

Original Investigation

Evaluation of Surveillance Bias and the Validity of the Venous Thromboembolism Quality Measure

Karl Y. Bilimoria, MD, MS; Jeanette Chung, PhD; Mila H. Ju, MD; Elliott R. Haut, MD; David J. Bentrem, MD, MS; Clifford Y. Ko, MD, MS; David W. Baker, MD, MPH

IMPORTANCE Postoperative venous thromboembolism (VTE) rates are widely reported quality metrics soon to be used in pay-for-performance programs. Surveillance bias occurs when some clinicians use imaging studies to detect VTE more frequently than other clinicians. Because they look more, they find more VTE events, paradoxically worsening their hospital's VTE quality measure performance. A surveillance bias may influence VTE measurement if (1) greater hospital VTE prophylaxis adherence fails to result in lower measured VTE rates, (2) hospitals with characteristics suggestive of higher quality (eg, more accreditations) have greater VTE prophylaxis adherence rates but worse VTE event rates, and (3) higher hospital VTE imaging utilization use rates are associated with higher measured VTE event rates.

OBJECTIVE To examine whether a surveillance bias influences the validity of reported VTE rates.

DESIGN, SETTING, AND PARTICIPANTS 2010 Hospital Compare and American Hospital Association data from 2838 hospitals were merged. Next, 2009-2010 Medicare claims data for 954 926 surgical patient discharges from 2786 hospitals who were undergoing 1 of 11 major operations were used to calculate VTE imaging (duplex ultrasonography, chest computed tomography/magnetic resonance imaging, and ventilation-perfusion scans) and VTE event rates.

MAIN OUTCOMES AND MEASURES The association between hospital VTE prophylaxis adherence and risk-adjusted VTE event rates was examined. The relationship between a summary score of hospital structural characteristics reflecting quality (hospital size, numbers of accreditations/quality initiatives) and performance on VTE prophylaxis and risk-adjusted VTE measures was examined. Hospital-level VTE event rates were compared across VTE diagnostic imaging rate quartiles and with a quantile regression.

RESULTS Greater hospital VTE prophylaxis adherence rates were weakly associated with worse risk-adjusted VTE event rates ($r^2 = 4.2\%$; $P = .03$). Hospitals with increasing structural quality scores had higher VTE prophylaxis adherence rates (93.3% vs 95.5%, lowest vs highest quality quartile; $P < .001$) but worse risk-adjusted VTE rates (4.8 vs 6.4 per 1000, lowest vs highest quality quartile; $P < .001$). Mean VTE diagnostic imaging rates ranged from 32 studies per 1000 in the lowest imaging use quartile to 167 per 1000 in the highest quartile ($P < .001$). Risk-adjusted VTE rates increased significantly with VTE imaging use rates in a stepwise fashion, from 5.0 per 1000 in the lowest quartile to 13.5 per 1000 in the highest quartile ($P < .001$).

CONCLUSIONS AND RELEVANCE Hospitals with higher quality scores had higher VTE prophylaxis rates but worse risk-adjusted VTE rates. Increased hospital VTE event rates were associated with increasing hospital VTE imaging use rates. Surveillance bias limits the usefulness of the VTE quality measure for hospitals working to improve quality and patients seeking to identify a high-quality hospital.

JAMA. 2013;310(14):1482-1489. doi:10.1001/jama.2013.280048
Published online October 7, 2013.

← Editorial page 1453

+ Author Video Interview at jama.com

+ Supplemental content at jama.com

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Karl Y. Bilimoria, MD, MS, Surgical Outcomes and Quality Improvement Center, Department of Surgery, Feinberg School of Medicine, Northwestern University and Northwestern Memorial Hospital, 676 St Clair St, Arkes Pavilion Ste 6-650, Chicago, IL 60611 (k-bilimoria@northwestern.edu).

