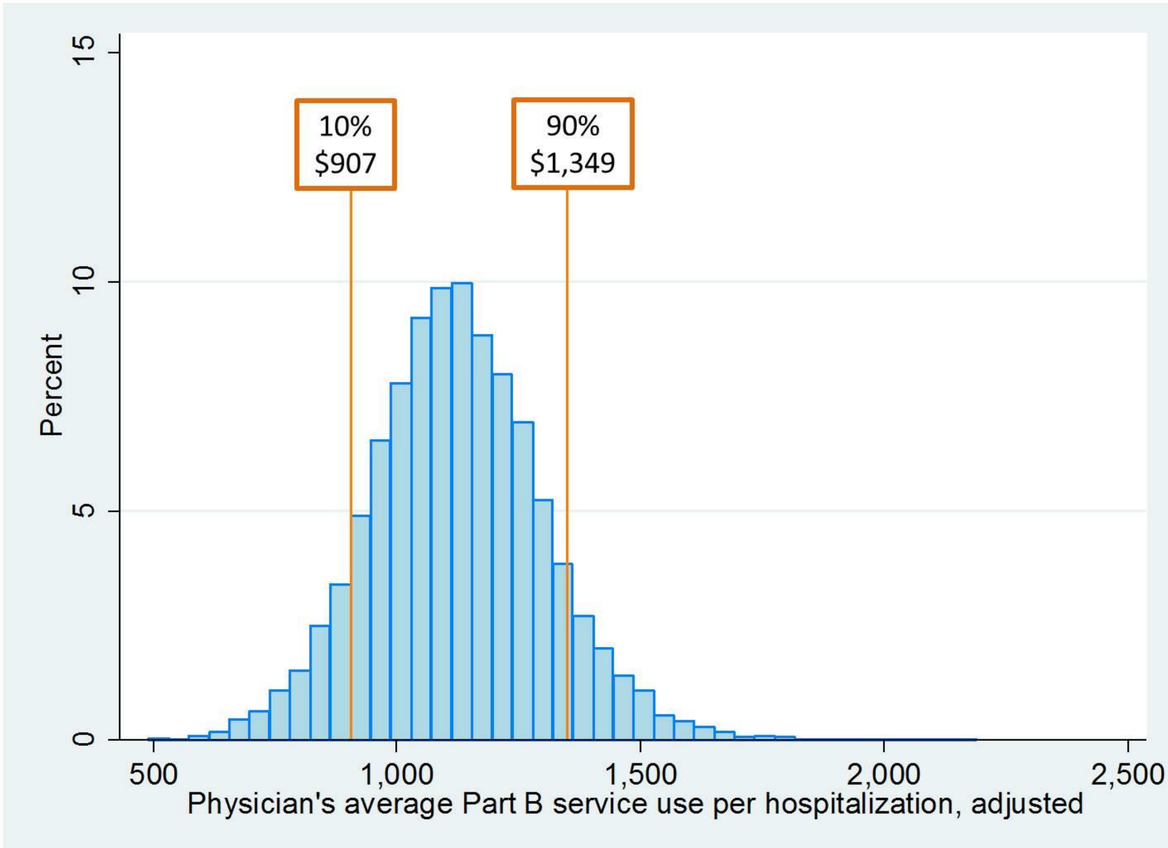

Figure 1.1 | Distribution of physician's average Part B spending (adjusted) per hospitalization. Vertical lines show 10th and 90th percentile of the distribution. Data are from a sample representative of general internists treating 10 or more hospitalizations in 2011.

Table 1.2 Proportion of variation in adjusted Medicare spending explained by hospitals, physicians, and patients

	Total Part B spending	Evaluation & management spending	Imaging & tests spending
General internists			
Variance partition coefficients (VPCs)			
Between hospitals	6.2%	6.6%	5.7%
Between physicians	10.9%	11.5%	4.4%
Between patients	82.9%	81.8%	89.9%
Inter-unit reliability (IUR)	0.80	0.81	0.60
Hospitalist physicians			
Variance partition coefficients (VPCs)			
Between hospitals	7.9%	8.5%	6.7%
Between physicians	8.2%	8.7%	3.2%
Between patients	83.9%	82.8%	90.1%
Inter-unit reliability (IUR)	0.76	0.77	0.53

We classified physician spending into Berenson-Eggers Type of Service (BETOS) categories of evaluation-and-management, and imaging and tests (e.g., lab tests, electrocardiograms). Although we included spending on procedures in our total Part B spending measures, we did not investigate variation in procedure spending as a distinct category because approximately 87% of hospitalizations had no procedure spending. See text for method of standardization.

Physician characteristics related to high spending

Among hospitalist physicians, total spending was approximately 40% higher among physicians in the highest spending quartile compared to the lowest spending quartile ($1,231 vs $863 per hospitalization) (Table 1.3). Higher spending physicians were more likely to be older and to be female. For example, the mean age of physicians in the highest adjusted spending quartile was 44.3 years versus 43.4 years in the lowest spending quartile (p=0.001 for difference). Allopathic or osteopathic training were not associated with physician spending.

Table 1.3 Physician and patient characteristics across quartiles of physicians' adjusted spending, among hospitalist physicians

	Quartile of physician spending, adjusted				
	Q1 (Lowest)	**Q2**	**Q3**	**Q4 (Highest)**	**p-value**
Physician characteristics					
Adjusted Prat B spending ($)	$863	$1,001	$1,086	$1,231	
Number of patients per physician (n)	16.8 (7.2)	20.2 (11.8)	21.3 (11.5)	18.4 (9.2)	<0.001
Age, yrs (sd)	43.4 (7.9)	43.7 (7.9)	44.0 (8.2)	44.3 (8.4)	0.001
Female (%)	29.6%	28.6%	29.9%	32.7%	0.03
Credentials (%)					
MD (allopathic)	91.5%	91.1%	92.4%	92.9%	0.13
DO (osteopathic)	8.5%	8.9%	7.7%	7.2%	
Patient characteristics					
Number of patients (n)	61,036	71,820	75,546	66,858	
Age (yr)	80.2 (8.5)	80.2 (8.5)	80.2 (8.5)	80.3 (8.5)	0.08
Female (%)	60.9%	60.5%	60.8%	60.8%	0.58
Race (%)					
White	83.9%	84.7%	84.4%	82.8%	<0.001
Black	9.5%	8.6%	8.8%	9.7%	
Hispanic	4.0%	4.1%	4.3%	4.7%	
Other races	2.6%	2.6%	2.5%	2.8%	
Medicaid status (%)	22.9%	22.6%	23.1%	23.9%	<0.001
Coexisting condition (%)					
Congestive heart failure	19.4%	19.5%	19.7%	20.1%	0.01
Chronic pulmonary disease	24.7%	25.4%	25.3%	25.5%	0.003
Diabetes	31.3%	31.4%	31.4%	31.3%	0.95
Renal failure	22.0%	21.5%	21.8%	21.2%	0.03
Neurological disorders	15.4%	15.2%	15.3%	15.4%	0.63
Cancer	6.5%	6.7%	6.8%	7.0%	0.002
Mental illness	14.7%	14.5%	14.2%	14.6%	0.05
Discharge location					
Home	60.6%	58.5%	58.4%	57.6%	<0.001
Skilled nursing facility	25.5%	26.6%	26.7%	27.5%	
Rehabilitation facility	2.1%	2.4%	2.4%	2.5%	
Others	11.9%	12.6%	12.5%	12.4%	

Physician spending and patient mortality

We examined 272,979 hospitalizations treated by 8,489 hospitalists. Among hospitalist physicians, patient characteristics were largely similar across adjusted physician-level spending quartiles (Table 1.3); statistically significant differences in patient characteristics across spending quartiles were small in magnitude. The C-statistic of the risk-adjustment model was 0.83 indicating good discriminating power. (Aylin, Bottle et al. 2007)

The overall mortality rate was 11.0%. Within hospital, hospitalist physicians in the highest-spending quartile had lower 30-day mortality than physicians in the bottom-spending quartile, after adjusting for potential confounders (adjusted mortality rate 10.7% vs. 11.2%, adjusted odds ratio [OR] 0.94, 95%CI: 0.90 to 0.98, p=0.002) (Table 1.4).

Table 1.4 Association between adjusted physician spending and 30-day patient mortality, among hospitalist physicians

	30-day mortality rate			
	Unadjusted mortality rate	**Adjusted mortality rate***	**Adjusted odds ratio (95%CI)***	**p-value**
Quartile of physician spending, adjusted				
Q1 (Lowest)	**10.7%** (10.5% to 11.0%)	**11.2%** (11.0% to 11.5%)	Reference	
Q2	**11.0%** (10.8% to 11.2%)	**10.8%** (10.6% to 11.0%)	**0.95** (0.91 to 0.99)	0.02
Q3	**11.0%** (10.8% to 11.2%)	**11.0%** (10.8% to 11.2%)	**0.97** (0.93 to 1.01)	0.15
Q4 (Highest)	**10.6%** (10.4% to 10.9%)	**10.7%** (10.5% to 10.9%)	**0.94** (0.90 to 0.98)	0.002

*Adjusted for patient characteristics, physician characteristics, and hospital-specific fixed effects.

Physician spending and patient readmissions

We examined 268,885 hospitalizations treated by 8,507 hospitalists. The overall 30-day readmission rate was 14.5%. We observed no systematic association between physician-level spending and readmissions, after adjusting for potential confounders (adjusted readmission rates in top and bottom spending quartile 14.9% vs. 15.2%, adjusted OR 0.97, 95% CI: 0.94 to 1.004, p=0.08) (Table 1.5).

Table 1.5 Association between adjusted physician spending and 30-day patient readmissions, among hospitalist physicians

	30-day readmission rate			
	Unadjusted readmission rate	**Adjusted readmission rate***	**Adjusted odds ratio (95%CI)***	**p-value**
Quartile of physician spending, adjusted				
Q1 (Lowest)	**14.9%** (14.7% to 15.2%)	**15.2%** (15.0% to 15.5%)	Reference	
Q2	**14.8%** (14.6% to 15.1%)	**14.9%** (14.6% to 15.2%)	**0.97** (0.94 to 1.003)	0.08
Q3	**14.9%** (14.6% to 15.1%)	**14.9%** (14.6% to 15.1%)	**0.97** (0.94 to 1.003)	0.07
Q4 (Highest)	**15.0%** (14.7% to 15.3%)	**14.9%** (14.6% to 15.2%)	**0.97** (0.94 to 1.004)	0.08

*Adjusted for patient characteristics, physician characteristics, and hospital-specific fixed effects.

Sensitivity analyses

Our findings were robust to several sensitivity analyses, including identifying hospitalist physicians through an alternative method, using different methods for treating outlier cases, incorporating into the analysis patients who were transferred to other acute care hospitals, excluding patients with cancer, and repeating analyses among general internists overall (Tables S1.1-S1.6 in Supplemental Materials).

1.4 DISCUSSION

Using national data on hospitalized Medicare beneficiaries, we found wide variation across individual physicians in risk-adjusted standardized spending – which is effectively an aggregate measure of utilization – after accounting for specific hospital of practice. Between-physician variation in spending was larger than between-hospital variation, supporting the importance of physician-level factors in influencing overall health care spending. Within the same hospital, hospitalist physicians in the highest spending quartile had lower mortality rates than physicians in the lowest quartile, despite treating observably similar patients. Taken together, these findings suggest that not only does physician spending vary substantially even with the same hospital, but that the variation may have important consequences for patient outcomes.

Policy efforts to improve the efficiency of health care should focus on units of observation where key decisions are made, including – and maybe especially – physicians. (Institute of Medicine 2013) However, many federal efforts to improve the quality and value of inpatient care, such as hospital value-based purchasing, penalties for 30-day readmissions, and pay-for-performance, have tended to focus on hospitals, with the notion that hospitals can shape individual physician behavior. Much this is true, given that variation in spending is larger across physicians than across hospitals, policy interventions may be made more effective by broadening the focus to target variation in practice patterns of individual physicians.

Our findings are consistent with several hospital-level studies that demonstrate better patient outcomes in higher-spending hospitals. (Jha, Orav et al. 2009, Yasaitis, Fisher et al. 2009, Chen, Jha et al. 2010, Lagu, Rothberg et al. 2011, Romley, Jena et al. 2011, Stukel, Fisher et al. 2012, Hussey, Wertheimer et al. 2013) Understanding the mechanisms of this relationship is important. If more intensive care itself is responsible for better outcomes, hospital efforts to constrain health care spending among their highest spending physicians may have unintended consequences of negatively affecting patient outcomes. Conversely, physicians' intensity of care may be correlated with other characteristics of physicians that positively influence patient outcomes. In this case, it may be possible to reduce physicians' health care spending without compromising patient outcomes. In addition, in contrast to mortality rates, we did not observe a systematic relationship between physician-level spending and patient readmissions, suggesting the possibility that readmissions may be influenced more by the quality and coordination of care that patients receive after hospital discharge. (Joynt and Jha 2012)

The most important threat to internal validity of our study is that the relationship between physician-level spending and patient outcomes could be confounded by several factors. For example, physicians with higher adjusted spending may treat patients with greater unobserved illness severity. This would, however, lead us to find higher, not lower, mortality among patients treated by higher spending physicians. Alternatively, physicians may differ in the proportion of their patients with advanced directives (e.g., DNR directives), which may be associated with increased mortality and lower use of health care services. However, our results were unaffected by excluding patients with cancer, who may have higher likelihood of DNR directives. More generally, to address confounding, we specifically analyzed hospitalist physicians, among whom patients are plausibly randomly quasi-randomized assigned based on physician work schedules. Consistent with this hypothesis, we demonstrated that observable patient characteristics were generally similar across spending quartiles of hospitalist physicians within the same hospital.

Our study has other limitations. First, we calculated physician-level estimates of Part B spending based on a limited number of hospitalizations per physician. Although we excluded outliers and demonstrated the robustness of our findings to a different threshold number of cases, estimates developed from a larger set of patients seen by each physician could differ. (Adams, McGlynn et al. 2010, Adams, Mehrotra et al. 2010, Mehrotra, Adams et al. 2010, Mehrotra, Adams et al. 2010) Second, the relationship between physician spending and patient mortality was not monotonic; the lowest-spending physicians had the highest 30-day mortality (11.2%), but mortality rates were not substantially different between medium- and highest-spending physicians (e.g., 10.8%, 11.0%, and 10.7%, in quartiles two through four, respectively). This pattern may reflect differential effects of physician spending on mortality depending on the level of spending; for example, most of the benefit of added spending may occur at the lowest levels of spending. Third, Medicare claims data do not permit measurement of inpatient drugs and laboratory tests. If these substituted for the inpatient services we measured, we would overstate variation across physicians. That mortality is related to the variation we measured, however, suggests the variation we did measure is important. Lastly, our analysis of physician spending and outcomes was limited to Medicare patients treated by hospitalists in order to exploit the plausibly random assignment of patients across hospitalists. Results may differ for surgical conditions or medical conditions treated by other specialists.

In summary, we found substantial variation in standardized Part B spending across physicians after accounting for specific hospital of practice. Variation in spending across physicians exceeds variation in spending across hospitals. Among hospitalist physicians in the same hospital, those in the highest spending quartile have lower patient mortality compared to physicians in the bottom most spending quartile, but no differences in readmissions. Given larger variation in spending across physicians than across hospitals, our findings suggest that policies focused on individual physicians may be as or more effective than those targeted toward hospitals or regions. Moreover, interventions targeted at high-spending physicians to reduce spending, without accounting for their quality of care, may have the unintended consequence of negatively impacting patient outcomes.

Chapter 2: Physician Experience and Outcomes of Hospitalized Medicare Beneficiaries[2]

[2] A version of this chapter is submitted to a journal for publication. Please do not circulate.

2.1 INTRODUCTION

The purpose of continuing medical education and board recertification is to support physicians in maintaining clinical competence as clinical knowledge evolves. Although clinical skills accumulated by more experienced physicians may improve quality of care, clinical guidelines change frequently and new technologies are developed rapidly. Learning new guidelines and incorporating them into clinical practice is time consuming may at times be overwhelming. (Meskauskas and Webster 1975, Ray, Federspiel et al. 1976, Norcini, Lipner et al. 1985, Shapiro 1989, Salem-Schatz, Avorn et al. 1990, Gemson, Colombotos et al. 1991, Ramsey, Carline et al. 1991, Schwartz, Lewis et al. 1991, Ayanian, Hauptman et al. 1994, Czaja, McFall et al. 1994, Streja and Rabkin 1999, Fehrenbach, Budnitz et al. 2001) As a result, physicians' knowledge and skills may become outdated or practicing physicians may lack the time to learn and implement into practice. (Meskauskas and Webster 1975, Ray, Federspiel et al. 1976, Norcini, Lipner et al. 1985, Shapiro 1989, Salem-Schatz, Avorn et al. 1990, Gemson, Colombotos et al. 1991, Ramsey, Carline et al. 1991, Schwartz, Lewis et al. 1991, Ayanian, Hauptman et al. 1994, Czaja, McFall et al. 1994, Streja and Rabkin 1999, Fehrenbach, Budnitz et al. 2001) Furthermore, at some point age itself might degrade some essential skills. In this context, the relationship between duration of clinical practice and patient outcomes has important implications for ensuring the consistent provision of high quality care, and inform recent debates within medicine over the role of recertification. (Bauchner, Fontanarosa et al. 2015, Teirstein and Topol 2015)

Several studies of changes in physician knowledge and quality of care processes over a physician's career have found a negative relationship between years in practice and clinical knowledge (Meskauskas and Webster 1975, Norcini, Lipner et al. 1985, Shapiro 1989, Salem-Schatz, Avorn et al. 1990, Gemson, Colombotos et al. 1991, Ramsey, Carline et al. 1991, Ayanian, Hauptman et al. 1994), quality of care with respect to diagnosis, screening and preventive care (Gemson, Colombotos et al. 1991, Schwartz, Lewis et al. 1991, Czaja, McFall et al. 1994, Streja and Rabkin 1999), and adherence to standards of appropriate therapy.(Ray, Federspiel et al. 1976, Fehrenbach, Budnitz et al. 2001) However, evidence is limited regarding the association between length of time in practice and clinical outcomes, especially for medical conditions. (Choudhry, Fletcher et al. 2005) Previous studies have been limited in size (Davidson, Molloy et al. 1995, Hartz, Kuhn et al. 1999, Katon, Rutter et al. 2000, O'Neill, Lanska et al. 2000, Blanc, Trupin et al. 2003, Southern, Bellin et al. 2011) or disease scope (Burns and Wholey 1991, Norcini, Kimball et al. 2000, Epstein, Srinivas et al. 2013), and have not been nationally representative. Therefore, whether physicians' length of time in practice is associated with better or worse clinical outcomes among patients with medical conditions is largely unknown.