Venous thromboembolism (VTE), which includes deep vein thrombosis (DVT) and pulmonary embolism (PE), is a common postoperative complication that remains a leading potentially preventable cause of postoperative morbidity and mortality.¹ The Agency for Healthcare Research and Quality developed a risk-adjusted postoperative VTE rate measure, Patient Safety Indicator 12 (PSI-12).² Endorsed by the National Quality Forum in 2008, the VTE outcome measure has been incorporated into numerous quality improvement programs and public reporting initiatives, including the Centers for Medicare & Medicaid Services 2015 Value-based Purchasing program.³

However, measuring VTE rates may be flawed because of surveillance bias,⁴⁻⁶ in which variation in outcomes reflects variation in screening and detection, or “the more you look, the more you find” phenomenon. This can occur in a number of ways: hos-

DVT deep vein thrombosis

PE pulmonary embolism

PSI-12 Patient Safety Indicator 12

SCIP-VTE-2 Surgical Care Improvement Project for VTE

VTE venous thromboembolism

pitals may use screening protocols, in which asymptomatic patients routinely undergo VTE imaging studies on a certain postoperative day,^{1,7,8} or clinicians have a lower threshold to order a VTE imaging study

for patients with minimal or equivocal signs or symptoms (eg, any leg swelling prompts a venous duplex). Hospitals that are more vigilant and perform more imaging studies for VTE may identify more VTE events, thus resulting in paradoxically worse performance on the VTE outcome measure.

If VTE rates are subject to a surveillance bias, then (1) higher VTE prophylaxis adherence should not be associated with lower VTE rates (ie, process-outcome dissociation), (2) hospitals with more characteristics reflecting quality (ie, accreditations and characteristics that generally are thought to be associated with better performance and lower complication rates) should demonstrate better VTE prophylaxis adherence rates but worse VTE rates, and (3) hospital risk-adjusted VTE rates should be related to hospital VTE imaging use rates.

To examine the effect of surveillance bias on the validity of VTE as a quality measure, 3 analyses were performed. First, we used national hospital-level data to examine the relationship between hospital VTE prophylaxis adherence and VTE event rates. Second, these data were also used to examine the associations between hospital characteristics reflecting higher quality and hospital VTE prophylaxis and event rates. Third, the associations between hospital VTE imaging and VTE event rates were examined with patient-level Medicare claims data.

Methods

This study was approved by the Northwestern University Institutional Review Board.

Comparison of VTE Rates With Process and Structural Measures

Data Sources

To evaluate the association between VTE rates (PSI-12), VTE prophylaxis adherence (SCIP-VTE-2 [Surgical Care Improvement

Project for VTE]), and structural measures of hospital quality, we used publicly reported hospital-level data on VTE prophylaxis rates and risk-adjusted VTE performance from the 2010 Centers for Medicare & Medicaid Services Hospital Compare program (October 2011 release)⁹ and data on hospital characteristics from the 2010 American Hospital Association Annual Hospital Survey.

Measures

SCIP-VTE-2 measures the proportion of surgical patients who received appropriate VTE perioperative prophylaxis (mechanical or chemoprophylaxis, depending on the operation group).¹⁰ PSI-12 measures the hospital inpatient postoperative VTE rate per 1000 surgical discharges, adjusted for differences in case mix.^{2,11} Inclusion and exclusion criteria for SCIP-VTE-2 and PSI-12 are similar.^{2,10}

To measure the structural quality of hospitals, we a priori selected 8 hospital characteristics previously used to examine health care quality and which reflect a hospital's resources and focus on programs intended to provide higher-quality care¹²⁻¹⁸: (1) bed size (<300 beds, ≥300 beds); (2) accreditation by the Joint Commission, (3) Commission on Cancer, (4) or as a level 1 trauma center; (5) presence of a residency training approved by the Accreditation Council for Graduate Medical Education; (6) provision of burn care services or (7) transplant surgery services; and (8) whether a hospital disseminated reports to its community on quality and costs of health care services. A summary score was created for each hospital's overall structural quality (0 to 8, with 8 representing the highest quality).

Statistical Analysis

The association between hospital risk-adjusted VTE rates and VTE prophylaxis adherence rates was examined with bivariate Spearman correlations. To evaluate the association between hospitals' summary score of structural quality and their performance on the VTE process and outcomes measures, we divided hospitals into quartiles according to their structural quality summary scores. Risk-adjusted VTE rates and VTE prophylaxis adherence rates were examined across quartiles of structural quality scores with 1-way analysis of variance, with Bonferroni correction for pairwise differences in means and Cuzick extension of the Wilcoxon signed rank test for trends.¹⁹ All tests were 2-sided, with significance set at .05.

Assessment of VTE Imaging Use and Event Rates

Data Source and Patients

To construct hospital-level VTE imaging rates and risk-adjusted VTE rates, patient-level data were obtained from the Medicare Provider and Analysis Review, Carrier, and Outpatient claims files. Fee-for-service Medicare beneficiaries aged 65 years or older who had one of 11 major operations (abdominal aortic aneurysm repair, coronary artery bypass graft, craniotomy, colectomy, cystectomy, esophagectomy, gastric bypass, lung resection, pancreatic resection, proctectomy, or total knee arthroplasty) between January 1, 2009, and December 31, 2010, were identified with *International Classification of Diseases, Ninth Revision (ICD-9)* codes (eTable 1 in the Supplement). Agency for Healthcare Research and Quality PSI-12 exclusions were applied.²

Measures

Hospital-level VTE surveillance imaging rates were ascertained by identifying VTE imaging studies with Centers for Medicare & Medicaid Services Healthcare Common Procedure Coding System codes for services occurring during the inpatient postoperative period (eTable 1 in the Supplement). Deep vein thrombosis imaging included venous duplex ultrasonography of an upper or lower extremity. Pulmonary embolism imaging included chest computed tomography scans, ventilation-perfusion scans, and chest magnetic resonance imaging.

Patient-level postoperative VTEs were identified with secondary ICD-9 diagnosis codes based on PSI-12 specifications. Hospital-level observed rates of VTE, DVT, and PE were calculated as the number of patients with VTE, DVT, and PE per 1000 discharges. To construct hospital-level risk-adjusted VTE, DVT, and PE rates, the Agency for Healthcare Research and Quality PSI-12 risk-adjustment methodology was used.^{2,11} Logistic regression models adjusted for patient age, transfer status, Elixhauser comorbidities, modified diagnosis related groups, and major diagnostic categories as hierarchical models with random hospital effects may overadjust hospital-level variation underlying VTE imaging use. Hospital-level-predicted VTE, DVT, and PE rates were calculated as the sum of predicted events per 1000 discharges. Risk-adjusted VTE, DVT, and PE rates were calculated as the ratio of observed to predicted rates multiplied by the overall rate of each complication in the analytic sample.

Statistical Analysis

To investigate surveillance bias, hospitals were grouped into quartiles according to rates of VTE-related imaging. Quartiles were selected a priori to allow evaluation of the effect of surveillance bias at the high and low ends of VTE imaging rates because the relationship between VTE imaging and events is not linear. Mean risk-adjusted VTE event rates were examined across quartiles of hospital VTE imaging use with 1-way analysis of variance, with Bonferroni correction for pairwise comparisons and Cuzick extension of the Wilcoxon signed-rank test to test for overall trends.¹⁹ All tests were 2-sided, with significance set at .05. As a supplemental approach to characterize surveillance bias that preserves full range of variation in both VTE imaging and complication rates, we used quantile regression to examine whether imaging rates had differential effects across the distribution of risk-adjusted VTE event rates. The association of VTE imaging and VTE event rates was assessed at the 25th, 50th (median), 75th, 85th, 95th, and 99th quantiles of the VTE event rate distribution.

Additional Analyses

Four additional analyses were performed. First, we assessed whether publicly reported hospital VTE performance from Hospital Compare 2010 also demonstrated a surveillance bias.

Second, to better account for differences in patient risk factors across VTE imaging use quartiles, a more comprehensive risk-adjustment approach was performed with additional patient-level covariates (eTable 2 in the Supplement).

Third, analyses were repeated with a hierarchical regression model with hospital random intercepts.

Fourth, a competing hypothesis is that high VTE imaging use hospitals may not provide adequate VTE prophylaxis, thus

resulting in higher VTE rates and correspondingly requiring more VTE imaging studies. We tested this alternative explanation by comparing SCIP-VTE-2 adherence across VTE imaging use quartiles.

Results

Comparison of VTE Rates With Process and Structural Measures

In 2010, 2838 hospitals reported VTE process and outcome measures to Hospital Compare (Table 1). The mean hospital VTE prophylaxis (SCIP-VTE-2) adherence rate was 94.51% (95% CI, 94.29% to 94.73%). The mean risk-adjusted PSI-12 VTE rate per 1000 discharges was 5.35 (95% CI, 5.24 to 5.45). SCIP-VTE-2 adherence was weakly correlated with hospital PSI-12 rates ($r^2 = 4.2\%$; $P = .03$) (Figure 1) in the opposite direction than expected: hospitals with higher SCIP-VTE-2 adherence also had higher (not lower) PSI-12 rates.