Using nationally representative data on Medicare beneficiaries hospitalized with a medical condition in 2011-2013, and employing the quasi-randomization of these patients to hospital-based physicians ("hospitalists") of varying years of experience within the same hospital, we sought answers to three questions. First, is a physician's duration in practice since completion of residency associated with patient mortality? Second, do such effects vary with the physician's volume of patients treated? Third, how are years in practice associated with other measures of care quality and utilization such as readmission rates and costs of care?

2.2 METHODS

Data

We used the Medicare 20% Carrier and corresponding Inpatient Files to identify Medicare fee-for-service (FFS) beneficiaries age 65 years or older who were hospitalized in acute care hospitals between January 1, 2011 and December 31, 2013 for medical conditions as identified by a medical diagnosis related group (MS-DRG). To allow a sufficient follow-up period, we excluded patients admitted in December 2013 from analyses of 30-day mortality and patients discharged in December 2013 from readmission analyses. We also excluded elective hospitalizations and patients who left against medical advice.

The physician who accounted for the most Medicare Part B costs during each hospitalization was assigned to that hospitalization using the National Provider Identifier (NPI). (McWilliams, Landon et al. 2014) On average, these physicians were responsible for 52% of total Part B spending for hospitalizations. Part B costs are comprised of professional and other fees determined by the physician. For patients transferred to other acute care hospitals (1.2% of hospitalizations), we attributed the multi-hospital episode of care and associated outcomes to the assigned physician of the initial hospitalization.(Ross, Normand et al. 2010, Drye, Normand et al. 2012)

Physician characteristics were collected by linking the NPI to a database obtained from Doximity, a company that provides online professional networking services for U.S. physicians. The database included information on years of residency completion, age, sex, specialty, allopathic (MD) versus osteopathic (DO) training, and medical school attended (which we used to identify international medical graduates). Specialty information in the Doximity database was obtained from multiple sources and data partnerships, including the National Plan and Provider Enumeration System NPI Registry, the American Board of Medical Specialties, other specialty societies, state licensing boards, and collaborating hospitals and medical schools. Descriptions of the database and validation of its accuracy have been published elsewhere. (Jena, Khullar et al.

2015, Olmsted, Geisen et al. 2015)

To address the possibility that physicians in practice longer may treat patients with greater or lesser unmeasured illness severity, we restricted the study population to patients treated by hospital-based physicians ("hospitalists"). We hypothesized that, within the same hospital, patients treated by hospitalists may plausibly be quasi-randomized to a physician based on that physician's work schedule. We defined hospitalists as general internists (identified using specialty data in the Doximity database) with at least 20 E&M claims in a given year, who filed at least 90% of their total E&M claims in an inpatient setting as defined by Current Procedural Terminology [CPT] codes 99221-99223, 99231-99233, and 99251-99255, a previously validated method. (Kuo, Sharma et al. 2009) As a sensitivity analysis, we also employed different thresholds for identifying hospitalists (80% and 95% rather than 90%). We assessed the validity of the quasi-randomization assumption by testing balance of a broad range of patient characteristics across categories of hospitalists' years in practice.

Physician years in practice

We measured physician years in practice by years since residency completion, both as a continuous and a categorical variable (in 10-year increments, categorized into 1-4, 5-9, 10-14, and ≥15 years). To evaluate the impact of missing data on year of residency completion (missing for 35.5% of physicians), in a secondary analysis, we used physician age – missing for only 6.1% of physicians – as a proxy for physicians' years in practice.

Adjustment variables

We adjusted for three major categories of variables: patient characteristics, physician characteristics, and hospital-specific fixed effects. Patient characteristics included: age in 5-year increments (from 65-69 years, 70 to 74, and so on through 90 to 94, and ≥ 95), sex, race or ethnic group (non-Hispanic white, non-Hispanic black, Hispanic, other), primary diagnosis (defined by MS-DRG), Elixhauser comorbidity index (Elixhauser, Steiner et al. 1998), an indicator for dual Medicaid coverage, and year indicators. Physician characteristics included sex, allopathic versus osteopathic training, and whether a physician graduated from a foreign medical school. We excluded physician age from our models because of its close relationship with physicians' years in practice (correlation coefficient, 0.86). As a secondary analysis, we evaluated the impact of substituting physician age for years in practice. Because the average level of physician years in practice may vary across hospitals and be correlated at the hospital level with unobserved patient illness severity, we included hospital-specific fixed effects, effectively comparing patient outcomes across physicians with varying years in practice within the same hospital.

Statistical analysis

First, we examined the association between physician years in practice and 30-day mortality by estimating multivariable linear probability models that adjusted for the patient and physician characteristics and hospital-specific fixed effects (details about adjustment variables are described above). (Wooldridge 2010) We calculated risk-adjusted 30-day mortality rates by physician years in practice through estimating predicted probabilities of outcomes for each hospitalization averaged over the distribution of covariates in our national sample, fixing physician years in practice at each level (known as the marginal standardization form of predictive margins (Williams 2012, Zhang, Baik et al. 2012)). To account for potential correlations among outcomes for patients of the same physicians, we clustered standard errors at the physician level. We repeated the same analysis using physician years in practice as a continuous variable.

Next, because physicians who care for a high volume of patients may better maintain clinical knowledge and skills (Halm, Lee et al. 2002, Moscucci, Share et al. 2005, Nallamothu, Gurm et al. 2011, Joynt, Orav et al. 2013), we examined whether the association between physician years in practice and patient mortality was modified by patient volume. Patient volume was estimated from the number of hospitalized Medicare patients treated by each physician in our data (based on the data that Medicare FFS beneficiaries account for approximately 35% of all hospitalizations in the U.S. (Argency for Healthcare Research and Quality)). We classified physicians into terciles of patient volume: low (≤66 hospitalizations per year), medium (67-142 hospitalizations), and high (≥143 hospitalizations). Within each group, we examined the association between physician years in practice and patient mortality, adjusting for patient and physician characteristics and hospital-specific fixed effects. We also conducted a Wald test for interaction on the logit scale, which would signify a disproportionate multiplicative effect of physician years in different patient groups.

Finally, we evaluated the association between physician years in practice and 30-day readmission rates and costs of care. We standardized costs to account for variation in factor prices and malpractice premiums across areas. (Centers for Medicare & Medicaid Services 2013) We assessed the associations between physician years in practice and 30-day readmissions and costs of care with the same regression model used for the mortality analysis. We focused on Part B costs because, conditional on admission, Part A spending is largely invariant to physician decisions about an individual patient's treatment (since MS-DRG payments are fixed within hospital except for a small set of outlier cases).

Secondary analyses

We conducted additional secondary analyses. First, to address the possibility that some physicians have better patient outcomes by keeping their patients in the hospital longer, we adjusted for patients' length of stay (LOS) (continuous variable with quadratic and cubic terms to allow for non-linear relationship). Second, to test the generalizability of our findings, we repeated our analyses using all general internists (as opposed to focusing on hospitalists) treating Medicare patients in the U.S. Third, to focus on more homogenous patient populations, we evaluated mortality effects separately in sepsis, pneumonia, and congestive heart failure (CHF), common conditions treated by general internists and with relatively high mortality. We identified these conditions using International Classification of Diagnosis, 9th Edition (ICD-9) codes (Table S2.1 in Supplemental Materials). Fourth, we analyzed whether the association between physician years in practice and patient mortality varied by underlying patient risk, as categorized by 3 risk groups (low, medium, and high, defined on the basis of tercile of predicted mortality). Fifth, as physicians in practice for very long time may have prominent influence on our estimates, we reanalyzed the data after excluding physicians who were in practice for more than 20 years. Lastly, we evaluated whether our findings depended on model specification by repeating our analyses using logistic regression. To improve convergence of the likelihood maximization algorithm, we combined DRG codes without an outcome into clinically similar categories. This study was approved by the institutional review board at Harvard Medical School.

2.3 RESULTS

Physician and patient characteristics

The median and mean years in practice among 14,650 hospitalist physicians in our sample in 2013 were 8.0 and 9.7 years, respectively (Figure 2.1). A broad range of patient characteristics were balanced across physicians with different years in practice, suggestive that within hospital, we might plausibly regard assignment of patients to physicians with varying years in practice as random (Table 2.1).

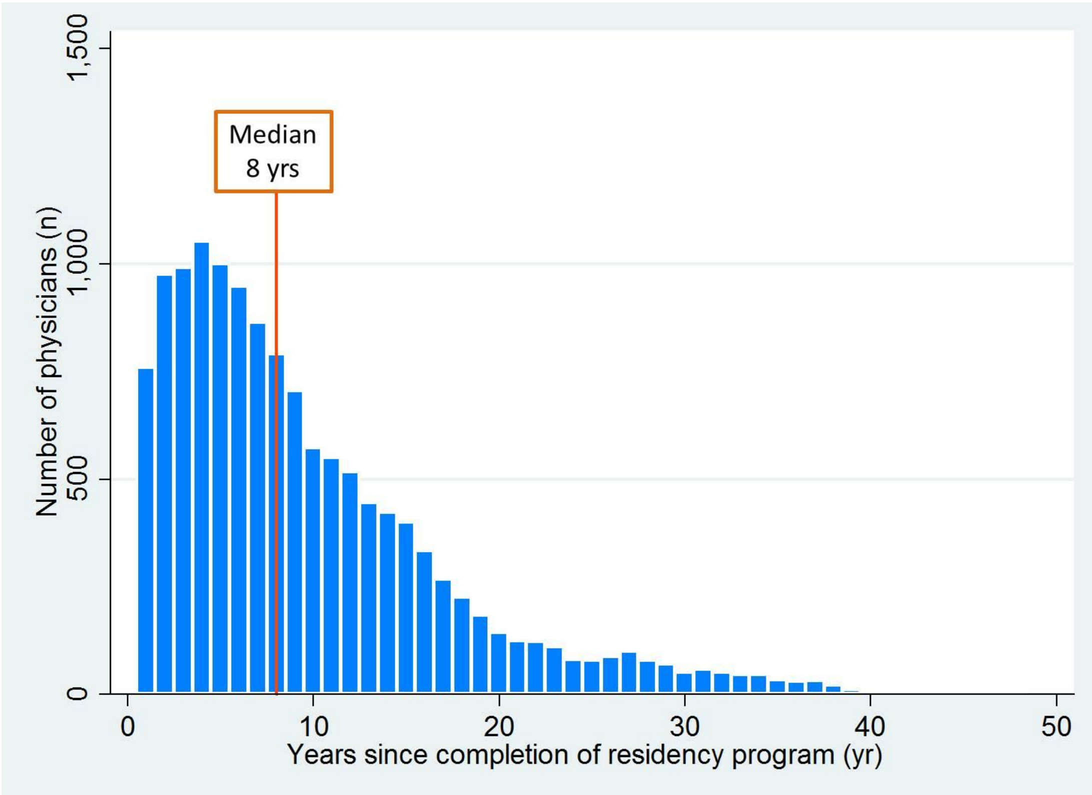

Figure 2.1 | Years since completion of residency programs. The median and mean years in practice in 2013 were 8.0 and 9.7 years, respectively. Data on hospitalist physicians caring for hospitalized Medicare beneficiaries during 2011-2013. Physician years in practice, defined as years since residency completion, measured as of 2013.

Table 2.1 Patient characteristics according to years of physician years in practice

	Years since completion of residency			
	1-4 years (n=5,066)	**5-9 years (n=4,128)**	**10-14 years (n=2,543)**	**≥15 years (n=2,913)**
No. of Medicare hospitalizations in 2011-2013, n	108,918	124,495	74,844	77,902
No. of hospitalizations per physician per year*, n	71.7	100.5	98.1	89.1
Patient's age, yrs (sd)	80.5 (8.6)	80.6 (8.6)	80.7 (8.5)	80.6 (8.5)
Female, %	60.4%	60.2%	60.3%	60.7%
Race, %				
White	82.4%	82.6%	81.9%	83.2%
Black	10.8%	10.1%	10.2%	9.6%
Hispanic	4.1%	4.4%	4.9%	4.5%
Other races	2.7%	2.9%	3.0%	2.6%
Medicaid status	22.3%	22.3%	23.1%	23.5%
Coexisting condition, %				
CHF	19.8%	20.1%	20.2%	19.9%
COPD	24.3%	24.6%	24.8%	26.1%
Diabetes	31.7%	31.1%	31.5%	31.2%
Renal failure	22.5%	21.9%	21.8%	21.8%
Neurological disorders	14.7%	15.2%	15.2%	16.0%
Cancer	7.3%	7.1%	6.9%	7.1%
Mental illness	14.4%	14.1%	14.6%	14.7%
Discharge location, %				
Home	60.2%	59.2%	58.5%	57.1%
Skilled nursing facility	25.7%	26.0%	26.2%	25.9%
Rehabilitation facility	2.3%	2.4%	2.4%	2.4%
Others	11.7%	12.5%	12.9%	14.6%

*Estimated from 20% Medicare Inpatient Files, based on the data that Medicare patients comprise 35% of all hospitalizations in the U.S.

Physician years in practice and patient mortality

The overall 30-day mortality rate in our sample of 386,159 hospitalizations was 11.0%. After adjusting for patient and physician characteristics and hospital-specific fixed effects, physicians in practice longer had higher patient mortality than physicians in practice for fewer years; for example, the risk-adjusted 30-day mortality rate 11.6% among physicians with ≥15 years in practice vs. 10.7% among those with <5 years; adjusted risk difference +0.9%, 95%CI: +0.6% to +1.2%; $p<0.001$) (Table 2.2). Adjusted 30-day mortality rates increased monotonically with experience. Analysis of physician years in practice as a continuous predictor showed that 10 additional years in practice was associated on average with a 0.5 percentage point increase (95%CI: +0.3% to +0.7%, $p<0.001$) in patient 30-day mortality.

Table 2.2 Association between physicians' years in practice and 30-day patient mortality

Years in practice	N of hospitalizations (N of physicians)	Unadjusted 30-day patient mortality rate (95%CI)	Risk-adjusted 30-day patient mortality rate (95%CI)*	Adjusted risk difference*	
				Risk difference (95%CI)	p-value
Years in practice as categorical variable					
1-4	108,918 (5,066)	**10.2%** (10.0% to 10.4%)	**10.7%** (10.5% to 10.8%)	Reference	
5-9	124,495 (4,128)	**10.9%** (10.7% to 11.0%)	**10.9%** (10.7% to 11.1%)	**+0.3%** (+0.01% to +0.5%)	0.06
10-14	74,844 (2,543)	**11.2%** (11.0% to 11.5%)	**11.0%** (10.8% to 11.3%)	**+0.4%** (+0.06% to +0.7%)	0.02
≥15	77,902 (2,913)	**12.1%** (11.9% to 12.3%)	**11.6%** (11.3% to 11.8%)	**+0.9%** (+0.6% to +1.2%)	<0.001
Years in practice as continuous variable					
For every 10 yrs increase	386,159 (14,650)			**+0.5%** (+0.3% to +0.7%)	<0.001

*Adjusted for patient and physician characteristics and hospital-specific fixed effects.

Physician years in practice and mortality, by patient volume

Years in practice was positively associated with patient mortality among low- and medium-volume physicians, but not among high-volume physicians (Table 2.3). For example, compared to physicians in practice for <5 years, physicians in practice for ≥15 years had higher patient mortality when their patient volume was low (risk-adjusted mortality rate 11.4% vs. 15.0%; adjusted risk difference +3.6%, 95%CI: +2.3% to +4.9%, p<0.001) or medium (risk-adjusted mortality rate 10.5% vs. 12.0%; adjusted risk difference +1.5%, 95%CI: +0.7% to +2.2%, p<0.001). However, we observed no difference in patient mortality among high-volume physicians (risk-adjusted mortality rate 10.7% vs. 10.7%; adjusted risk difference -0.04%, 95%CI: -0.5% to +0.4%, p=0.85). There was a statistically significant interaction between years in practice and patient volume on the logit scale (p <0.001).

Table 2.3 Physicians' years in practice and 30-day patient mortality, by patient volume

Patient volume	N of hospitalizations (N of physicians)	Years in practice	Risk-adjusted 30-day mortality rate (95%CI)*	Adjusted risk difference*	
				Risk difference (95%CI)	p-value
Low-volume physicians (≤66 cases/yr)	14,016 (1,926)	1-4	**11.4%** (10.9% to 12.0%)	Reference	
	8,830 (1,268)	5-9	**14.0%** (13.2% to 14.8%)	**+2.6%** (+1.6% to +3.6%)	<0.001
	5,366 (772)	10-14	**13.9%** (12.9% to 14.9%)	**+2.5%** (+1.3% to +3.7%)	<0.001
	7,030 (1,066)	≥15	**15.0%** (14.0% to 16.0%)	**+3.6%** (+2.3% to +4.9%)	<0.001
Medium-volume physicians (67-142 cases/yr)	35,910 (1,718)	1-4	**10.5%** (10.2% to 10.8%)	Reference	
	32,567 (1,508)	5-9	**10.9%** (10.6% to 11.3%)	**+0.5%** (-0.04% to +1.0%)	0.07
	18,034 (843)	10-14	**11.5%** (11.0% to 12.0%)	**+1.0%** (+0.4% to +1.7%)	0.002
	17,824 (840)	≥15	**12.0%** (11.4% to 12.6%)	**+1.5%** (+0.7% to +2.2%)	<0.001
High-volume physicians (≥143 cases/yr)	58,992 (1,240)	1-4	**10.7%** (10.5% to 11.0%)	Reference	
	83,098 (1,616)	5-9	**10.7%** (10.4% to 10.9%)	**-0.09%** (-0.4% to +0.2%)	0.58
	51,444 (976)	10-14	**10.5%** (10.2% to 10.8%)	**-0.2%** (-0.6% to +0.2%)	0.24
	53,048 (967)	≥15	**10.7%** (10.4% to 11.0%)	**-0.04%** (-0.5% to +0.4%)	0.85

There was a statistically significant interaction between years in practice and patient volume on the logit scale ($p < 0.001$).

*Adjusted for patient and physician characteristics and hospital-specific fixed effects.

Patient readmission rates and costs of care

We found no association between physician years in practice and patient 30-day readmission rates (Table 2.4). Although differences in costs of care between physicians of varying years in practice were statistically significant, they were practically small. For example, physicians with ≥15 years in practice had adjusted Part B costs of $1,066 (95%CI: $1,054 to $1,077) compared to $1,010 (95%CI: $1,004 to $1,017) among physicians with <5 years in practice.