Compared with hospitals with the lowest structural quality scores (lowest quartile; fewest characteristics reflecting quality), those with the highest ones (highest quartile; most characteristics reflecting quality) had significantly better performance on the SCIP-VTE-2 process measure (93.27% for lowest quartile vs 95.47% for highest quartile; $P < .001$) but paradoxically worse performance on the PSI-12 VTE outcome measure (4.77 per 1000 for lowest quartile vs 6.37 per 1000 for highest quartile; $P < .001$) (Figure 2; eTable 3 in the Supplement).

Assessment of VTE Imaging Use and Event Rates

Facilities in the Medicare sample were excluded that reported fewer than 10 total cases ($n = 73$ hospitals) or no VTE imaging studies during the study period ($n = 445$ hospitals; $n = 9442$ patients) (eTable 4 in the Supplement). Of the hospitals with no VTE imaging studies reported, 436 (98%) reported no VTE events. Thus, the analytic sample included 954 526 patients from 2786 hospitals (Table 1). Patient-level event rates were 0.99% for VTE, 0.56% for DVT, and 0.50% for PE. Overall, 8.82% of patients received some form of VTE imaging.

The mean VTE imaging rate in the quartile with the highest VTE imaging use was more than 5-fold that in the lowest quartile (167.05 vs 31.55 per 1000 discharges; $P < .001$) (Table 2). The mean unadjusted VTE rate in the quartile of hospitals with the highest VTE imaging use was 15.07 per 1000 discharges compared with 4.34 in lowest-quartile hospitals ($P < .001$).

Risk-adjusted VTE rates increased significantly in a stepwise fashion across VTE imaging use quartiles (Table 2). The mean risk-adjusted VTE rate in the lowest VTE imaging use quartile was 4.99 per 1000 discharges compared with 13.46 in the highest quartile ($P < .001$ trend and pairwise) (Figure 3A). The mean risk-adjusted DVT rate in the lowest DVT imaging use quartile was 1.79 per 1000 discharges compared with 9.14 in the highest quartile ($P < .001$ trend and pairwise) (Figure 3B). The mean risk-adjusted PE rate in the lowest PE imaging use quartile was 2.81 per 1000 discharges compared with 7.38 in the highest quartile ($P < .001$ trend and pairwise) (Figure 3C). Similar findings were observed when each of the 11 procedures was examined

individually, except gastric bypass, which could not be modeled individually because of a very low VTE event rate.

Using quantile regression to characterize the association between VTE imaging rates and risk-adjusted VTE event rates, we found a positive association between imaging and VTE events across the distribution of VTE event rates. However, we found larger effects among hospitals with the highest VTE event rates (eTable 5 in the Supplement). For example, a unit change in VTE imaging rates increased the 25th quantile of risk-adjusted VTE by 0.025 (95% CI, 0.016-0.033) compared with 0.129 (95% CI, 0.102-0.156) for the 95th quantile and 0.457 (95% CI, 0.293-0.620) for the 99th quantile.

Four additional analyses were performed. First, publicly reported overall hospital risk-adjusted VTE rates from Hospital Compare increased across VTE imaging use quartiles (3.94 to 7.32 per 1000 discharges; $P < .001$), and the proportion of hospitals with “worse than national rate” designations also increased across VTE imaging use quartiles (0.15% to 19.03%; $P < .001$) (Table 2).

Second, because the findings could be due to sicker patients being treated at the highest VTE imaging use hospitals, the analysis was performed with additional patient-level covariates for risk adjustment than those specified for Agency for Healthcare Research and Quality PSI-12, but the association between hospital VTE imaging use and VTE rate persisted (eTable 6 in the Supplement).

Third, hierarchical models yielded a similar but slightly blunted relationship between VTE imaging rates and VTE event rates (eTable 6).

Fourth, although the findings could be affected by hospital variation in VTE prophylaxis use practices, mean hospital performance on VTE prophylaxis administration (SCIP-VTE-2) was not significantly worse in the highest VTE imaging use quartile (Table 2).

Discussion

Our results raise concerns about the validity of using VTE rates as a hospital quality metric. Instead of finding the expected relationship between higher VTE prophylaxis adherence rates and lower VTE events rates, we found that VTE prophylaxis rates were positively correlated with VTE event rates ($r^2 = 4.2\%$; $P = .03$). A paradoxical relationship was also found between a measure of hospital structural characteristics reflecting quality and VTE event rates: hospitals with higher structural quality scores had better VTE prophylaxis adherence rates (93.3% lowest quartile vs 95.5% highest quartile; $P < .001$), but they had unexpectedly higher risk-adjusted VTE rates (4.8 per 1000 lowest quartile vs 6.4 per 1000 highest quartile; $P < .001$). Most important, hospital VTE rates were associated with the intensity of detecting VTE with imaging studies. Hospitals in the lowest quartile for VTE imaging rates obtained a mean of 32 studies per 1000 discharges, whereas those in the highest quartile obtained 167 imaging studies per 1000 discharges. Lowest imaging rate quartile hospitals diagnosed 5.0 VTEs per 1000 discharges, whereas the highest imaging rate quartile hospitals found 13.5 VTEs per 1000 discharges. These findings sup-

Table 1. Characteristics of the 2 Data Sets Used in the Study: Hospitals Reporting to Hospital Compare (2010) and Medicare Beneficiaries in the Patient-Level Data Set (2009-2010)

Data Sets and Characteristics	No. (%)
Hospitals from Hospital Compare and AHA Data Sets	
No. of hospitals reporting to Hospital Compare	2838
Hospital structural quality characteristics	
Total bed size ≥ 300	768 (27.1)
Joint Commission accreditation	2428 (85.6)
Commission on Cancer accreditation	1303 (45.9)
ACGME accreditation	746 (26.3)
Level I trauma certification	309 (10.9)
Burn services	130 (4.6)
Transplant services	381 (13.4)
Quality and cost reporting to community	1853 (65.3)
No. of structural “quality” characteristics, mean (SD) [range]	2.79 (1.77) [0-8]
Patients from Medicare claims data set^a	
No. of Medicare beneficiaries in the patient-level data set	954 926
Age, mean (SD), y	74.68 (6.6)
No. of Elixhauser comorbidities, mean (SD)	2.04 (1.2)
Transfer status	33 863 (3.6)
Type of surgery	
Abdominal aortic aneurysm repair	42 429 (4.5)
Colectomy	159 058 (16.7)
Coronary artery bypass graft surgery	131 949 (13.8)
Craniotomy	54 597 (5.7)
Cystectomy	10 023 (1.1)
Esophageal resection	4423 (0.5)
Gastric bypass	3641 (0.4)
Lung resection	39 390 (4.1)
Pancreatic resection	9180 (1.0)
Proctectomy	18 871 (2.0)
Total knee arthroplasty	480 965 (50.4)
Inpatient events	
VTE	9423 (0.99)
DVT	5375 (0.56)
PE	4792 (0.50)
Imaging received	
VTE	844 155 (8.82)
DVT	58 244 (6.10)
PE	36 065 (3.78)

Abbreviations: ACGME, Accreditation Council for Graduate Medical Education; AHA, American Hospital Association; DVT, deep vein thrombosis; PE, pulmonary embolism; VTE, venous thromboembolism.

^aN = 2786 hospitals.

port the surveillance bias hypothesis, suggesting that VTE rates reflect the hospital staff’s vigilance in looking for postoperative VTE, not true quality of care.