Table 2.4 Association linking physician years in practice with 30-day readmission rates and costs of care

	N of hospitalizations (N of physicians)	Years in practice	Risk-adjusted patient outcomes (95%CI)*	Adjusted difference*	
				Difference (95%CI)	p-value
30-day readmission rate (%)	107,036 (5,104)	1-4	**15.0%** (14.8% to 15.2%)	Reference	
	121,300 (4,173)	5-9	**15.0%** (14.8% to 15.2%)	**-0.04%** (-0.3% to +0.3%)	0.81
	72,648 (2,551)	10-14	**14.9%** (14.6% to 15.2%)	**-0.1%** (-0.5% to +0.3%)	0.58
	75,183 (2,893)	≥15	**14.9%** (14.6% to 15.2%)	**-0.1%** (-0.5% to +0.3%)	0.54
Total Part B costs ($)	108,918 (4,884)	1-4	**$1,010** ($1,004 to $1,017)	Reference	
	124,495 (4,392)	5-9	**$1,025** ($1,018 to $1,032)	**+$14** (+$5 to +$24)	0.003
	74,844 (2,591)	10-14	**$1,043** ($1,033 to $1,054)	**+$32** (+$19 to +$45)	<0.001
	77,902 (2,873)	≥15	**$1,066** ($1,054 to $1,077)	**+$55** (+$41 to +$69)	<0.001

*Adjusted for patient and physician characteristics and hospital-specific fixed effects

Secondary analyses

Our findings were robust to using physician age as a proxy for their years in practice; we found that patients of older physicians had higher mortality than those of younger physicians (Table S2.2 in Supplemental Materials). For instance, physicians aged 60 years or older had higher patient mortality than physicians younger than 40 years (risk-adjusted 30-day mortality rate

12.0% vs. 10.7%; adjusted risk difference +1.3%, 95%CI: +0.8% to +1.8%, p<0.001). Our findings were not qualitatively affected by using different thresholds (80% or 95%) for identifying hospitalists (Table S2.3), or additional adjustment for LOS, suggesting that differences in LOS decisions by physicians did not mediate the effect physician years in practice on patient mortality (Table S2.4). In an analysis of all general internists (including both hospitalists and non-hospitalists), physicians in practice longer continued to have higher patient mortality (Table S2.5). When we stratified by primary diagnosis, we found that physicians with greater years in practice had higher patient mortality for sepsis and CHF, but not for pneumonia; however, a formal interaction test was not statistically significant on the logit scale (p=0.50) (Table S2.6). The influence of physician years in practice appeared more prominent among severely ill patients, but the interaction test was not statistically significant on the logit scale (p=0.17) (Table S2.7). The relationship remained unchanged when we excluded physicians who were in practice for more than 20 years (every 10 years in practice was associated with 0.5% increase in patient mortality [95%CI: +0.2% to +0.7%, p<0.001]) (data not shown), or when we used a logistic regression model (Table S2.8).

2.4 DISCUSSION

In a national sample of Medicare beneficiaries hospitalized with medical conditions, patients treated by physicians with greater years in practice had higher 30-day mortality compared to those cared for by physicians in practice for fewer years, despite similar patient characteristics across physicians of varying years in practice. Each additional 10 years in practice was associated with a 0.5 percentage point increase in patient 30-day mortality. These effects were found among low- and medium-volume physicians, but not among high-volume physicians. Readmission rates and costs of care did not meaningfully vary with physician years in practice. Taken together, these findings suggest that outcomes of hospitalized patients may be worse among physicians with greater years in practice, except for physicians who continue to treat a large number of patients.

Several mechanisms may explain the observed association. As previous studies suggested, physicians with more years elapsed since training may be less likely to utilize newly proven therapies, may less frequently adhere to evidence-based guidelines, and may more frequently rely on evidence that is not up-to-date. (Freiman 1985, Hlatky, Cotugno et al. 1988) In addition, while intense exposures to a large number of patients during residency training may enable physicians shortly out of residency training to provide high quality care, the benefits of this training experience may wane over time if physicians care for fewer inpatients after residency. The lack of association between physician years in practice and patient mortality among physicians with higher patient volume is consistent with this potential mechanism.

Our findings have important implications for the current debate regarding the role of specialty boards in ensuring clinical competence over a physician's career. (Bauchner, Fontanarosa et al. 2015, Teirstein and Topol 2015) Given rapidly changing evidence in medicine, specialty boards such as the American Board of Internal Medicine (ABIM) have instituted several requirements for maintenance of certification (MOC), including specific MOC programs, requirements for Continuing Medical Education (CME) credits, and written examinations. These requirements have been criticized, however, for lacking evidence linking them to improved physician performance, particularly since just updating knowledge and completing examinations may be insufficient to assure that quality of care does not diminish over a clinician's career. (Bauchner, Fontanarosa et al. 2015, Teirstein and Topol 2015) Although our findings suggest the critical importance of some form of continuing medical education, further studies should directly evaluate the impact of recertification and other continuing education practices on maintenance of physician expertise.

Our study has limitations. First, our study design was premised on the notion that patients are quasi-randomized to hospitalist physicians of varying experience within the same hospital, a study design that we employed to address the issue that physicians with greater years in practice may treat patients with higher unobserved illness severity. Although this design is supported by our observation that hospitalist physicians of varying years in practice treat patients that are largely similar on a range of observed demographic and clinical characteristics, residual confounding is nonetheless possible. Second, the data on years since residency completion was missing for some physicians; however, our findings were robust to using physician age as a proxy for years in practice. Third, we found that the positive association between years since residency and patient mortality was driven primarily by physicians treating a low to medium volume of patients, suggestive that high volumes may be 'protective' of clinical skills. However, the causal relationship between practice volume and skills is bidirectional. Physicians whose skills are declining might either self-select or be encouraged by others to leave positions in which they are responsible for clinical management of large numbers of patients, and may therefore treat fewer patients over time. It is nonetheless important to understand that low- and medium-volume physicians with greater years in practice have worse patient outcomes, and may require interventions such as rigorous assessments for the safety and appropriateness of providing health care services to patients. Finally, our findings may not generalize to the non-Medicare population or to patients cared for by surgeons or other specialists.

In summary, we found that hospitalized patients treated by physicians with greater years in practice have higher mortality than patients cared for by physicians with fewer years in practice, except for physicians with high patient volume. Our findings suggest that evaluating patient outcomes, particularly among older physicians with low patient volume, may be necessary to guarantee that high quality care is provided by physicians throughout their careers.

Chapter 3: Physician Sex and Outcomes of Hospitalized Medicare Patients[3]

[3] A version of this chapter is submitted to a journal for publication. Please do not circulate.

3.1 INTRODUCTION

There is clear evidence that men and women may practice medicine differently. Female physicians are more likely to adhere to clinical guidelines (Kim, McEwen et al. 2005, Berthold, Gouni-Berthold et al. 2008, Baumhakel, Muller et al. 2009), more often provide preventive care (Franks and Clancy 1993, Lurie, Slater et al. 1993, Andersen and Urban 1997, Franks and Bertakis 2003), employ more patient-centered communication (Bertakis, Helms et al. 1995, Krupat, Rosenkranz et al. 2000, Roter, Hall et al. 2002, Roter and Hall 2004), perform as well or better on standardized examinations (Ferguson, James et al. 2002), and provide more psychosocial counseling (Roter, Hall et al. 2002) to their patients than their male peers. Although the data are consistent, studies that compare practice patterns between male and female physicians have often been small, limited in focus, and have not analyzed patient outcomes. Whether patient outcomes differ between male and female physicians is largely unknown.

Female physicians now account for approximately one third of the U.S. physician workforce (Kaiser Family Foundation 2016) and comprise half of all U.S. medical school graduates (Association of American Medical Colleges). Understanding differences in patient outcomes between male and female physicians is important because, despite evidence suggesting that female physicians may provide higher quality care (Franks and Clancy 1993, Lurie, Slater et al. 1993, Bertakis, Helms et al. 1995, Andersen and Urban 1997, Krupat, Rosenkranz et al. 2000, Ferguson, James et al. 2002, Roter, Hall et al. 2002, Franks and Bertakis 2003, Roter and Hall 2004, Kim, McEwen et al. 2005, Berthold, Gouni-Berthold et al. 2008, Baumhakel, Muller et al. 2009), some have argued that higher salaries among male physicians (Nonnemaker 2000, Jagsi, Griffith et al. 2012, Seabury, Chandra et al. 2013) may be justified by several factors that may in theory compromise the quality of care provided by female physicians, including career interruptions for childrearing, higher rates of part-time employment, and greater tradeoffs between home and work responsibilities (Franks and Clancy 1993, Lurie, Slater et al. 1993, Bertakis, Helms et al. 1995, Andersen and Urban 1997, Krupat, Rosenkranz et al. 2000, Ferguson, James et al. 2002, Roter, Hall et al. 2002, Franks and Bertakis 2003, Roter and Hall 2004, Kim, McEwen et al. 2005, Berthold, Gouni-Berthold et al. 2008, Baumhakel, Muller et al. 2009, Jolly, Griffith et al. 2014). Concrete evidence of whether patient outcomes differ between male and female physicians is therefore needed.

The hospital as a context for studying patient outcomes between male and female physicians offers a unique advantage to studying differences in the outpatient setting – within a given hospital, there is plausibly less selection by the patient (of the physician) or the physician (of the patient). While some patients choose their primary care physician and may factor in physician sex into their decision, patients hospitalized urgently or emergently are less likely to select their physicians. Thus, an inpatient setting may provide a better context for assessing differences in practice patterns and clinical outcomes between male and female physicians.

Therefore, we used data for a national sample of Medicare beneficiaries hospitalized with medical conditions to investigate three key questions: First, are there important differences in clinical outcomes for patients cared for by female physicians compared to those treated by male physicians? Second, does the influence of physician sex on patient outcomes differ across a variety of conditions for which patients are commonly hospitalized? And finally, do these differences vary with the underlying severity of a patient's illness?

3.2 METHODS

Data

We analyzed the Medicare 20% Carrier File and corresponding Inpatient File to identify Medicare fee-for-service (FFS) beneficiaries aged 65 years or older who were hospitalized in acute care hospitals from January 1, 2011 through December 31, 2013. Our study population was restricted to hospitalizations due to medical conditions identified by the presence of a medical diagnosis related group (DRG). In order to allow for sufficient follow-up, patients admitted in December 2013 were excluded from the analysis of 30-day mortality, and patients discharged in December 2013 were excluded from 30-day readmission analyses. Patients who were out of hospital for less than 30 days at the time of admission were excluded from the denominator in readmission analysis. Patients who were transferred between acute care hospitals were linked into a single episode of care with the patient outcome attributed to the first hospitalization. (Ross, Normand et al. 2010, Drye, Normand et al. 2012) We restricted our sample to non-elective hospitalizations and excluded patients who left against medical advice.

We assigned each hospitalization to a physician based on the National Provider Identifier (NPI) in the Carrier File that accounted for the most Medicare Part B spending during that hospitalization. (McWilliams, Landon et al. 2014) We focused on hospitalizations in which the assigned physician was a general internist, which was identified by linking NPI to a comprehensive physician database obtained from Doximity, a company that provides online professional networking services for U.S. physicians. Specialty information in the database was obtained from multiple sources and data partnerships, including the National Plan and Provider Enumeration System (NPPES) NPI Registry, the American Board of Medical Specialties, other specialty societies, state licensing boards, and collaborating hospitals and medical schools.

Approximately 98% of physicians in our Medicare sample were matched to the database. Details of the database and validation have been published elsewhere. (Jena, Khullar et al. 2015)

Adjustment variables

We accounted for several differences in patient and physician characteristics between male and female physicians. Patient characteristics included patient age in 5-year increments (from 65-69 years, 70 to 74, and so on through 90 to 94, and ≥ 95), sex, race or ethnic group (non-Hispanic white, non-Hispanic black, Hispanic, and other), MS-DRG, Elixhauser comorbidity index (27 coexisting conditions),(Elixhauser, Steiner et al. 1998) and an indicator for dual Medicaid coverage (beneficiaries who were eligible for Medicaid for at least 1 month in a given year). Physician characteristics included physician age in 5-year increments (<35 years, 35-39, and so on through 65-69, and ≥ 70), years since completion of residency, and allopathic (MD) versus osteopathic (DO) training.

Statistical analysis

We first compared differences in physician and patient characteristics between male and female physicians. We then examined the association between physician sex and 30-day mortality and readmissions by using two sets of multivariable linear probability models. (Wooldridge 2010) The first model accounted for patient characteristics described above and hospital-specific fixed effects. The second model included additional adjustment for physician characteristics. Having hospital-specific fixed effects in our models allowed for comparison of male and female physicians within the same hospital. We calculated risk-adjusted 30-day mortality and readmissions by estimating predicted probabilities of outcomes for each hospitalization with the distribution of covariates in the national sample, and fixing physician sex at female versus male (known as the marginal standardization form of predictive margins). (Williams 2012) To address potential correlation of patient outcomes within the same physician, standard errors were clustered at the physician-level.

Next, we assessed whether differences in patient outcomes between male and female physicians differed according to the primary condition for which a patient was admitted. We evaluated the eight most common medical conditions treated by general internists in the Medicare data: Congestive heart failure (CHF), sepsis, pneumonia, chronic obstructive pulmonary disease (COPD), urinary tract infection (UTI), acute renal failure, arrhythmia, and gastrointestinal bleeding (GI bleeding). A list of International Classification of Diagnosis, 9th Edition (ICD-9) codes for these conditions is available in Table S3.1 in Supplemental Materials.

Finally, we assessed whether differences in outcomes between male and female physicians varied according to illness severity. We used a risk-adjustment model, a logistic regression model with 30-day mortality as an outcome and patient-level adjustment variables listed above as explanatory variables, to predict each patient's likelihood of death and then classified patients into quintiles of predicted mortality. Within each quintile of predicted mortality, we examined differences in outcomes between patients who were treated by male versus female physicians, adjusting for patient and physician characteristics and hospital-specific fixed effects. We formally tested if differences in patient outcomes between male and female physicians varied by patient illness severity using Wald tests.

Analysis of potential mechanisms

We explored potential mechanisms for observed differences between male and female physicians, including differences in average utilization of care, length of stay (LOS), patient volume, or discharge location (e.g., differences in 30-day mortality between male and female physicians may occur if the likelihood of discharge home versus skilled nursing facility differs by physician sex). Utilization of care was measured by total Part B spending per hospitalization, using standardized costs calculated with the method established by the Centers for Medicare and Medicaid Services (CMS). (Centers for Medicare & Medicaid Services 2013) LOS and patient volume were categorized into deciles to allow non-linear relationships. We also investigated whether the relationship between physician sex and patient outcome varied according to physician age, by performing a stratified analysis by physician age. In doing so, we categorized physician age into <40, 40-49, and ≥50 years.

Sensitivity analyses

We conducted a series of sensitivity analyses. First, in order to address the possibility that female physicians may treat healthier patients, we restricted the study population to patients treated by hospitalist physicians (internists who specialize in hospital-based care). We hypothesized that within the same hospital, patients treated by hospitalist physicians are quasi-randomized to a given physician based on that physician's particular work schedule or patient census. We defined hospitalists using a validated approach; general internists with at least 20 E&M billings in a given year who filed at least 90% of their total E&M billings in an inpatient setting as defined by Current Procedural Terminology [CPT] codes 99221-99223, 99231-99233, and 99251-99255). (Kuo, Sharma et al. 2009) Second, we addressed the possibility that within some hospitals internists may practice in intensive care units (in lieu of, or in addition to, critical care specialists), and that male internists are more likely to work in intensive care units and have severely ill patients. To address this issue, we restricted our analysis to hospitals without a medical ICU to determine if the patterns of care were similar. Third, to account for the possibility that our categorization of age variables may not be able to fully account for differences between female and male physicians, we used physician age and patient age as continuous variables, with quadratic and cubic terms to allow for non-linear relationships with patient outcomes. Fourth, we used logistic regression models instead of linear probability models. To overcome a failure of the likelihood maximization algorithm to converge, we combined DRG codes that had no event into clinically similar categories. Lastly, our initial analysis showed that black patients had lower mortality and higher readmissions compared to white patients. We found similar patterns for patients who were also eligible for Medicaid ("dual-eligibles") as well. Although such relationships have been shown in previous studies (Jha, Shlipak et al. 2001, Pippins, Fitzmaurice et al. 2007, Gordon, Nowlin et al. 2010, Joynt, Orav et al. 2011, Bennett and Probst 2015, Rahman, Tyler et al. 2015), we addressed the possibility of measurement error in these variables influencing our findings by reanalyzing our data after excluding these two variables from our regression models. This study was approved by the institutional review board at Harvard Medical School.

3.3 RESULTS

Characteristics of female and male physicians

In 2011-2013, there were 54,511 general internists who treated at least one Medicare beneficiary who was hospitalized with a medical condition (Table 3.1). Among those, 31.5% of physicians were female and 68.5% were male. Female physicians were younger (43.4 vs 48.4 years old), graduated more recently from residency (10.6 vs 15.4 years), were more likely to have undergone osteopathic training (8.3% vs 6.8%), and treated fewer patients (30.4 vs 41.1 Medicare hospitalizations per year), compared to male physicians. Patient characteristics were relatively balanced between those treated by female physicians versus those treated by male physicians with one modest exception: female physicians treated slightly higher proportions of female patients than male physicians (62.2% vs 60.2%, $p<0.001$).

Table 3.1 Physician and patient characteristics, by physician sex

	Physician gender	
	Female physicians N=17,194	**Male physicians** N=37,317
Physician characteristics		
Physician's age, yrs	43.4 (9.3)	48.4 (11.2)
Years of experience since completion of residency, yrs	10.6 (8.0)	15.4 (10.8)
Credentials, %		
MD (allopathic)	91.7%	93.2%
DO (osteopathic)	8.3%	6.8%
Annual number of Medicare hospitalizations per physician, n	30.4	41.1
Patient characteristics		
Number of patients, n (%)	313,588	919,377
Patient's age, yrs	80.8 (8.5)	80.6 (8.5)
Female, %	62.2%	60.2%
Race, %		
White	81.1%	81.1%
Black	11.2%	10.6%
Hispanic	4.7%	5.5%
Other races	3.1%	2.8%
Medicaid status	23.5%	24.8%
Coexisting condition, %		
CHF	19.8%	20.0%
COPD	24.7%	26.4%
Diabetes	31.6%	32.4%
Renal failure	22.0%	21.6%
Neurological disorders	15.8%	16.2%
Cancer	7.2%	6.9%
Mental illness	15.9%	15.4%

All p-values <0.001 except CHF (p-value 0.03) and renal failure (p-value 0.01)

Physician sex and patient mortality

The data sample that we used to analyze 30-day mortality rates included 1,173,993 hospitalizations treated by 51,347 physicians. The overall 30-day mortality rate for the entire sample was 11.3%. Patients cared for by female physicians had lower 30-day mortality rates both before and after adjustment for patient characteristics and hospital-specific fixed effects (crude mortality rate 10.6% vs 11.5%, risk difference -1.0%, 95% confidence interval [CI]: -1.2% to -0.8%, p<0.001; adjusted mortality rate 10.9% vs 11.4%, adjusted risk difference -0.6%, 95%CI: -0.7% to -0.5%, p <0.001) (Table 3.2). Further adjusting for physician characteristics had little impact on these results (adjusted mortality rate 10.9% vs 11.4%, adjusted risk difference -0.5%, 95%CI: -0.7% to -0.4%, p-value <0.001). The parameter estimates of covariates included in our regression model for mortality and readmission analyses are shown in Tables S3.2 and S3.3 in Supplemental Materials, respectively.

Physician sex and patient readmissions

To examine sex differences in 30-day readmission rates, we used a sample consisting of 1,142,062 hospitalizations treated by 51,270 physicians. The overall 30-day readmission rate for this sample was 15.6%. Female physicians had significantly lower unadjusted readmission rates than male physicians (crude readmission rates 15.1% vs 15.8%, risk difference -0.6%, 95%CI: -0.8% to -0.5%, p<0.001) (Table 3.2). This difference in readmission rates between female and male physicians remained essentially unchanged after adjusting for hospital-specific fixed effects and patients' demographic and clinical characteristics (adjusted readmission rates 15.1% vs 15.8%, adjusted risk difference -0.6%, 95%CI: -0.8% to -0.5%, p<0.001). Similarly, accounting for physician characteristics had little, if any effect on these readmission rates (adjusted readmission rates: 15.1% for female physicians versus 15.8% for male physicians, p<0.001).

Table 3.2 Association between physician sex and 30-day patient mortality and readmissions

	Patient outcomes		Adjusted risk difference (95%CI) Female − Male	p-value
	Female physicians	Male physicians		
30-day mortality rate				
Unadjusted	**10.6%** (10.5% to 10.7%)	**11.5%** (11.5% to 11.6%)	**-1.0%** (-1.2% to -0.8%)	<0.001
Adjusted for patient characteristics and hospital fixed-effects	**10.9%** (10.7% to 11.0%)	**11.4%** (11.4% to 11.5%)	**-0.6%** (-0.7% to -0.5%)	<0.001
Adjusted for patient and physician characteristics, and hospital fixed-effects	**10.9%** (10.8% to 11.0%)	**11.4%** (11.4% to 11.5%)	**-0.5%** (-0.7% to -0.4%)	<0.001
30-day readmission rate				
Unadjusted	**15.1%** (15.0% to 15.2%)	**15.8%** (15.7% to 15.8%)	**-0.6%** (-0.8% to -0.5%)	<0.001
Adjusted for patient characteristics and hospital fixed-effects	**15.1%** (15.0% to 15.3%)	**15.8%** (15.7% to 15.8%)	**-0.6%** (-0.8% to -0.5%)	<0.001
Adjusted for patient and physician characteristics, and hospital fixed-effects	**15.1%** (15.0% to 15.3%)	**15.8%** (15.7% to 15.9%)	**-0.7%** (-0.8% to -0.5%)	<0.001

Standard errors were clustered at the physician-level.

Physician sex and patient outcomes by primary diagnoses

Female physicians had lower mortality and readmission rates across all medical conditions that we examined, although the magnitude of the difference varied by condition and was not always statistically significant. Some of the differences in mortality were small while some, such as those for sepsis, pneumonia and GI bleeding, were sizable (Table 3.3). Patterns for readmissions were similar across all conditions that we studied (Table 3.4).

Table 3.3 Association between physician sex and 30-day patient mortality, by condition

	N of hospitalizations (N of physicians)	Adjusted 30-day mortality rate (95%CI)		Adjusted risk difference (95%CI) Female − Male	p-value
		Female physicians	Male physicians		
Sepsis	102,597 (29,776)	**23.4%** (22.8% to 23.9%)	**25.3%** (25.0% to 25.6%)	**-1.9%** (-2.6% to -1.3%)	<0.001
Pneumonia	91,698 (30,398)	**10.1%** (9.7% to 10.5%)	**11.1%** (10.9% to 11.4%)	**-1.0%** (-1.5% to -0.6%)	<0.001
CHF	83,634 (28,901)	**11.7%** (11.3% to 12.2%)	**12.1%** (11.9% to 12.4%)	**-0.4%** (-0.9% to +0.2%)	0.16
COPD	66,532 (25,210)	**5.0%** (4.7% to 5.4%)	**5.4%** (5.2% to 5.6%)	**-0.4%** (-0.8% to +0.05%)	0.09
UTI	63,152 (25,348)	**5.9%** (5.5% to 6.3%)	**6.2%** (6.0% to 6.4%)	**-0.3%** (-0.7% to +0.2%)	0.23
Acute renal failure	48,938 (22,454)	**12.5%** (11.9% to 13.1%)	**13.5%** (13.2% to 13.8%)	**-1.0%** (-1.7% to -0.3%)	0.005
Arrhythmia	39,389 (20,329)	**5.0%** (4.5% to 5.4%)	**6.0%** (5.7% to 6.2%)	**-1.0%** (-1.7% to -0.4%)	0.001
GI bleeding	21,462 (14,074)	**9.3%** (8.4% to 10.2%)	**10.4%** (10.0% to 10.9%)	**-1.1%** (-2.2% to -0.1%)	0.04

Adjusted for patient and physician characteristics, and hospital-specific fixed effects. Standard errors were clustered at physician-level. CHF denotes congestive heart failure, COPD denotes chronic obstructive pulmonary disease, UTI denotes urinary tract infection, and GI bleeding denotes gastrointestinal bleeding.

Table 3.4 Association between physician sex and 30-day patient readmissions, by condition

	N of hospitalizations (N of physicians)	Adjusted 30-day readmission rate (95%CI)		Adjusted risk difference (95%CI) Female − Male	p-value
		Female physicians	Male physicians		
Sepsis	90,136 (28,254)	**15.8%** (15.3% to 16.3%)	**16.5%** (16.2% to 16.8%)	**-0.7%** (-1.2% to -0.06%)	0.03
Pneumonia	88,969 (30,006)	**14.1%** (13.6% to 14.6%)	**15.3%** (15.0% to 15.5%)	**-1.2%** (-1.8% to -0.6%)	<0.001
CHF	81,222 (28,486)	**20.2%** (19.7% to 20.8%)	**21.3%** (21.0% to 21.6%)	**-1.1%** (-1.7% to -0.4%)	0.003
COPD	66,389 (25,196)	**18.5%** (17.8% to 19.1%)	**19.3%** (18.9% to 19.6%)	**-0.8%** (-1.6% to -0.04%)	0.04
UTI	63,025 (25,345)	**14.2%** (13.7% to 14.8%)	**15.1%** (14.8% to 15.4%)	**-0.9%** (-1.5% to -0.2%)	0.01
Acute renal failure	47,865 (22,157)	**17.2%** (16.5% to 17.9%)	**17.5%** (17.1% to 17.9%)	**-0.3%** (-1.2% to +0.5%)	0.45
Arrhythmia	39,277 (20,308)	**14.3%** (13.5% to 15.0%)	**14.7%** (14.3% to 15.1%)	**-0.4%** (-1.3% to +0.5%)	0.38
GI bleeding	21,106 (13,927)	**14.2%** (13.2% to 15.3%)	**16.4%** (15.9% to 17.0%)	**-2.2%** (-3.5% to -0.9%)	0.001

Adjusted for patient and physician characteristics, and hospital-specific fixed effects. Standard errors were clustered at physician-level. CHF denotes congestive heart failure, COPD denotes chronic obstructive pulmonary disease, UTI denotes urinary tract infection, and GI bleeding denotes gastrointestinal bleeding.

Physician sex and patient outcomes by severity of illness

The association between physician sex and patient outcomes varied with patients' severity of illness (Figures 3.1 and 3.2). With regard to mortality, female physicians had significantly lower mortality rates than male physicians for patients in the highest quintile of predicted mortality (adjusted mortality rate 32.0% vs 33.3%, adjusted risk difference 1.3%, 95%CI: 0.8% to 1.8%, p<0.001). In contrast, we observed smaller associations between physician sex and patient mortality for patients in the lowest quintile of predicted mortality (1.0% vs 1.1%, adjusted risk difference 0.1%, 95%CI: -0.005% to 0.2%, p=0.06) (Table S3.4). The interaction between physician sex and predicted mortality of patients was statistically significant (p<0.001).

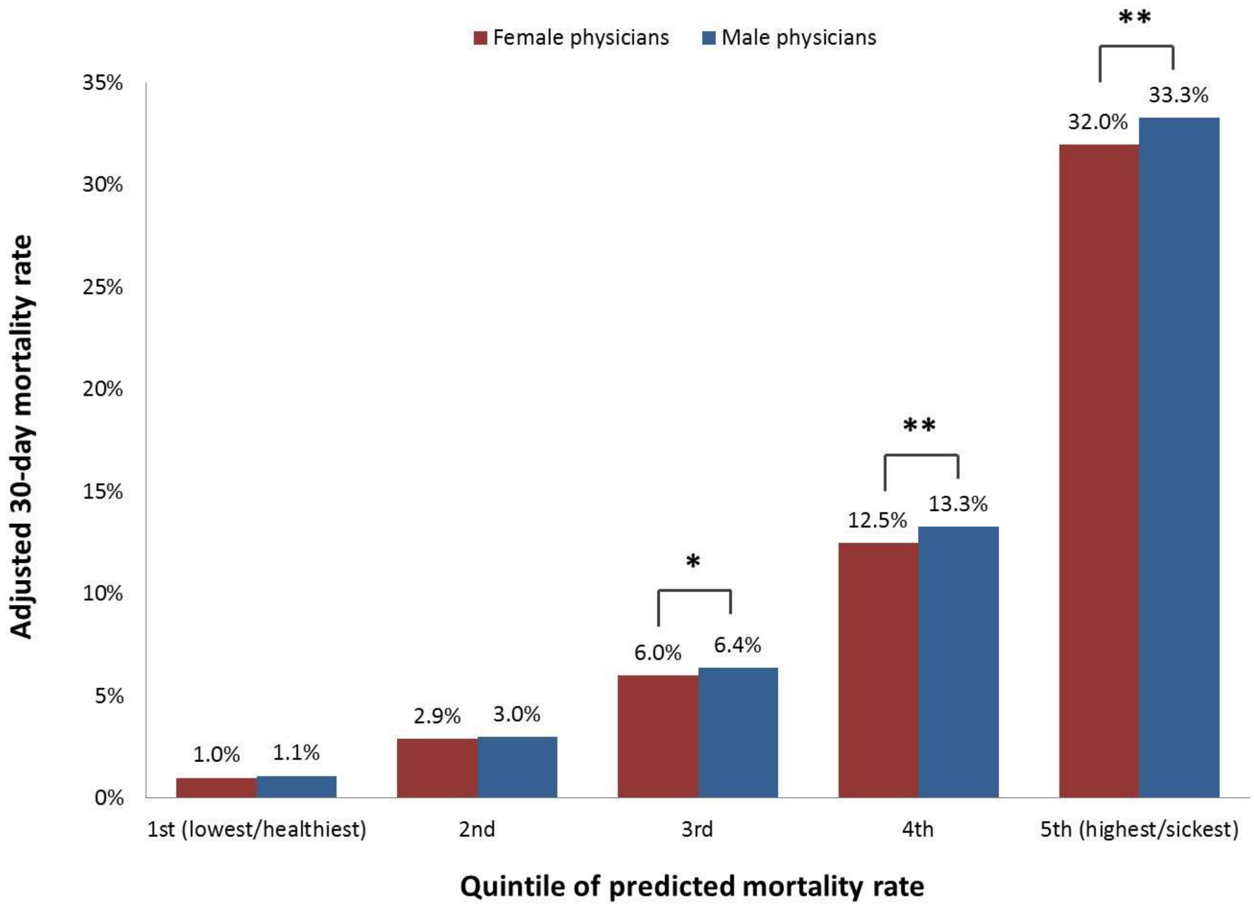

Figure 3.1 | Association between physician sex and patient 30-day mortality rate, by predicted mortality rates

*Statistically significant with p<0.01.

** Statistically significant with p<0.001.

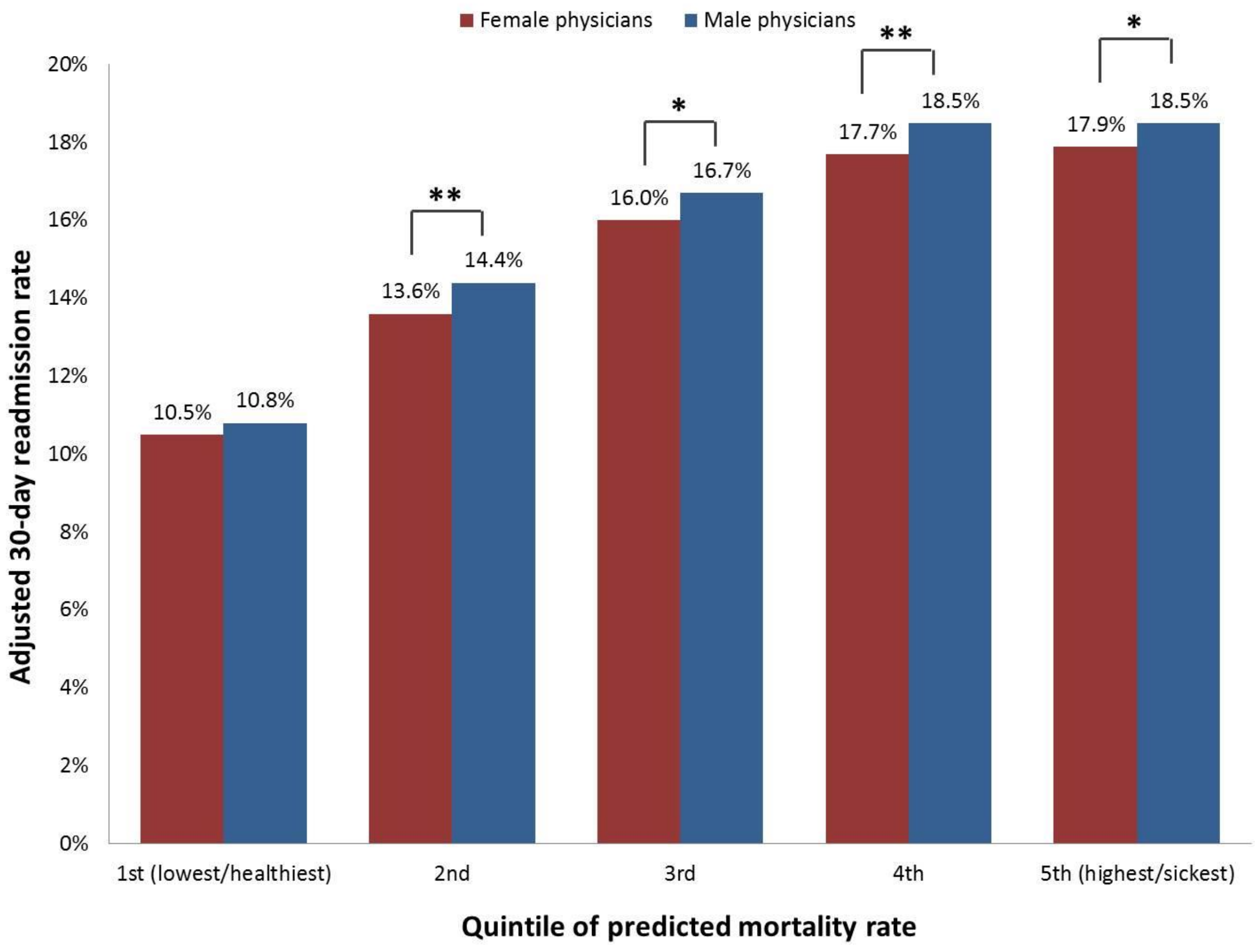

Figure 3.2 | Association between physician sex and patient 30-day readmission rate, by predicted mortality rates

*Statistically significant with p<0.01.

** Statistically significant with p<0.001.

We observed similar patterns for readmissions. The overall difference in readmission rates between female and male physicians was more prominent for the sickest patients (adjusted readmission rate 17.9% vs 18.5%, adjusted risk difference -0.6, 95%CI: -1.0% to -0.3%, p=0.001) compared to the least sick patients (10.5% vs 10.8%, adjusted risk difference -0.3%, 95%CI: -0.6% to +0.04%, p=0.09). The interaction between physician sex and patient illness severity was not statistically significant (p=0.13).

Analysis of potential mechanisms

We found that additional adjustment for utilization of care, LOS, patient volume, and discharge location did not affect our findings (Table S3.5). We also found that female physicians had lower patient mortality and readmissions across all age groups (Table S3.6).

Sensitivity analyses

We encountered similar patterns when we restricted our sample to patients treated by hospitalist physicians. For example, patients receiving care by female hospitalists had 0.4% lower adjusted 30-day mortality (95%CI: 0.2% to 0.6%, p<0.001) compared to patients receiving care by male hospitalists (Table S3.7). We observed a similar pattern when we analyzed patients admitted to hospitals without a medical ICU (Table S3.8). Our findings were unaffected by using physician age and patient age as continuous variables (with quadratic and cubic terms), estimating logistic regression models instead of linear probability models, and excluding patient race and Medicaid status from our regression models (Tables S3.9-S11).

3.4 DISCUSSION

We found that patients receiving inpatient care from female physicians had lower mortality and lower readmission rates compared to patients cared for by male physicians. This association was consistent across different conditions, and the impact of physician sex was largest among the sickest patients in our sample (i.e. those with the highest predicted mortality rates). These findings raise important questions about the etiology of differences in patient outcomes between male and female physicians that we observed, and what lessons they might offer for improving quality of care more broadly for hospitalized patients in the U.S.

Our findings demonstrate that differences in practice patterns between male and female physicians have important implications for patient outcomes. There is substantial evidence that compared to male physicians, female physicians are more likely to practice evidence-based medicine, perform as well or better on standardized examinations (Ferguson, James et al. 2002), and provide more patient-centered care (Bertakis, Helms et al. 1995, Krupat, Rosenkranz et al. 2000, Roter, Hall et al. 2002, Roter and Hall 2004). Higher performance by female physicians on these process measures of quality is important to note given evidence that patient-centered care has been associated with better patient outcomes in several studies. (Rathert, Wyrwich et al. 2013, Schoenthaler, Kalet et al. 2014). Male and female physicians may also differ in terms of their approach to solving complex problems and making decisions. Data from other industries suggest that men are more likely to make risky decisions or be less deliberate in their approach to complex problems. (Powell and Ansic 1997, Barber and Odean 2001, Charness and Gneezy 2012) If they also apply to how male and female doctors approach clinical problems and decisions, these patterns of behavior may provide a plausible causal mechanism linking physician sex with patient outcomes, especially given the larger gaps in outcomes were observed among the most complex patients.