Comparison of VTE Rates With Process and Structural Measures

Our results are consistent with those of 2 previous studies that did not find a relationship between better VTE prophylaxis adherence and lower VTE rates.^{17,20} Moreover, we found that better VTE prophylaxis adherence rates are weakly associated with

Figure 1. Hospital Venous Thromboembolism (VTE) Prophylaxis Adherence Rates and Risk-Adjusted VTE Event Rates

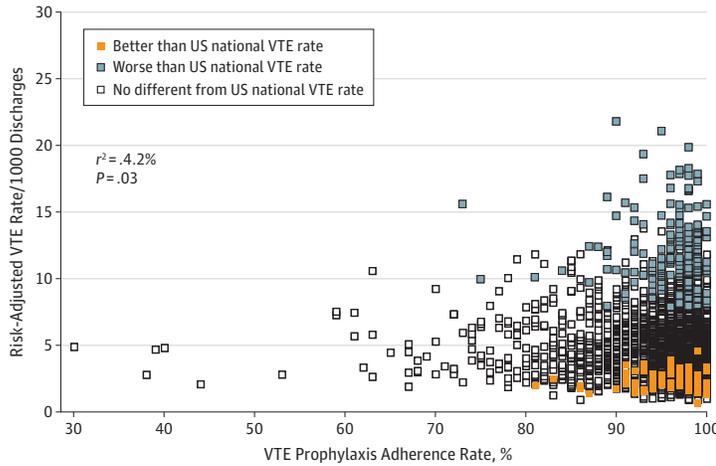
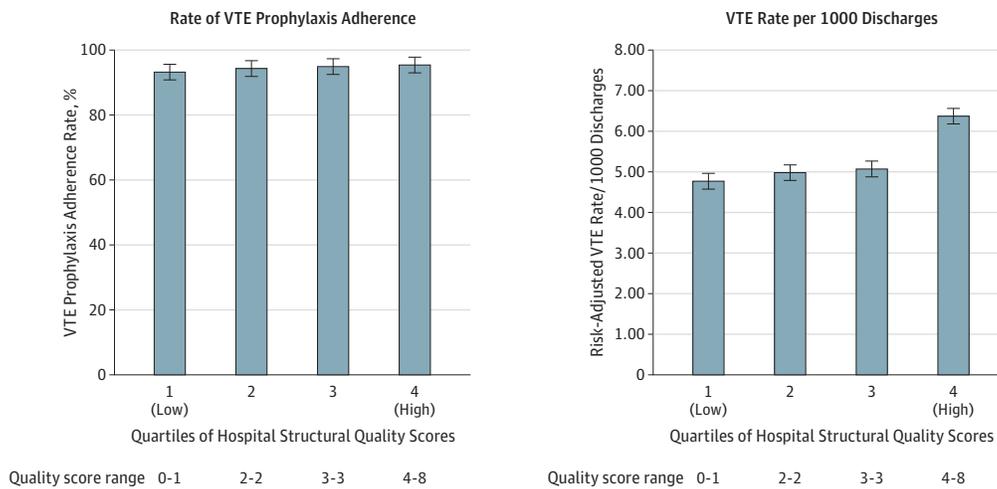


Figure 2. Venous Thromboembolism (VTE) Prophylaxis Adherence Rates and Risk-Adjusted VTE Rates by Hospital Structural Quality Scores



For both panels, $P < .001$ by 2-tailed test for differences in rates compared with quartile 4 rate after Bonferroni correction for multiple pairwise comparisons. $P < .001$ by Cuzick extension of the Wilcoxon rank-sum test for trend across ordered categories. Error bars indicate 95% CIs.

unexpectedly worse risk-adjusted VTE rates. Although this could be due to problems with the process measure, it suggests an issue with the VTE outcome measure, as would be expected if VTE measurement is subject to a surveillance bias.

We also found that hospitals with increasing numbers of structural quality characteristics (ie, larger hospitals with more accreditations and engagement in quality initiatives that typically suggest high-quality care) had better VTE prophylaxis (SCIP-VTE-2) adherence rates, as one would anticipate, given their increased resources and focus on quality; but they unexpectedly had higher VTE rates (PSI-12). These paradoxical findings suggest that the lack of an association between VTE process and outcome may more likely be due to problems with the VTE outcome measure, consistent with surveillance bias.