An important concern in interpreting our results is the possibility that our findings are driven by patient selection. However, we found that nearly all observable characteristics typically associated with illness severity were relatively well balanced between female and male physicians. Even if unobserved differences in patient characteristics do exist, they would have to be substantial in order to fully explain the relatively large difference in outcomes between male and female physicians which we observed. We ran a simple simulation to show how high the prevalence of confounder would need to be to fully explain a 0.5% difference in patient mortality between male and female physicians. We assumed that an unmeasured confounder increased the risk of 30-day mortality by either 1.6% (similar to that of renal failure) or 2.1% (paralysis). These are two comorbidities that represent substantial increase in likelihood of dying. If an unmeasured confounder increases mortality by 1.6% and the prevalence of the confounder among patients of female physicians is 20%, in order for the confounder to explain the observed difference of 0.5% in mortality, its prevalence among patients of male physicians must be 51%, which is highly unlikely (most of the differences we observed in prevalence of comorbidities between patients of female and male physicians was less than 2%). Similarly, when an unmeasured confounder is associated with 2.1% higher mortality (as paralysis is) and its prevalence for patients of female physicians is 20%, for the confounder to explain the observed effect of physician sex, the prevalence of the confounder among patients of male physicians needs to be at least 44%, which we believe to also be unlikely. Moreover, we found that even for hospitalist physicians, among whom patients are plausibly more likely to be 'randomly assigned,' female physicians still had lower mortality and readmission rates than male physicians. Importantly, our analyses included hospital-specific fixed effects and thereby effectively compared male and female physicians within a given hospital.

We are aware of only one other study examining the impact of physician sex on patient outcomes.

Jerant and colleagues analyzed data from the Medical Expenditure Panel Survey (MEPS) and found no evidence of associations between physician sex and patient survival in a relatively small cohort of healthy outpatients.(Jerant, Bertakis et al. 2013) However, several analyses investigating differences in processes of care between male and female physicians have yielded results that are in line with our findings. For instance, Kim et al. and Berthold et al. both found that female physicians outperform males on process measures for patients with diabetes. (Kim, McEwen et al. 2005, Berthold, Gouni-Berthold et al. 2008) Baumahkel and colleagues reported that female physicians were more likely to prescribe beta-blockers to patients with congestive heart failure (CHF) than male physicians.(Baumhakel, Muller et al. 2009)

Our study has limitations. First, because of our study's large sample size, even a modest difference in patient mortality between female and male physicians (i.e., small effect size) could be statistically significant. However, an observed effect size of 0.5% difference in mortality is a clinically meaningful difference, especially when considering its effects across a population. For example, there has been widespread recognition that patient outcomes have improved substantially over the past decade; all-cause mortality declined from 5.0% in 2003 to 4.5% in 2013 among Medicare beneficiaries. (Krumholz, Nuti et al. 2015) This drop is seen as the results of both large national investments in quality improvement and substantial new innovations that allow us to treat patients better. And the difference in mortality rates between male and female physician in our study was of a comparable magnitude. In addition, given that there are more than 10 million Medicare hospitalizations due to medical conditions in the U.S. annually, and 75% of those patients are currently treated by male physicians, our estimates suggest that if this relationship is truly causal, approximately 37,500 fewer people would die if male physicians could achieve the same outcomes as female physicians. (Weiss AJ (Truven Health Analytics) and Elixhauser A (AHRQ) October 2014) And this is just among patients in the traditional (FFS) Medicare program. The impact would be even larger if the relationships between physician sex and patient outcomes also hold for other (non-Medicare populations). Second, our risk-adjustment methodologies cannot account for unobserved differences in patient characteristics that could influence their risk of death or readmission. Third, we cannot explain exactly why female physicians have better outcomes than male physicians from our data. Given that physician sex by itself does not affect patient outcomes, sex likely represents a marker of differences in practice patterns between male and female physicians that meaningfully affect patient outcomes. Lastly, our analysis was limited to Medicare patients with medical conditions treated by general internists. Thus, our findings may not generalize to surgical conditions, medical conditions treated by specialists, non-Medicare patients, or outpatient care. Further studies are warranted to understand whether the observed findings in our study also exist in other areas of medicine and in the outpatient setting.

In conclusion, in a national sample of hospitalized Medicare beneficiaries, we found that patients who receive care from female general internists have lower 30-day mortality and readmission rates than patients cared for by male internists. These differences in outcomes were largest for the sickest patients. Although the exact mechanism underlying these differences remains unclear, understanding why these differences in care quality exist, and what we might do to alleviate them, is critical to ensuring that all patients get high quality care.

SUPPLEMENTAL MATERIALS

Table S1.1 Adjusted association between physician spending and 30-day patient outcomes, using different method to define hospitalist physicians

	30-day mortality rate			
	Unadjusted mortality rate	**Adjusted mortality rate***	**Adjusted odds ratio (95%CI)***	**p-value**
Quintile of physician spending, adjusted				
Q1 (Lowest)	**10.8%** (10.5% to 11.0%)	**11.3%** (11.1% to 11.5%)	Reference	
Q2	**11.0%** (10.8% to 11.2%)	**10.9%** (10.7% to 11.1%)	**0.95** (0.92 to 0.996)	0.03
Q3	**11.1%** (10.9% to 11.3%)	**11.0%** (10.8% to 11.2%)	**0.97** (0.93 to 1.01)	0.11
Q4 (Highest)	**10.8%** (10.6% to 11.0%)	**10.8%** (10.6% to 11.1%)	**0.95** (0.91 to 0.99)	0.008
	30-day readmission rate			
	Unadjusted readmission rate	**Adjusted readmission rate***	**Adjusted odds ratio (95%CI)***	**p-value**
Quintile of physician spending, adjusted				
Q1 (Lowest)	**14.9%** (14.6% to 15.2%)	**15.2%** (14.9% to 15.5%)	Reference	
Q2	**15.1%** (14.8% to 15.3%)	**14.9%** (14.6% to 15.2%)	**0.99** (0.96 to 1.02)	0.53
Q3	**15.0%** (14.7% to 15.2%)	**14.9%** (14.7% to 15.2%)	**0.98** (0.95 to 1.02)	0.31
Q4 (Highest)	**15.3%** (15.0% to 15.5%)	**14.8%** (14.6% to 15.1%)	**0.99** (0.96 to 1.02)	0.39

Hospitalist physicians were defined as general internists filing at least 80% of their total E&M billings in inpatient setting.

*Adjusted for patient characteristics, physician characteristics, and hospital-specific fixed effects.

Table S1.2 Adjusted association between physician spending and 30-day patient outcomes, trimming observations with top and bottom 3 percentiles of residuals

	30-day mortality rate			
	Unadjusted mortality rate	**Adjusted mortality rate***	**Adjusted odds ratio (95%CI)***	**p-value**
Quartile of physician spending, adjusted				
Q1 (Lowest)	**10.8%** (10.5% to 11.0%)	**11.2%** (11.0% to 11.5%)	Reference	
Q2	**11.0%** (10.8% to 11.2%)	**10.9%** (10.7% to 11.1%)	**0.96** (0.92 to 1.001)	0.06
Q3	**11.0%** (10.8% to 11.2%)	**11.0%** (10.8% to 11.2%)	**0.97** (0.93 to 1.02)	0.22
Q4 (Highest)	**10.7%** (10.5% to 10.9%)	**10.7%** (10.5% to 10.9%)	**0.94** (0.90 to 0.98)	0.003
	30-day readmission rate			
	Unadjusted readmission rate	**Adjusted readmission rate***	**Adjusted odds ratio (95%CI)***	**p-value**
Quartile of physician spending, adjusted				
Q1 (Lowest)	**14.8%** (14.5% to 15.0%)	**15.0%** (14.8% to 15.3%)	Reference	
Q2	**14.9%** (14.7% to 15.2%)	**15.1%** (14.8% to 15.4%)	**1.00** (0.97 to 1.04)	0.77
Q3	**14.8%** (14.5% to 15.0%)	**14.7%** (14.5% to 15.0%)	**0.97** (0.94 to 1.006)	0.11
Q4 (Highest)	**15.1%** (14.9% to 15.4%)	**15.0%** (14.8% to 15.3%)	**1.00** (0.97 to 1.03)	0.93

*Adjusted for patient characteristics, physician characteristics, and hospital-specific fixed effects.

Table S1.3 Adjusted association between physician spending and 30-day patient outcomes, trimming observations with top and bottom 7 percentiles of residuals

	30-day mortality rate			
	Unadjusted mortality rate	**Adjusted mortality rate***	**Adjusted odds ratio (95%CI)***	**p-value**
Quartile of physician spending, adjusted				
Q1 (Lowest)	**10.9%** (10.7% to 11.1%)	**11.3%** (11.1% to 11.6%)	Reference	
Q2	**10.9%** (10.6% to 11.1%)	**10.8%** (10.6% to 11.1%)	**0.94** (0.90 to 0.99)	0.01
Q3	**11.0%** (10.7% to 11.2%)	**10.9%** (10.7% to 11.1%)	**0.95** (0.91 to 0.99)	0.02
Q4 (Highest)	**10.6%** (10.4% to 10.9%)	**10.7%** (10.4% to 10.9%)	**0.92** (0.89 to 0.96)	<0.001
	30-day readmission rate			
	Unadjusted readmission rate	**Adjusted readmission rate***	**Adjusted odds ratio (95%CI)***	**p-value**
Quartile of physician spending, adjusted				
Q1 (Lowest)	**14.9%** (14.6% to 15.1%)	**15.1%** (14.8% to 15.4%)	Reference	
Q2	**14.9%** (14.7% to 15.2%)	**15.1%** (14.8% to 15.3%)	**0.99** (0.97 to 1.03)	0.79
Q3	**14.8%** (14.6% to 15.1%)	**14.8%** (14.5% to 15.1%)	**0.97** (0.94 to 1.01)	0.10
Q4 (Highest)	**15.0%** (14.7% to 15.3%)	**14.9%** (14.6% to 15.2%)	**0.98** (0.95 to 1.02)	0.29

*Adjusted for patient characteristics, physician characteristics, and hospital-specific fixed effects.

Table S1.4 Adjusted association between physician spending and 30-day patient outcomes, including patients who were transferred out

	30-day mortality rate			
	Unadjusted mortality rate	**Adjusted mortality rate***	**Adjusted odds ratio (95%CI)***	**p-value**
Quintile of physician spending, adjusted				
Q1 (Lowest)	**10.7%** (10.5% to 10.9%)	**11.1%** (10.9% to 11.4%)	Reference	
Q2	**11.0%** (10.8% to 11.2%)	**10.9%** (10.7% to 11.1%)	**0.97** (0.93 to 1.01)	0.17
Q3	**11.0%** (10.8% to 11.2%)	**11.0%** (10.7% to 11.2%)	**0.98** (0.94 to 1.02)	0.29
Q4 (Highest)	**10.7%** (10.4% to 10.9%)	**10.7%** (10.5% to 11.0%)	**0.95** (0.91 to 0.99)	0.02
	30-day readmission rate			
	Unadjusted readmission rate	**Adjusted readmission rate***	**Adjusted odds ratio (95%CI)***	**p-value**
Quintile of physician spending, adjusted				
Q1 (Lowest)	**14.9%** (14.6% to 15.1%)	**15.2%** (14.9% to 15.5%)	Reference	
Q2	**14.8%** (14.6% to 15.1%)	**14.9%** (14.6% to 15.2%)	**0.97** (0.94 to 1.007)	0.12
Q3	**14.9%** (14.7% to 15.2%)	**14.9%** (14.7% to 15.2%)	**0.98** (0.95 to 1.01)	0.16
Q4 (Highest)	**15.0%** (14.7% to 15.2%)	**14.8%** (14.6% to 15.1%)	**0.97** (0.94 to 1.001)	0.06

All spending of the transfer-in (recipient) hospitalizations was assigned to the initial hospitalization.

*Adjusted for patient characteristics, physician characteristics, and hospital-specific fixed effects.

Table S1.5 Adjusted association between physician spending and 30-day patient outcomes, after excluding patients with cancer

	30-day mortality rate			
	Unadjusted mortality rate	**Adjusted mortality rate***	**Adjusted odds ratio (95%CI)***	**p-value**
Quintile of physician spending, adjusted				
Q1 (Lowest)	**9.9%** (9.7% to 10.2%)	**10.3%** (10.1% to 10.6%)	Reference	
Q2	**10.1%** (9.9% to 10.3%)	**10.0%** (9.7% to 10.2%)	**0.95** (0.91 to 0.99)	0.03
Q3	**10.1%** (9.9% to 10.3%)	**10.1%** (9.9% to 10.3%)	**0.97** (0.93 to 1.01)	0.18
Q4 (Highest)	**9.8%** (9.6% to 10.0%)	**9.9%** (9.7% to 10.2%)	**0.95** (0.91 to 0.99)	0.02
	30-day readmission rate			
	Unadjusted readmission rate	**Adjusted readmission rate***	**Adjusted odds ratio (95%CI)***	**p-value**
Quintile of physician spending, adjusted				
Q1 (Lowest)	14.6% (14.4% to 14.9%)	15.0% (14.7% to 15.3%)	Reference	
Q2	14.6% (14.3% to 14.9%)	14.7% (14.4% to 15.0%)	0.97 (0.94 to 1.01)	0.13
Q3	14.6% (14.3% to 14.8%)	14.6% (14.3% to 14.8%)	0.97 (0.93 to 0.999)	0.04
Q4 (Highest)	14.7% (14.4% to 14.9%)	14.6% (14.3% to 14.8%)	0.97 (0.93 to 0.998)	0.04

*Adjusted for patient characteristics, physician characteristics, and hospital-specific fixed effects.

Table S1.6 Adjusted association between physician spending and 30-day patient outcomes, among general internists

	30-day mortality rate			
	Unadjusted mortality rate	**Adjusted mortality rate***	**Adjusted odds ratio (95%CI)***	**p-value**
Quartile of physician spending, adjusted				
Q1 (Lowest)	**10.8%** (10.6% to 11.0%)	**11.2%** (11.0% to 11.4%)	Reference	
Q2	**11.1%** (10.9% to 11.3%)	**11.0%** (10.9% to 11.2%)	**0.98** (0.94 to 1.01)	0.20
Q3	**11.1%** (10.9% to 11.2%)	**11.1%** (10.9% to 11.2%)	**0.98** (0.95 to 1.02)	0.31
Q4 (Highest)	**11.0%** (10.8% to 11.2%)	**10.9%** (10.7% to 11.1%)	**0.96** (0.93 to 0.99)	0.04
	30-day readmission rate			
	Unadjusted readmission rate	**Adjusted readmission rate***	**Adjusted odds ratio (95%CI)***	**p-value**
Quartile of physician spending, adjusted				
Q1 (Lowest)	**15.2%** (14.9% to 15.4%)	**15.4%** (15.2% to 15.7%)	Reference	
Q2	**15.4%** (15.2% to 15.6%)	**15.5%** (15.3% to 15.7%)	**1.01** (0.98 to 1.03)	0.60
Q3	**15.4%** (15.2% to 15.6%)	**15.4%** (15.2% to 15.6%)	**1.00** (0.97 to 1.03)	0.96
Q4 (Highest)	**15.9%** (15.7% to 16.1%)	**15.6%** (15.4% to 15.9%)	**1.02** (0.99 to 1.04)	0.21

*Adjusted for patient characteristics, physician characteristics, and hospital-specific fixed effects.

Table S2.1 ICD-9 (International Classification of Diseases, 9th Edition) codes

Condition	ICD-9 codes
Sepsis	0031, 0202, 0223, 0362, 0380, 0381, 03810, 03811, 03812, 03819, 0382, 0383, 03840, 03841, 03842, 03843, 03844, 03849, 0388, 0389, 0545, 449, 77181, 7907, 99591, 99592
Pneumonia	00322, 0203, 0204, 0205, 0212, 0221, 0310, 0391, 0521, 0551, 0730, 0830, 1124, 1140, 1144, 1145, 11505, 11515 11595, 1304, 1363, 4800, 4801, 4802, 4803, 4808, 4809, 481, 4820, 4821, 4822, 4823, 48230, 48231, 48232, 48239, 4824, 48240, 48241, 48242, 48249, 4828, 48281, 48282, 48283, 48284, 48289, 4829, 483, 4830, 4831, 4838, 4841, 4843, 4845, 4846, 4847, 4848, 485, 486, 5130, 5171
Congestive heart failure	39891, 4280, 4281, 42820, 42821, 42822, 42823, 42830, 42831, 42832, 42833, 42840, 42841, 42842, 42843, 4289

Table S2.2 Association between physicians' age and 30-day patient mortality

Physician age (yr)	N of hospitalizations (N of physicians)	Unadjusted 30-day patient mortality rate (95%CI)	Risk-adjusted 30-day patient mortality rate (95%CI)*	Adjusted risk difference*	
				Risk difference (95%CI)	p-value
Physician age as categorical variable					
<40	263,105 (9,860)	**10.5%** (10.4% to 10.7%)	**10.7%** (10.6% to 10.9%)	Reference	
40-49	272,485 (7,675)	**11.1%** (10.9% to 11.2%)	**11.0%** (10.9% to 11.2%)	**+0.3%** (+0.1% to +0.5%)	0.002
50-59	110,383 (3,063)	**11.5%** (11.3% to 11.8%)	**11.4%** (11.2% to 11.6%)	**+0.6%** (+0.4% to +0.9%)	<0.001
≥60	32,459 (1,098)	**12.5%** (12.0% to 13.0%)	**12.0%** (11.5% to 12.5%)	**+1.3%** (+0.8% to +1.8%)	<0.001
Physician age as continuous variable					
For every 10 yrs increase	678,432 (21,696)			**+0.4%** (+0.3% to +0.5%)	<0.001

*Adjusted for patient characteristics, physician characteristics, and hospital-specific fixed effects.

Table S2.3 Association between physicians' years in practice and 30-day patient mortality, using different definitions for identifying hospitalists

Years in practice	N of hospitalizations (N of physicians)	Unadjusted 30-day patient mortality rate (95%CI)	Risk-adjusted 30-day patient mortality rate (95%CI)*	Adjusted risk difference*	
				Risk difference (95%CI)	p-value
Hospitalists defined as physicians with ≥80% of E&M claims in inpatient setting					
1-4	113,950 (5,806)	**10.2%** (10.0% to 10.4%)	**10.7%** (10.6% to 10.9%)	Reference	
5-9	132,040 (4,879)	**11.1%** (10.9% to 11.2%)	**11.1%** (10.9% to 11.3%)	**+0.4%** (+0.1% to +0.6%)	0.005
10-14	80,989 (2,917)	**11.5%** (11.3% to 11.7%)	**11.3%** (11.0% to 11.5%)	**+0.5%** (+0.2% to +0.8%)	0.001
≥15	88,598 (3,488)	**12.4%** (12.2% to 12.6%)	**11.8%** (11.5% to 12.0%)	**+1.0%** (+0.7% to +1.4%)	<0.001
Hospitalists defined as physicians with ≥95% of E&M claims in inpatient setting					
1-4	107,462 (5,533)	**10.2%** (10.0% to 10.4%)	**10.6%** (10.4% to 10.8%)	Reference	
5-9	120,289 (4,191)	**10.9%** (10.7% to 11.0%)	**10.9%** (10.7% to 11.1%)	**+0.3%** (+0.02% to +0.5%)	0.04
10-14	70,756 (2,545)	**11.2%** (11.0% to 11.5%)	**11.0%** (10.8% to 11.3%)	**+0.4%** (+0.1% to +0.7%)	0.01
≥15	73,142 (2,886)	**12.2%** (12.0% to 12.5%)	**11.6%** (11.4% to 11.9%)	**+1.0%** (+0.7% to +1.4%)	<0.001

*Adjusted for patient characteristics, physician characteristics, and hospital-specific fixed effects.

Table S2.4 Analysis of potential mechanism, additional adjustment for length of stay

Years in practice	N of hospitalizations (N of physicians)	Unadjusted 30-day patient mortality rate (95%CI)	Risk-adjusted 30-day patient mortality rate (95%CI)*	Adjusted risk difference*	
				Risk difference (95%CI)	p-value
1-4	108,891 (4,884)	**10.2%** (10.0% to 10.4%)	**10.6%** (10.4% to 10.8%)	Reference	
5-9	124,459 (4,392)	**10.9%** (10.7% to 11.0%)	**10.9%** (10.7% to 11.1%)	**+0.3%** (+0.04% to +0.5%)	0.047
10-14	74,826 (2,591)	**11.2%** (11.0% to 11.5%)	**11.0%** (10.8% to 11.3%)	**+0.4%** (+0.1% to +0.7%)	0.01
≥15	77,897 (2,873)	**12.1%** (11.9% to 12.3%)	**11.6%** (11.3% to 11.9%)	**+1.0%** (+0.6% to +1.3%)	<0.001

Length of stay was used as a continuous variable with quadratic and cubic terms to allow for potential non-linear relationship.

*Adjusted for patient characteristics, physician characteristics, and hospital-specific fixed effects.

Table S2.5 Association between physicians' years in practice and 30-day patient mortality, among general internists

Years in practice	N of hospitalizations (N of physicians)	Unadjusted 30-day patient mortality rate (95%CI)	Risk-adjusted 30-day patient mortality rate (95%CI)*	Adjusted risk difference*	
				Risk difference (95%CI)	p-value
Years in practice as categorical variable					
1-4	123,830 (7,189)	**10.2%** (10.0% to 10.3%)	**10.7%** (10.5% to 10.9%)	Reference	
5-9	159,977 (7,175)	**11.4%** (11.2% to 11.6%)	**11.3%** (11.2% to 11.5%)	**+0.6%** (+0.4% to +0.8%)	<0.001
10-14	123,910 (5,820)	**11.6%** (11.4% to 11.8%)	**11.5%** (11.3% to 11.7%)	**+0.8%** (+0.5% to +1.0%)	<0.001
≥15	243,516 (13,641)	**11.8%** (11.6% to 11.9%)	**11.6%** (11.4% to 11.7%)	**+0.8%** (+0.6% to +1.1%)	<0.001
Years in practice as continuous variable					
For every 10 yrs increase	651,233 (33,825)			**+0.2%** (+0.1% to +0.3%)	<0.001

*Adjusted for patient characteristics, physician characteristics, and hospital-specific fixed effects.

Table S2.6 Association between physicians' years in practice and 30-day patient mortality, by primary diagnosis

	Years in practice	N of hospitalizations (N of physicians)	Risk-adjusted 30-day mortality rate (95%CI)*	Adjusted risk difference*	
				Risk difference (95%CI)	p-value
Sepsis (n=36,746)	1-4	10,419 (3,340)	**22.0%** (21.0% to 22.6%)	Reference	
	5-9	11,743 (3,198)	**22.4%** (21.7% to 23.2%)	**+0.6%** (-0.5% to +1.7%)	0.31
	10-14	7,049 (1,931)	**23.4%** (22.4% to 24.4%)	**+1.6%** (+0.2% to +2.9%)	0.02
	≥15	7,535 (2,026)	**24.0%** (22.9% to 25.0%)	**+2.2%** (+0.8% to +3.5%)	0.002
Pneumonia (n=28,773)	1-4	7,943 (3,179)	**10.1%** (9.5% to 10.8%)	Reference	
	5-9	9,034 (3,110)	**9.9%** (9.2% to 10.5%)	**-0.3%** (-1.2% to +0.7%)	0.54
	10-14	5,693 (1,842)	**10.0%** (9.2% to 10.8%)	**-0.2%** (-1.3% to +0.9%)	0.76
	≥15	6,103 (1,905)	**10.1%** (9.3% to 10.9%)	**-0.03%** (-1.2% to +1.1%)	0.96
CHF (n=27,337)	1-4	7,722 (3,108)	**11.1%** (10.4% to 11.9%)	Reference	
	5-9	8,722 (3,035)	**11.7%** (11.0% to 12.3%)	**+0.5%** (-0.5% to +1.6%)	0.31
	10-14	5,349 (1,770)	**11.5%** (10.5% to 12.4%)	**+0.3%** (-0.9% to +1.6%)	0.60
	≥15	5,544 (1,842)	**13.1%** (12.1% to 14.1%)	**+2.0%** (+0.7% to +3.3%)	0.003

The interaction test was not statistically significant (p=0.50).

*Adjusted for patient characteristics, physician characteristics, and hospital-specific fixed effects.

Table S2.7 Association between physicians' years in practice and 30-day patient mortality, stratified by severity of illness

Severity of illness	Years in practice	N of hospitalizations (N of physicians)	Risk-adjusted 30-day mortality rate (95%CI)*	Adjusted risk difference*	
				Risk difference (95%CI)	p-value
Low (Healthiest)	1-4	36,533 (4,794)	**1.5%** (1.3% to 1.6%)	Reference	
	5-9	41,025 (3,997)	**1.7%** (1.6% to 1.9%)	**+0.3%** (+0.08% to +0.4%)	0.01
	10-14	24,334 (2,397)	**1.6%** (1.4% to 1.7%)	**+0.1%** (-0.1% to +0.3%)	0.40
	≥15	25,236 (2,567)	**1.7%** (1.5% to 1.9%)	**+0.3%** (+0.03% to +0.5%)	0.03
Medium	1-4	36,340 (4,730)	**6.0%** (5.8% to 6.3%)	Reference	
	5-9	41,537 (3,997)	**6.1%** (5.9% to 6.3%)	**+0.08%** (-0.3% to +0.4%)	0.68
	10-14	25,061 (2,446)	**6.3%** (5.9% to 6.6%)	**+0.2%** (-0.2% to +0.7%)	0.31
	≥15	25,906 (2,653)	**6.8%** (6.5% to 7.2%)	**+0.8%** (+0.4% to +1.3%)	<0.001
High (Sickest)	1-4	36,044 (4,654)	**24.4%** (23.8% to 24.8%)	Reference	
	5-9	41,932 (4,090)	**24.6%** (24.2% to 25.0%)	**+0.3%** (-0.4% to +0.9%)	0.39
	10-14	25,448 (2,431)	**25.0%** (24.4% to 25.6%)	**-0.7%** (-0.1% to +1.4%)	0.10
	≥15	26,758 (2,684)	**25.7%** (25.1% to 26.3%)	**+1.3%** (+0.5% to +2.2%)	0.001

The interaction test was not statistically significant (p=0.17).

*Adjusted for patient characteristics, physician characteristics, and hospital-specific fixed effects.

Table S2.8 Association between physicians' years in practice and 30-day patient mortality, using logistic regression model

Years in practice	N of hospitalizations (N of physicians)	Unadjusted 30-day patient mortality rate (95%CI)	Risk-adjusted 30-day patient mortality rate (95%CI)*	Adjusted odds ratio*	
				Odds ratio (95%CI)	p-value
Years in practice as categorical variable					
1-4	108,918 (5,066)	**10.2%** (10.0% to 10.4%)	**10.7%** (10.5% to 10.9%)	Reference	
5-9	124,495 (4,128)	**10.9%** (10.7% to 11.0%)	**10.9%** (10.8% to 11.1%)	**1.03** (0.99 to 1.06)	0.12
10-14	74,844 (2,543)	**11.2%** (11.0% to 11.5%)	**11.1%** (10.8% to 11.3%)	**1.04** (1.003 to 1.09)	0.03
≥15	77,902 (2,913)	**12.1%** (11.9% to 12.3%)	**11.6%** (11.3% to 11.8%)	**1.11** (1.07 to 1.16)	<0.001
Years in practice as continuous variable					
For every 10 yrs increase	386,159 (14,650)			**1.006** (1.004 to 1.008)	<0.001

*Adjusted for patient characteristics, physician characteristics, and hospital-specific fixed effects.

Table S3.1 ICD-9 (International Classification of Diseases, 9th Edition) codes

Condition	ICD-9 codes
Sepsis	0031, 0202, 0223, 0362, 0380, 0381, 03810, 03811, 03812, 03819, 0382, 0383, 03840, 03841, 03842, 03843, 03844, 03849, 0388, 0389, 0545, 449, 77181, 7907, 99591, 99592
Pneumonia	00322, 0203, 0204, 0205, 0212, 0221, 0310, 0391, 0521, 0551, 0730, 0830, 1124, 1140, 1144, 1145, 11505, 11515 11595, 1304, 1363, 4800, 4801, 4802, 4803, 4808, 4809, 481, 4820, 4821, 4822, 4823, 48230, 48231, 48232, 48239, 4824, 48240, 48241, 48242, 48249, 4828, 48281, 48282, 48283, 48284, 48289, 4829, 483, 4830, 4831, 4838, 4841, 4843, 4845, 4846, 4847, 4848, 485, 486, 5130, 5171
Congestive heart failure	39891, 4280, 4281, 42820, 42821, 42822, 42823, 42830, 42831, 42832, 42833, 42840, 42841, 42842, 42843, 4289
Chronic obstructive pulmonary disease (COPD)	490, 4910, 4911, 4912, 49120, 49121, 49122, 4918, 4919, 4920, 4928, 494, 4940, 4941, 496
Urinary tract infection	03284, 59000, 59001, 59010, 59011, 5902, 5903, 59080, 59081, 5909, 5950, 5951, 5952, 5953, 5954, 59581, 59582, 59589, 5959, 5970, 59780, 59781, 59789, 59800, 59801, 5990
Acute renal failure	5845, 5846, 5847, 5848, 5849, 586
Arrhythmia	4270, 4271, 4272, 42731, 42732, 42760, 42761, 42769, 42781, 42789, 4279, 7850, 7851
Gastrointestinal bleeding	4560, 45620, 5307, 53082, 53100, 53101, 53120, 53121, 53140, 53141, 53160, 53161, 53200, 53201, 53220, 53221, 53240, 53241, 53260, 53261, 53300, 53301, 53320, 53321, 53340, 53341, 53360, 53361, 53400, 53401, 53420, 53421, 53440, 53441, 53460, 53461, 5693, 5780, 5781, 5789

Table S3.2 Parameter estimates covariates in our regression model for mortality analysis

Variables	Adjusted risk difference (95%CI)	p-value
Physician age		
<35	Reference	
35-39	+0.3% (+0.1% to +0.5%)	0.01
40-44	+0.5% (+0.2% to +0.7%)	<0.001
45-49	+0.6% (+0.3% to +0.8%)	<0.001
50-54	+0.8% (+0.5% to +1.0%)	<0.001
55-59	+0.7% (+0.4% to +1.0%)	<0.001
60-64	+0.6% (+0.3% to +1.0%)	<0.001
65-69	+0.7% (+0.3% to +1.1%)	0.001
≥70	+0.3% (-0.3% to +0.8%)	0.36
Credential		
DO	Reference	
MD	-0.04% (-0.3% to +0.2%)	0.78
Patient age		
65-69	Reference	
70-74	+1.4% (+1.2% to +1.5%)	<0.001
75-79	+2.8% (+2.6% to +3.0%)	<0.001
80-84	+4.6% (+4.4% to +4.8%)	<0.001
85-89	+6.9% (+6.7% to +7.1%)	<0.001
90-94	+10.1% (+9.9% to +10.3%)	<0.001
≥95	+15.1% (+14.7% to +15.5%)	<0.001
Patient sex		
Male	Reference	
Female	-1.1% (-1.2% to -0.9%)	<0.001
Patient race		
White	Reference	
Black	-1.7% (-1.9% to -1.5%)	<0.001
Hispanic	-1.1% (-1.4% to -0.8%)	<0.001
Other races	-0.9% (-1.2% to -0.5%)	<0.001
Medicaid status	-0.4% (-0.5% to -0.2%)	<0.001

Table S3.2 (Continued) Parameter estimates covariates in our regression model for mortality analysis

Variables	Adjusted risk difference(95%CI)*	p-value
Elixhauser comorbidity index		
Congestive heart failure	+2.5% (+2.3% to +2.7%)	<0.001
Valvular disease	-0.9% (-1.1% to -0.7%)	<0.001
Pulmonary circulation disorders	+0.8% (+0.4% to +1.1%)	<0.001
Peripheral vascular disorders	+0.7% (+0.5% to +0.9%)	<0.001
Hypertension	-2.8% (-3.0% to -2.7%)	<0.001
Paralysis	+2.1% (+1.8% to +2.5%)	<0.001
Other neurological disorders	+2.1% (+2.0% to +2.3%)	<0.001
COPD	-0.7% (-0.9% to -0.6%)	<0.001
Diabetes uncomplicated	-0.6% (-0.7% to -0.5%)	<0.001
Diabetes complicated	-1.4% (-1.6% to -1.2%)	<0.001
Hypothyroidism	-1.1% (-1.2% to -1.0%)	<0.001
Renal failure	+1.6% (+1.4% to +1.7%)	<0.001
Liver disease	+3.7% (+3.3% to +4.2%)	<0.001
Peptic ulcer excluding bleeding	-2.1% (-5.0% to +0.8%)	0.16
AIDS	-2.0% (-5.4% to +1.4%)	0.24
Lymphoma	+4.7% (+4.1% to 5.3%)	<0.001
Metastatic cancer	+22.2% (+21.7% to +22.7%)	<0.001
Solid tumor without metastasis	+7.4% (+7.1% to +7.8%)	<0.001
Rheumatoid arthritis/collagen vascular diseases	-0.3% (-0.5% to -0.004%)	0.047
Coagulopathy	+2.8% (+2.5% to +3.1%)	<0.001
Obesity	-1.4% (-1.6% to -1.3%)	<0.001
Weight loss	+7.4% (+7.1% to +7.7%)	<0.001
Fluid and electrolyte disorders	+3.1% (+3.0% to +3.3%)	<0.001
Blood loss anemia	-2.0% (-2.5% to -1.4%)	<0.001
Deficiency anemia	-1.4% (-1.6% to -1.3%)	<0.001
Psychosis	-0.5% (-0.8% to -0.3%)	<0.001
Depression	-0.6% (-0.7% to 0.4%)	<0.001

*The reference group is patients without indicated coexisting conditions.

Table S3.3 Parameter estimates covariates in our regression model for readmission analysis

Variables	Adjusted risk difference (95%CI)	p-value
Physician age		
<35	Reference	
35-39	+0.1% (-0.1% to +0.4%)	0.32
40-44	0% (-0.3% to +0.3%)	0.89
45-49	+0.3% (0% to 0.6%)	0.08
50-54	0% (-0.3% to +0.4%)	0.79
55-59	-0.1% (-0.4% to +0.3%)	0.71
60-64	-0.2% (-0.6% to +0.2%)	0.29
65-69	-0.4% (-0.9% to +0.1%)	0.08
≥70	+0.6% (-0.1% to +1.2%)	0.10
Credential		
DO	Reference	
MD	-0.1% (-0.4% to +0.1%)	0.35
Patient age		
65-69	Reference	
70-74	+0.1% (-0.2% to +0.3%)	0.61
75-79	+0.1% (-0.2% to +0.3%)	0.64
80-84	+0.3% (+0.1% to +0.6%)	0.01
85-89	+0.2% (-0.1% to +0.5%)	0.14
90-94	-0.2% (-0.5% to 0%)	0.09
≥95	-1.4% (-1.8% to -1.0%)	<0.001
Patient sex		
Male	Reference	
Female	-0.5% (-0.6% to -0.3%)	<0.001
Patient race		
White	Reference	
Black	+1.4% (+1.1% to +1.7%)	<0.001
Hispanic	-0.3% (-0.7% to +0.1%)	0.14
Other races	-0.4% (-0.9% to 0%)	0.07
Medicaid status	+2.0% (+1.8% to +2.2%)	<0.001

Table S3.3 (Continued) Parameter estimates covariates in our regression model for readmission analysis

Variables	Adjusted risk difference (95%CI)*	p-value
Elixhauser comorbidity index		
Congestive heart failure	+4.1% (+3.9% to +4.3%)	<0.001
Valvular disease	+0.8% (+0.5% to +1.1%)	<0.001
Pulmonary circulation disorders	+1.5% (+1.1% to +1.9%)	<0.001
Peripheral vascular disorders	+0.9% (+0.6% to +1.1%)	<0.001
Hypertension	-0.7% (-0.9% to -0.5%)	<0.001
Paralysis	+1.6% (+1.2% to +2.1%)	<0.001
Other neurological disorders	+0.4% (+0.2% to +0.6%)	<0.001
COPD	+2.8% (+2.7% to +3.0%)	<0.001
Diabetes uncomplicated	+1.5% (+1.3% to +1.6%)	<0.001
Diabetes complicated	+2.5% (+2.1% to +2.8%)	<0.001
Hypothyroidism	+0.5% (+0.3% to +0.6%)	<0.001
Renal failure	+3.2% (+3.1% to +3.4%)	<0.001
Liver disease	+3.2% (+2.7% to +3.8%)	<0.001
Peptic ulcer excluding bleeding	+0.3% (-3.4% to +4.0%)	0.87
AIDS	+6.8% (+1.6% to +12.1%)	0.01
Lymphoma	5.1% (4.4% to 5.8%)	<0.001
Metastatic cancer	4.9% (4.4% to 5.4%)	<0.001
Solid tumor without metastasis	+2.6% (+2.2% to +3.0%)	<0.001
Rheumatoid arthritis/collagen vascular diseases	+1.2% (+0.9% to +1.6%)	<0.001
Coagulopathy	+1.2% (+0.9% to +1.5%)	<0.001
Obesity	-0.3% (-0.6% to -0.1%)	0.02
Weight loss	+1.7% (+1.4% to +2.0%)	<0.001
Fluid and electrolyte disorders	+1.2% (+1.0% to +1.3%)	<0.001
Blood loss anemia	+2.6% (+1.9% to +3.3%)	<0.001
Deficiency anemia	+2.4% (+2.3% to +2.6%)	<0.001
Psychosis	+1.3% (+0.9% to +1.6%)	<0.001
Depression	+0.6% (+0.4% to +0.8%)	<0.001

*The reference group is patients without indicated coexisting conditions.