Assessment of VTE Imaging Use and Event Rates

The presence of a surveillance bias may explain the lack of expected correlations between VTE rates, VTE prophylaxis rates, and structural quality measures and is consistent with previous observations in trauma patients. A previous study found no differences in DVT rates for the lowest 3 quartiles of venous duplex ultrasonographic use were found, yet higher DVT rates in the highest quartile of venous duplex ultrasonography were observed.⁵ This study was limited by the quality of the self-reported data, reflected in that data from only 147 of 700 hospitals could be analyzed.²¹ To our knowledge, no previous studies have investigated surveillance bias outside the trauma patient population. Using Medicare claims data offered us more complete information on imaging and VTE event

Table 2. Venous Thromboembolism Rates by VTE Imaging Use Rate Quartiles

	VTE Imaging Use Rate Quartiles			
	1	2	3	4
No. of hospitals	697	696	708	685
VTE imaging rate per 1000 discharges, mean (95% CI) [range] ^{a,c}	31.55 ^b (30.70-32.41) [2.72-48.54]	61.61 ^b (61.06-62.17) [48.57-74.69]	91.21 ^b (90.43-91.98) [74.70-111.11]	167.05 (161.95-172.15) [111.52-612.90]
VTE event rate per 1000 discharges, mean (95% CI) ^{a,c}	4.34 ^b (3.84-4.85)	7.53 ^b (6.89-8.18)	10.70 ^b (9.84-11.56)	15.07 (13.96-16.18)
Hospital Compare				
VTE prophylaxis adherence				
No. of hospitals reporting	633	674	685	670
VTE prophylaxis adherence (SCIP-VTE-2), mean (SD) ^{d,e}	91.81 (91.21-92.41)	92.58 (92.07-93.09)	92.73 (92.25-93.22)	93.03 (92.50-93.56)
Risk-adjusted VTE rate (PSI-12) ^d				
No. of hospitals reporting	669	675	676	662
VTE rate per 1000 discharges, mean (SD) ^{c,d}	3.94 ^b (3.80-4.07)	4.74 ^b (4.58-4.89)	5.61 ^b (5.43-5.79)	7.32 (7.05-7.59)
Hospitals with "worse than US national" VTE rate designation, No. (%) ^f	1 (0.15)	4 (0.59)	38 (5.62)	126 (19.03)

Abbreviations: PSI-12, Patient Safety Indicator 12; SCIP-VTE-2, Surgical Care Improvement Project for VTE; VTE, venous thromboembolism.

^aResults from the patient-level Medicare data covering 11 operations.

^bP < .001, 2-tailed test for differences in means compared with quartile 4 mean after Bonferroni correction for multiple pairwise comparisons.

^cP < .001 Cuzick extension of Wilcoxon rank-sum test for trend across ordered categories.

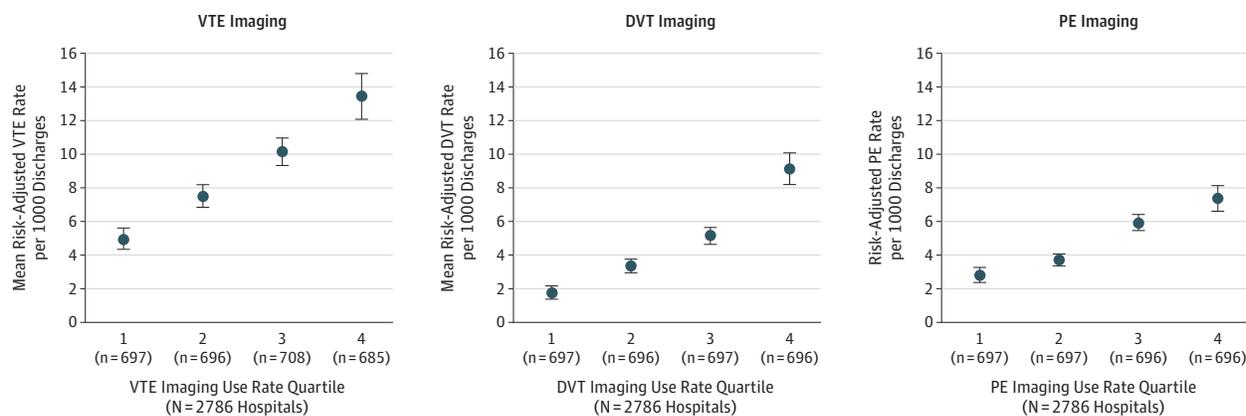
^dResults from the hospital-level publicly available Hospital Compare data

covering all of the operations included in the risk-adjusted VTE event rate measure (PSI-12). The hospital VTE diagnostic imaging rates calculated from the Medicare data are used to assess hospital-wide VTE rates from Hospital Compare.