Table S3.4 Association between physician sex and patient outcomes, by predicted mortality rates

		N of hospitalizations (N of physicians)	Adjusted 30-day patient outcomes*		Adjusted risk difference* (95%CI) Female − Male	p-value
			Female physicians	**Male physicians**		
Mortality rate						
Quintile of predicted mortality	**1st (Lowest)**	234,881 (50,600)	**1.0%** (0.9% to 1.1%)	**1.1%** (1.1% to 1.2%)	**-0.1%** (-0.2% to +0.005%)	0.06
	2nd	236,641 (40,894)	**2.9%** (2.7% to 3.0%)	**3.0%** (3.0% to 3.1%)	**-0.2%** (-0.3% to +0.02%)	0.053
	3rd	234,733 (40,758)	**6.0%** (6.3% to 6.5%)	**6.4%** (6.3% to 6.5%)	**-0.4%** (-0.6% to -0.1%)	0.002
	4th	234,624 (40,518)	**12.5%** (12.3% to 12.8%)	**13.3%** (13.2% to 13.5%)	**-0.8%** (-1.1% to -0.4%)	<0.001
	5th (Highest)	235,055 (39,737)	**32.0%** (31.6% to 32.4%)	**33.3%** (33.0% to 33.5%)	**-1.3%** (-1.8% to -0.8%)	<0.001
Readmission rate						
Quintile of predicted mortality	**1st (Lowest)**	228,406 (40,496)	**10.5%** (10.3% to 10.8%)	**10.8%** (10.6% to 10.9%)	**-0.3%** (-0.6% to +0.04%)	0.09
	2nd	228,301 (40,835)	**13.6%** (13.3% to 13.9%)	**14.4%** (14.3% to 14.6%)	**-0.8%** (-1.2% to -0.5%)	<0.001
	3rd	228,290 (40,658)	**16.0%** (15.7% to 16.3%)	**16.7%** (16.5% to 16.9%)	**-0.7%** (-1.0% to -0.3%)	0.001
	4th	228,403 (40,337)	**17.7%** (17.4% to 18.0%)	**18.5%** (18.3% to 18.7%)	**-0.8%** (-1.2% to -0.5%)	<0.001
	5th (Highest)	228,562 (39,696)	**17.9%** (17.5% to 18.2%)	**18.5%** (18.3% to 18.7%)	**-0.6%** (-1.0% to -0.3%)	0.001

Standard errors were clustered at physician-level. The p-for-interaction was <0.001 for the mortality analysis, and 0.13 for the readmission analysis.

*Adjusted for patient characteristics, physician characteristics, and hospital-specific fixed effects.

.

Table S3.5 Association between physician sex and patient outcomes, sensitivity analysis

	N of hospitalizations (N of physicians)	Patients' adjusted outcomes*		Adjusted risk difference* (95%CI) Female − Male	p-value
		Female physicians	**Male physicians**		
30-day mortality rate					
Additional adjustment for LOS	1,173,712 (51,346)	**10.9%** (10.8% to 11.0%)	**11.4%** (11.4% to 11.5%)	**-0.5%** (-0.7% to -0.4%)	<0.001
Additional adjustment for patient volume	1,173,993 (51,347)	**10.8%** (10.6% to 10.9%)	**11.5%** (11.4% to 11.5%)	**-0.7%** (-0.9% to -0.6%)	<0.001
Additional adjustment for utilization of care†	1,173,993 (51,347)	**10.9%** (10.8% to 11.0%)	**11.4%** (11.4% to 11.5%)	**-0.5%** (-0.7% to -0.4%)	<0.001
Additional adjustment for discharge location	1,173,993 (51,347)	**11.1%** (11.0% to 11.2%)	**11.4%** (11.3% to 11.4%)	**-0.3%** (-0.4% to -0.2%)	<0.001
30-day readmission rate					
Additional adjustment for LOS	1,064,739 (50,858)	**16.2%** (16.1% to 16.4%)	**16.9%** (16.9% to 17.0%)	**-0.7%** (-0.9% to -0.5%)	<0.001
Additional adjustment for patient volume	1,141,582 (50,861)	**15.2%** (15.1% to 15.4%)	**15.8%** (15.7% to 15.8%)	**-0.5%** (-0.7% to -0.3%)	<0.001
Additional adjustment for utilization of care†	1,065,052 (50,861)	**16.2%** (16.1% to 16.4%)	**16.9%** (16.8% to 17.0%)	**-0.7%** (-0.9% to -0.5%)	<0.001
Additional adjustment for discharge location	1,065,052 (50,861)	**16.2%** (16.1% to 16.4%)	**16.9%** (16.9% to 17.0%)	**-0.7%** (-0.9% to -0.5%)	<0.001
Excluding patients who died out of hospital within 30 days of discharge	1,065,052 (50,861)	**16.2%** (16.1% to 16.4%)	**16.9%** (16.9% to 17.0%)	**-0.7%** (-0.9% to -0.6%)	<0.001

Adjusted for patient and physician characteristics, and hospital-specific fixed effects. LOS denotes length of stay. Standard errors were clustered at physician-level.
*Adjusted for patient characteristics, physician characteristics, and hospital-specific fixed effects.
†Defined as total Part B spending per hospitalization, standardized for geographical differences using the CMS method.

Table S3.6 Association between physician sex and patient outcomes, stratified by physician age

	N of hospitalizations (N of physicians)	Patients' adjusted outcomes*		Adjusted risk difference* (95%CI) Female − Male	p-value
		Female physicians	Male physicians		
30-day mortality rate					
Physician age < 40	317,057 (15,673)	**11.0%** (10.9% to 11.2%)	**11.6%** (11.5% to 11.7%)	**-0.5%** (-0.7% to -0.2%)	<0.001
Physician age 40-49	427,434 (16,291)	**11.0%** (10.8% to 11.2%)	**11.6%** (11.4% to 11.7%)	**-0.6%** (-0.8% to -0.3%)	<0.001
Physician age ≥ 50	429,502 (19,383)	**11.0%** (10.8% to 11.3%)	**11.6%** (11.5% to 11.7%)	**-0.6%** (-0.9% to -0.3%)	<0.001
30-day readmission rate					
Physician age < 40	311,735 (15,624)	**15.1%** (14.9% to 15.3%)	**15.7%** (15.5% to 15.9%)	**-0.6%** (-0.9% to -0.3%)	<0.001
Physician age 40-49	415,821 (16,270)	**15.2%** (15.0% to 15.5%)	**15.8%** (15.7% to 15.9%)	**-0.6%** (-0.8% to -0.3%)	<0.001
Physician age ≥ 50	414,506 (19,376)	**14.9%** (14.6% to 15.2%)	**15.9%** (15.7% to 16.0%)	**-0.9%** (-1.3% to -0.6%)	<0.001

Standard errors were clustered at physician-level.
*Adjusted for patient characteristics, physician characteristics, and hospital-specific fixed effects.

Table S3.7 Association between physician sex and patient outcomes, among hospitalist physicians only

	N of hospitalizations (N of physicians)	Patient outcomes		Adjusted risk difference (95%CI) Female − Male	p-value
		Female physicians	Male physicians		
30-day mortality rate					
Unadjusted	725,821 (25,550)	**10.5%** (10.4% to 10.6%)	**11.4%** (11.3% to 11.4%)	**-0.9%** (-1.1% to -0.7%)	<0.001
Adjusted for patient characteristics and hospital fixed-effects	725,821 (25,550)	**10.7%** (10.6% to 10.9%)	**11.2%** (11.2% to 11.3%)	**-0.5%** (-0.7% to -0.3%)	<0.001
Adjusted for patient and physician characteristics, and hospital fixed-effects	685,441 (23,793)	**10.8%** (10.7% to 10.9%)	**11.2%** (11.1% to 11.3%)	**-0.4%** (-0.6% to -0.2%)	<0.001
30-day readmission rate					
Unadjusted	704,174 (23,205)	**14.8%** (14.6% to 14.9%)	**15.3%** (15.2% to 15.4%)	**-0.5%** (-0.7% to -0.3%)	<0.001
Adjusted for patient characteristics and hospital fixed-effects	704,174 (23,205)	**14.6%** (14.5% to 14.8%)	**15.3%** (15.2% to 15.4%)	**-0.7%** (-0.9% to -0.5%)	<0.001
Adjusted for patient and physician characteristics, and hospital fixed-effects	665,285 (21,689)	**14.7%** (14.5% to 14.8%)	**15.3%** (15.2% to 15.4%)	**-0.7%** (-0.9% to -0.5%)	<0.001

Standard errors were clustered at physician-level.

Table S3.8 Association between physician sex and 30-day patient mortality, hospitals without medical ICU

	N of hospitalizations (N of physicians)	Patient outcomes		Adjusted risk difference (95%CI) Female − Male	p-value
		Female physicians	Male physicians		
30-day mortality rate					
Unadjusted	162,841 (10,979)	**10.9%** (10.6% to 11.2%)	**11.7%** (11.6% to 11.9%)	**-0.9%** (-1.3% to -0.4%)	<0.001
Adjusted for patient characteristics and hospital fixed-effects	162,841 (10,979)	**11.2%** (10.9% to 11.5%)	**11.7%** (11.5% to 11.8%)	**-0.5%** (-0.9% to -0.06%)	0.02
Adjusted for patient and physician characteristics, and hospital fixed-effects	154,689 (10,349)	**11.2%** (10.8% to 11.6%)	**11.7%** (11.5% to 11.8%)	**-0.5%** (-0.9% to -0.05%)	0.03
30-day readmission rate					
Unadjusted	156,093 (10,331)	**15.2%** (14.9% to 15.6%)	**16.2%** (16.0% to 16.4%)	**-1.0%** (-1.5% to -0.5%)	<0.001
Adjusted for patient characteristics and hospital fixed-effects	156,093 (10,331)	**15.7%** (15.3% to 16.1%)	**16.1%** (15.9% to 16.3%)	**-0.4%** (-0.9% to +0.04%)	0.08
Adjusted for patient and physician characteristics, and hospital fixed-effects	148,231 (9,732)	**15.7%** (15.2% to 16.1%)	**16.1%** (15.9% to 16.3%)	**-0.5%** (-0.9% to +0.03%)	0.07

Standard errors were clustered at physician-level.

Table S3.9 Association between physician sex and patient outcomes, using age as a continuous variable†

	N of hospitalizations (N of physicians)	Patient outcomes		Adjusted risk difference (95%CI) Female − Male	p-value
		Female physicians	Male physicians		
30-day mortality rate					
Adjusted for patient characteristics and hospital fixed-effects	1,232,965 (54,511)	**10.9%** (10.7% to 11.0%)	**11.4%** (11.4% to 11.5%)	**-0.6%** (-0.7% to -0.5%)	<0.001
Adjusted for patient and physician characteristics, and hospital fixed-effects	1,173,993 (51,347)	**10.9%** (10.8% to 11.0%)	**11.4%** (11.4% to 11.5%)	**-0.5%** (-0.7% to -0.4%)	<0.001
30-day readmission rate					
Adjusted for patient characteristics and hospital fixed-effects	1,199,438 (54,526)	**15.1%** (15.0% to 15.3%)	**15.8%** (15.7% to 15.8%)	**-0.6%** (-0.8% to -0.5%)	<0.001
Adjusted for patient and physician characteristics, and hospital fixed-effects	1,142,062 (51,270)	**15.1%** (15.0% to 15.3%)	**15.8%** (15.7% to 15.9%)	**-0.7%** (-0.8% to -0.5%)	<0.001

Standard errors were clustered at physician-level.

†Both physician age and patient age were used as continuous variables, with quadratic and cubic terms to allow for non-linear relationship.

Table S3.10 Association between physician sex and patient outcomes, using logistic regression models

	N of hospitalizations (N of physicians)	Patient outcomes		Adjusted odds ratio (95%CI) Female − Male	p-value
		Female physicians	Male physicians		
30-day mortality rate					
Unadjusted	1,232,965 (54,511)	**10.6%** (10.5% to 10.7%)	**11.5%** (11.5% to 11.6%)	**0.90** (0.89 to 0.92)	<0.001
Adjusted for patient characteristics and hospital fixed-effects	1,231,955 (54,399)	**10.9%** (10.7% to 11.0%)	**11.5%** (11.4% to 11.5%)	**0.93** (0.91 to 0.94)	<0.001
Adjusted for patient and physician characteristics, and hospital fixed-effects	1,172,906 (51,243)	**10.9%** (10.8% to 11.0%)	**11.4%** (11.4% to 11.5%)	**0.94** (0.92 to 0.95)	<0.001
30-day readmission rate					
Unadjusted	1,199,438 (54,426)	**15.1%** (15.0% to 15.2%)	**15.8%** (15.7% to 15.8%)	**0.95** (0.94 to 0.96)	<0.001
Adjusted for patient characteristics and hospital fixed-effects	1,198,708 (54,278)	**15.2%** (15.0% to 15.3%)	**15.8%** (15.7% to 15.8%)	**0.95** (0.94 to 0.97)	<0.001
Adjusted for patient and physician characteristics, and hospital fixed-effects	1,141,371 (51,131)	**15.1%** (15.0% to 15.3%)	**15.8%** (15.7% to 15.9%)	**0.95** (0.94 to 0.96)	<0.001

Standard errors were clustered at physician-level.

Table S3.11 Association between physician sex and patient outcomes, excluding patient race and an indicator for Medicaid eligibility

	N of hospitalizations (N of physicians)	Patients' adjusted outcomes*		Adjusted risk difference* (95%CI) Female − Male	p-value
		Female physicians	Male physicians		
30-day mortality rate					
Excluding patient race and Medicaid indicator from the model	1,173,993 (51,347)	**10.9%** (10.8% to 11.0%)	**11.4%** (11.4% to 11.5%)	**-0.5%** (-0.7% to -0.4%)	<0.001
30-day readmission rate					
Excluding patient race and Medicaid indicator from the model	1,142,062 (51,270)	**15.1%** (15.0% to 15.3%)	**15.8%** (15.7% to 15.9%)	**-0.7%** (-0.8% to -0.5%)	<0.001

Standard errors were clustered at physician-level.
*Adjusted for patient characteristics, physician characteristics, and hospital-specific fixed effects.

REFERENCES

Adams, J. L., E. A. McGlynn, J. W. Thomas and A. Mehrotra (2010). "Incorporating statistical uncertainty in the use of physician cost profiles." BMC Health Services Research **10**: 57.

Adams, J. L., A. Mehrotra, J. W. Thomas and E. A. McGlynn (2010). "Physician cost profiling--reliability and risk of misclassification." N Engl J Med **362**(11): 1014-1021.

Allison, P. D. (2008). Convergence failures in logistic regression. SAS Global Forum, Citeseer.

Andersen, M. R. and N. Urban (1997). "Physician gender and screening: do patient differences account for differences in mammography use?" Women Health **26**(1): 29-39.

Argency for Healthcare Research and Quality. "HCUPnet ", from http://hcupnet.ahrq.gov.

Association of American Medical Colleges U.S. Medical School Applicants and Students 1982-1983 to 2011-2012. Washington, DC.

Ayanian, J. Z., P. J. Hauptman, E. Guadagnoli, E. M. Antman, C. L. Pashos and B. J. McNeil (1994). "Knowledge and practices of generalist and specialist physicians regarding drug therapy for acute myocardial infarction." N Engl J Med **331**(17): 1136-1142.

Aylin, P., A. Bottle and A. Majeed (2007). "Use of administrative data or clinical databases as predictors of risk of death in hospital: comparison of models." BMJ **334**(7602): 1044.

Barber, B. M. and T. Odean (2001). "Boys will be boys: Gender, overconfidence, and common stock investment." Quarterly journal of Economics: 261-292.

Barnato, A. E., M. B. Herndon, D. L. Anthony, P. M. Gallagher, J. S. Skinner, J. P. Bynum and E. S. Fisher (2007). "Are regional variations in end-of-life care intensity explained by patient preferences?: A Study of the US Medicare Population." Medical care **45**(5): 386-393.

Bauchner, H., P. B. Fontanarosa and A. E. Thompson (2015). "Professionalism, governance, and self-regulation of medicine." JAMA **313**(18): 1831-1836.

Baumhakel, M., U. Muller and M. Bohm (2009). "Influence of gender of physicians and patients on guideline-recommended treatment of chronic heart failure in a cross-sectional study." Eur J Heart Fail **11**(3): 299-303.

Bennett, K. J. and J. C. Probst (2015). "Thirty-Day Readmission Rates Among Dual-Eligible Beneficiaries." The Journal of Rural Health **32**(2): 188-195.

Bertakis, K. D., L. J. Helms, E. J. Callahan, R. Azari and J. A. Robbins (1995). "The influence of gender on physician practice style." Med Care **33**(4): 407-416.

Berthold, H. K., I. Gouni-Berthold, K. P. Bestehorn, M. Bohm and W. Krone (2008). "Physician gender is associated with the quality of type 2 diabetes care." J Intern Med **264**(4): 340-350.

Blanc, P. D., L. Trupin, G. Earnest, M. San Pedro, P. P. Katz, E. H. Yelin and M. D. Eisner (2003). "Effects of physician-related factors on adult asthma care, health status, and quality of life." Am J Med **114**(7): 581-587.

Burns, L. R. and D. R. Wholey (1991). "The effects of patient, hospital, and physician characteristics on length of stay and mortality." Med Care **29**(3): 251-271.

Centers for Medicare & Medicaid Services (2013). Standardization Methodology For Allowed Amount, Centers for Medicare & Medicaid Services.

Charness, G. and U. Gneezy (2012). "Strong evidence for gender differences in risk taking." Journal of Economic Behavior & Organization **83**(1): 50-58.

Chen, L. M., A. K. Jha, S. Guterman, A. B. Ridgway, E. J. Orav and A. M. Epstein (2010). "Hospital cost of care, quality of care, and readmission rates: penny wise and pound foolish?" Arch Intern Med **170**(4): 340-346.

Choudhry, N. K., R. H. Fletcher and S. B. Soumerai (2005). "Systematic review: the relationship between clinical experience and quality of health care." Ann Intern Med **142**(4): 260-273.

Czaja, R., S. L. McFall, R. B. Warnecke, L. Ford and A. D. Kaluzny (1994). "Preferences of community physicians for cancer screening guidelines." Ann Intern Med **120**(7): 602-608.

Davidson, W., D. W. Molloy and M. Bedard (1995). "Physician characteristics and prescribing for elderly people in New Brunswick: relation to patient outcomes." CMAJ **152**(8): 1227-1234.

Diggle, P., P. Heagerty, K.-Y. Liang and S. Zeger (2002). Analysis of longitudinal data, Oxford University Press.

Drye, E. E., S. L. Normand, Y. Wang, J. S. Ross, G. C. Schreiner, L. Han, M. Rapp and H. M. Krumholz (2012). "Comparison of hospital risk-standardized mortality rates calculated by using in-hospital and 30-day models: an observational study with implications for hospital profiling." Ann Intern Med **156**(1 Pt 1): 19-26.

Elixhauser, A., C. Steiner, D. R. Harris and R. M. Coffey (1998). "Comorbidity measures for use with administrative data." Medical care **36**(1): 8-27.

Epstein, A. J., S. K. Srinivas, S. Nicholson, J. Herrin and D. A. Asch (2013). "Association between physicians' experience after training and maternal obstetrical outcomes: cohort study." BMJ **346**: f1596.

Fehrenbach, S. N., D. S. Budnitz, J. A. Gazmararian and H. M. Krumholz (2001). "Physician characteristics and the initiation of beta-adrenergic blocking agent therapy after acute myocardial infarction in a managed care population." Am J Manag Care **7**(7): 717-723.

Ferguson, E., D. James and L. Madeley (2002). "Factors associated with success in medical school: systematic review of the literature." BMJ **324**(7343): 952-957.

Ferguson, T. B., Jr., E. D. Peterson, L. P. Coombs, M. C. Eiken, M. L. Carey, F. L. Grover, E. R. DeLong, S. Society of Thoracic and D. the National Cardiac (2003). "Use of continuous quality improvement to increase use of process measures in patients undergoing coronary artery bypass graft surgery: a randomized controlled trial." JAMA **290**(1): 49-56.

Fisher, E. S., D. E. Wennberg, T. A. Stukel, D. J. Gottlieb, F. L. Lucas and E. L. Pinder (2003). "The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care." Ann Intern Med **138**(4): 273-287.

Fisher, E. S., D. E. Wennberg, T. A. Stukel, D. J. Gottlieb, F. L. Lucas and E. L. Pinder (2003). "The implications of regional variations in Medicare spending. Part 2: health outcomes and satisfaction with care." Ann Intern Med **138**(4): 288-298.

Franks, P. and K. D. Bertakis (2003). "Physician gender, patient gender, and primary care." J Womens Health (Larchmt) **12**(1): 73-80.

Franks, P. and C. M. Clancy (1993). "Physician gender bias in clinical decisionmaking: screening for cancer in primary care." Med Care **31**(3): 213-218.

Freiman, M. P. (1985). "The rate of adoption of new procedures among physicians. The impact of specialty and practice characteristics." Medical care **23**(8): 939-945.

Gemson, D. H., J. Colombotos, J. Elinson, E. J. Fordyce, M. Hynes and R. Stoneburner (1991). "Acquired immunodeficiency syndrome prevention. Knowledge, attitudes, and practices of primary care physicians." Arch Intern Med **151**(6): 1102-1108.

Goldstein, H., W. Browne and J. Rasbash (2002). "Partitioning variation in multilevel models." Understanding Statistics: Statistical Issues in Psychology, Education, and the Social Sciences **1**(4): 223-231.

Gordon, H. S., P. R. Nowlin, D. Maynard, M. L. Berbaum and A. Deswal (2010). "Mortality after hospitalization for heart failure in blacks compared to whites." Am J Cardiol **105**(5): 694-700.

Grover, F. L., R. R. Johnson, A. L. Shroyer, G. Marshall and K. E. Hammermeister (1994). "The Veterans Affairs Continuous Improvement in Cardiac Surgery Study." Ann Thorac Surg **58**(6): 1845-1851.

Halm, E. A., C. Lee and M. R. Chassin (2002). "Is volume related to outcome in health care? A systematic review and methodologic critique of the literature." Ann Intern Med **137**(6): 511-520.

Hannan, E. L., H. Kilburn, Jr., J. F. O'Donnell, G. Lukacik and E. P. Shields (1990). "Adult open heart surgery in New York State. An analysis of risk factors and hospital mortality rates." JAMA **264**(21): 2768-2774.

Hannan, E. L., H. Kilburn, Jr., M. Racz, E. Shields and M. R. Chassin (1994). "Improving the outcomes of coronary artery bypass surgery in New York State." JAMA **271**(10): 761-766.

Hartz, A. J., E. M. Kuhn and J. Pulido (1999). "Prestige of training programs and experience of bypass surgeons as factors in adjusted patient mortality rates." Med Care **37**(1): 93-103.

Hlatky, M. A., H. Cotugno, C. O'Connor, D. B. Mark, D. B. Pryor and R. M. Califf (1988). "Adoption of thrombolytic therapy in the management of acute myocardial infarction." Am J Cardiol **61**(8): 510-514.

Hussey, P. S., S. Wertheimer and A. Mehrotra (2013). "The association between health care quality and cost: a systematic review." Ann Intern Med **158**(1): 27-34.

Institute of Medicine (2013). Variation in Health Care Spending: Target Decision Making, Not Geography. Washington, DC, The National Academies Press.

Jagsi, R., K. A. Griffith, A. Stewart, D. Sambuco, R. DeCastro and P. A. Ubel (2012). "Gender differences in the salaries of physician researchers." JAMA **307**(22): 2410-2417.

Jena, A. B., D. Khullar, O. Ho, A. R. Olenski and D. M. Blumenthal (2015). "Sex Differences in Academic Rank in US Medical Schools in 2014." JAMA **314**(11): 1149-1158.

Jerant, A., K. D. Bertakis, J. J. Fenton and P. Franks (2013). "Gender of physician as the usual source of care and patient health care utilization and mortality." J Am Board Fam Med **26**(2): 138-148.

Jha, A. K., E. J. Orav, A. Dobson, R. A. Book and A. M. Epstein (2009). "Measuring efficiency: the association of hospital costs and quality of care." Health Aff (Millwood) **28**(3): 897-906.

Jha, A. K., M. G. Shlipak, W. Hosmer, C. D. Frances and W. S. Browner (2001). "Racial differences in mortality among men hospitalized in the Veterans Affairs health care system." JAMA **285**(3): 297-303.

Jolly, S., K. A. Griffith, R. DeCastro, A. Stewart, P. Ubel and R. Jagsi (2014). "Gender differences in time spent on parenting and domestic responsibilities by high-achieving young physician-researchers." Annals of internal medicine **160**(5): 344-353.

Joynt, K. E. and A. K. Jha (2012). "Thirty-day readmissions—truth and consequences." N Engl J Med **366**(15): 1366-1369.

Joynt, K. E., E. J. Orav and A. K. Jha (2011). "Thirty-day readmission rates for Medicare beneficiaries by race and site of care." JAMA **305**(7): 675-681.

Joynt, K. E., E. J. Orav and A. K. Jha (2013). "Physician volume, specialty, and outcomes of care for patients with heart failure." Circ Heart Fail **6**(5): 890-897.

Kaiser Family Foundation (2016). Distribution of physicians by gender. Menlo Park, CA, Kaiser Family Foundation.

Katon, W., C. M. Rutter, E. Lin, G. Simon, M. Von Korff, T. Bush, E. Walker and E. Ludman (2000). "Are there detectable differences in quality of care or outcome of depression across primary care providers?" Med Care **38**(6): 552-561.

Kim, C., L. N. McEwen, R. B. Gerzoff, D. G. Marrero, C. M. Mangione, J. V. Selby and W. H. Herman (2005). "Is physician gender associated with the quality of diabetes care?" Diabetes Care **28**(7): 1594-1598.

Krumholz, H. M., S. V. Nuti, N. S. Downing, S. L. Normand and Y. Wang (2015). "Mortality, Hospitalizations, and Expenditures for the Medicare Population Aged 65 Years or Older, 1999-2013." JAMA **314**(4): 355-365.

Krupat, E., S. L. Rosenkranz, C. M. Yeager, K. Barnard, S. M. Putnam and T. S. Inui (2000). "The practice orientations of physicians and patients: the effect of doctor-patient congruence on satisfaction." Patient Educ Couns **39**(1): 49-59.

Kuo, Y. F., G. Sharma, J. L. Freeman and J. S. Goodwin (2009). "Growth in the care of older patients by hospitalists in the United States." N Engl J Med **360**(11): 1102-1112.

Lagu, T., M. B. Rothberg, B. H. Nathanson, P. S. Pekow, J. S. Steingrub and P. K. Lindenauer (2011). "The relationship between hospital spending and mortality in patients with sepsis." Arch Intern Med **171**(4): 292-299.

Lurie, N., J. Slater, P. McGovern, J. Ekstrum, L. Quam and K. Margolis (1993). "Preventive care for women. Does the sex of the physician matter?" N Engl J Med **329**(7): 478-482.

McWilliams, J. M., B. E. Landon, M. E. Chernew and A. M. Zaslavsky (2014). "Changes in Patients' Experiences in Medicare Accountable Care Organizations." N Engl J Med **371**(18): 1715-1724.

Mehrotra, A., J. L. Adams, J. W. Thomas and E. A. McGlynn (2010). "Cost profiles: should the focus be on individual physicians or physician groups?" Health affairs **29**(8): 1532-1538.

Mehrotra, A., J. L. Adams, J. W. Thomas and E. A. McGlynn (2010). "The effect of different attribution rules on individual physician cost profiles." Ann Intern Med **152**(10): 649-654.

Meskauskas, J. A. and G. D. Webster (1975). "The American Board of Internal Medicine recertification examination: process and results." Ann Intern Med **82**(4): 577-581.

Moscucci, M., D. Share, D. Smith, M. J. O'Donnell, A. Riba, R. McNamara, T. Lalonde, A. C. Defranco, K. Patel, E. Kline Rogers, C. D'Haem, M. Karve and K. A. Eagle (2005). "Relationship between operator volume and adverse outcome in contemporary percutaneous coronary intervention practice: an analysis of a quality-controlled multicenter percutaneous coronary intervention clinical database." J Am Coll Cardiol **46**(4): 625-632.

Nallamothu, B. K., H. S. Gurm, H. H. Ting, P. P. Goodney, M. A. Rogers, J. P. Curtis, J. B. Dimick, E. R.

Bates, H. M. Krumholz and J. D. Birkmeyer (2011). "Operator experience and carotid stenting outcomes in Medicare beneficiaries." JAMA **306**(12): 1338-1343.

Newhouse, J. P. and A. M. Garber (2013). "Geographic variation in health care spending in the United States: insights from an Institute of Medicine report." JAMA **310**(12): 1227-1228.

Newhouse, J. P. and A. M. Garber (2013). "Geographic variation in Medicare services." N Engl J Med **368**(16): 1465-1468.

Nonnemaker, L. (2000). "Women physicians in academic medicine: new insights from cohort studies." N Engl J Med **342**(6): 399-405.

Norcini, J. J., H. R. Kimball and R. S. Lipner (2000). "Certification and specialization: do they matter in the outcome of acute myocardial infarction?" Acad Med **75**(12): 1193-1198.

Norcini, J. J., R. S. Lipner, J. A. Benson, Jr. and G. D. Webster (1985). "An analysis of the knowledge base of practicing internists as measured by the 1980 recertification examination." Ann Intern Med **102**(3): 385-389.

O'Connor, G. T., S. K. Plume, E. M. Olmstead, L. H. Coffin, J. R. Morton, C. T. Maloney, E. R. Nowicki, J. F. Tryzelaar, F. Hernandez, L. Adrian and et al. (1991). "A regional prospective study of in-hospital mortality associated with coronary artery bypass grafting. The Northern New England Cardiovascular Disease Study Group." JAMA **266**(6): 803-809.

O'Connor, G. T., S. K. Plume, E. M. Olmstead, J. R. Morton, C. T. Maloney, W. C. Nugent, F. Hernandez, Jr., R. Clough, B. J. Leavitt, L. H. Coffin, C. A. Marrin, D. Wennberg, J. D. Birkmeyer, D. C. Charlesworth, D. J. Malenka, H. B. Quinton and J. F. Kasper (1996). "A regional intervention to improve the hospital mortality associated with coronary artery bypass graft surgery. The Northern New England Cardiovascular Disease Study Group." JAMA **275**(11): 841-846.

O'Neill, L., D. J. Lanska and A. Hartz (2000). "Surgeon characteristics associated with mortality and morbidity following carotid endarterectomy." Neurology **55**(6): 773-781.

Olmsted, M. G., E. Geisen, J. Murphy, D. Bell, M. Morley and M. Stanley (2015). Methodology: U.S. News & World Report Best Hospitals 2015-16, RTI International.

Pippins, J. R., G. M. Fitzmaurice and J. S. Haas (2007). "Hospital characteristics and racial disparities in hospital mortality from common medical conditions." J Natl Med Assoc **99**(9): 1030-1036.

Powell, M. and D. Ansic (1997). "Gender differences in risk behaviour in financial decision-making: An experimental analysis." Journal of economic psychology **18**(6): 605-628.

Rahman, M., D. Tyler, K. S. Thomas, D. C. Grabowski and V. Mor (2015). "Higher Medicare SNF care utilization by dual-eligible beneficiaries: can Medicaid long-term care policies be the answer?" Health Serv Res **50**(1): 161-179.

Ramsey, P. G., J. D. Carline, T. S. Inui, E. B. Larson, J. P. LoGerfo, J. J. Norcini and M. D. Wenrich (1991). "Changes over time in the knowledge base of practicing internists." JAMA **266**(8): 1103-1107.

Rathert, C., M. D. Wyrwich and S. A. Boren (2013). "Patient-centered care and outcomes: a systematic review of the literature." Med Care Res Rev **70**(4): 351-379.

Ray, W. A., C. F. Federspiel and W. Schaffner (1976). "Prescribing of chloramphenicol in ambulatory practice. An epidemiologic study among Tennessee Medicaid recipients." Ann Intern Med **84**(3): 266-270.

Romley, J. A., A. B. Jena and D. P. Goldman (2011). "Hospital spending and inpatient mortality: evidence from California: an observational study." Ann Intern Med **154**(3): 160-167.

Ross, J. S., S. L. Normand, Y. Wang, D. T. Ko, J. Chen, E. E. Drye, P. S. Keenan, J. H. Lichtman, H. Bueno, G. C. Schreiner and H. M. Krumholz (2010). "Hospital volume and 30-day mortality for three common medical conditions." N Engl J Med **362**(12): 1110-1118.

Roter, D. L. and J. A. Hall (2004). "Physician gender and patient-centered communication: a critical review of empirical research." Annu Rev Public Health **25**: 497-519.

Roter, D. L., J. A. Hall and Y. Aoki (2002). "Physician gender effects in medical communication: a meta-analytic review." JAMA **288**(6): 756-764.

Salem-Schatz, S. R., J. Avorn and S. B. Soumerai (1990). "Influence of clinical knowledge, organizational context, and practice style on transfusion decision making. Implications for practice change strategies." JAMA **264**(4): 476-483.

Schoenthaler, A., A. Kalet, J. Nicholson and M. Lipkin, Jr. (2014). "Does improving patient-practitioner communication improve clinical outcomes in patients with cardiovascular diseases? A systematic review of the evidence." Patient Educ Couns **96**(1): 3-12.

Schwartz, J. S., C. E. Lewis, C. Clancy, M. S. Kinosian, M. H. Radany and J. P. Koplan (1991). "Internists' practices in health promotion and disease prevention. A survey." Ann Intern Med **114**(1): 46-53.

Seabury, S. A., A. Chandra and A. B. Jena (2013). "Trends in the earnings of male and female health care professionals in the United States, 1987 to 2010." JAMA Intern Med **173**(18): 1748-1750.

Shapiro, J. A. (1989). "General practitioners' attitudes towards AIDS and their perceived information needs." BMJ **298**(6687): 1563-1566.

Southern, W. N., E. Y. Bellin and J. H. Arnsten (2011). "Longer lengths of stay and higher risk of mortality among inpatients of physicians with more years in practice." Am J Med **124**(9): 868-874.

Streja, D. A. and S. W. Rabkin (1999). "Factors associated with implementation of preventive care measures in patients with diabetes mellitus." Arch Intern Med **159**(3): 294-302.

Stukel, T. A., E. S. Fisher, D. A. Alter, A. Guttmann, D. T. Ko, K. Fung, W. P. Wodchis, N. N. Baxter, C. C. Earle and D. S. Lee (2012). "Association of hospital spending intensity with mortality and readmission rates in Ontario hospitals." JAMA **307**(10): 1037-1045.

Sutherland, J. M., E. S. Fisher and J. S. Skinner (2009). "Getting past denial—the high cost of health care in the United States." N Engl J Med **361**(13): 1227-1230.

Teirstein, P. S. and E. J. Topol (2015). "The role of maintenance of certification programs in governance and professionalism." JAMA **313**(18): 1809-1810.

Tsai, T. C., K. E. Joynt, E. J. Orav, A. A. Gawande and A. K. Jha (2013). "Variation in surgical-readmission rates and quality of hospital care." N Engl J Med **369**(12): 1134-1142.

Weiss AJ (Truven Health Analytics) and Elixhauser A (AHRQ) (October 2014). Overview of Hospital Stays in the United States, 2012. HCUP Statistical Brief #180. Rockville, MD, Agency for Healthcare Research and Quality.

Wennberg, J. and Gittelsohn (1973). "Small area variations in health care delivery." Science **182**(4117): 1102-1108.

Williams, R. (2012). "Using the margins command to estimate and interpret adjusted predictions and marginal effects." Stata Journal **12**(2): 308.

Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data, MIT press.

Yasaitis, L., E. S. Fisher, J. S. Skinner and A. Chandra (2009). "Hospital quality and intensity of spending: is there an association?" Health Aff (Millwood) **28**(4): w566-w572.

Zaslavsky, A. M. (2001). "Statistical issues in reporting quality data: small samples and casemix variation." Int J Qual Health Care **13**(6): 481-488.

Zhang, B., A. A. Wright, H. A. Huskamp, M. E. Nilsson, M. L. Maciejewski, C. C. Earle, S. D. Block, P. K. Maciejewski and H. G. Prigerson (2009). "Health care costs in the last week of life: associations with end-of-life conversations." Arch Intern Med **169**(5): 480-488.

Zhang, Y., S. H. Baik, A. M. Fendrick and K. Baicker (2012). "Comparing local and regional variation in health care spending." N Engl J Med **367**(18): 1724-1731.

Zuckerman, S., T. Waidmann, R. Berenson and J. Hadley (2010). "Clarifying sources of geographic differences in Medicare spending." N Engl J Med **363**(1): 54-62.