^eP = .002 Cuzick extension of Wilcoxon rank-sum test for trend across ordered categories.

^fP < .001 Cochran Armitage trend test.

Figure 3. Mean Risk-Adjusted Event Rates by Imaging Use Rate Quartile



VTE indicates venous thromboembolism; DVT, deep vein thrombosis; and PE, pulmonary embolism. For all panels, P < .001 by trend and pairwise for comparison of differences in rates for quartile 4 (highest) compared with each

other quartile. Numbers in parentheses under the x-axis represent number of hospitals. Error bars indicate 95% CIs around the intraquartile means.

rates across a broad, nationally representative sample of operations and hospitals.

The trauma patient population is unique because routine screening of asymptomatic patients for DVT is common at many trauma centers, but there is variation in the use and application of screening protocols.⁷ Surveillance bias might explain high VTE rates in trauma patients because routine screening in asymptomatic patients results in more screening tests, with correspondingly higher VTE event rates.^{6,7} Similarly, routine screening protocols have been used for orthopedics, neu-

rosurgery, and bariatric patients and may contribute to higher VTE rates at hospitals using such practices.^{8,22-24} The significant association between surveillance imaging use and higher risk-adjusted VTE rates that we found when examining individual procedures, such as colorectal surgery, suggests that surveillance bias influences VTE rates for patients and procedures where VTE screening protocols are not generally used. Importantly, this indicates that there are differences in hospital practices or culture in regard to thresholds for ordering VTE diagnostic imaging studies in symptomatic patients

(eg, any tachycardia prompts a computed tomography scan for PE), and this influences apparent VTE rates. Our findings suggest that performance on the risk-adjusted PSI-12 VTE measure may mostly reflect differences in VTE imaging use rather than differences in the ability to truly prevent VTE.

Two alternative explanations should be considered. First, it may be that some hospitals perform more duplexes in response to signs or symptoms and have truly higher VTE rates because they deliver poorer quality care by not adequately providing postoperative VTE prophylaxis. Our data suggest that this is not the case: VTE prophylaxis rates based on SCIP-VTE-2 were similar across surveillance quartiles, if not slightly better in the highest surveillance quartile.

Second, it may be that the differences in VTE surveillance imaging rates are due to hospital-level variation in case mix and severity of illness. Although we focused on using the modeling approach and risk-adjustment variables specified for Agency for Healthcare Research and Quality PSI-12, we did estimate risk-adjusted VTE rates with additional risk factors (eg, additional surgical details, emergency admission status). There may be some residual confounding from other VTE risk factors (eg, central lines, estrogen use) for which we do not have information. However, many of these factors affecting patient-level risk estimates do not greatly influence hospital-level quality comparisons.^{25,26} We think it is unlikely that residual differences in case-mix adjustment explain our findings.

Some additional limitations should be considered. First, there may be many hospital accreditations or surgical quality initiatives reflecting a hospital's commitment to providing higher-quality, specialized care that we did not include in the structural quality indicator. We did examine inclusion of many other hospital characteristics, but because the findings were comparable, we retained those selected a priori.

Second, our analyses were based on Medicare administrative data for patients who underwent 11 major operations. Our findings may not be generalizable to younger patients or to those undergoing other procedures.

Third, we excluded 458 hospitals that did not report any VTE surveillance imaging studies. It is unclear why these hospitals reported neither these studies nor VTE events, but the institutions accounted for less than 1% of the patients in the study.

Fourth, although there are known limitations in accurately ascertaining outcomes from administrative data, we used them because they are the basis of the public reporting and pay-for-performance metrics.

Fifth, chest computed tomography and magnetic resonance imaging are performed for a variety of indications other than to evaluate a patient for PE. We could not ascertain why imaging tests were conducted from the Medicare data, so there is likely some error in assuming all the imaging tests for PE were truly for the evaluation of PE. However, surveillance bias was present in separate analyses of DVT in which indications for the imaging study, venous duplex, are specific.

Implications

Our study calls into question the merit of the PSI-12 VTE outcome measure as a quality measure and its use in public reporting and performance-based payments. Hospitals reported to have the highest risk-adjusted VTE rates may in fact be providing vigilant care by ordering imaging studies to ensure that VTE events are not missed. Patients selecting hospitals according to publicly available metrics may be misled by currently reported VTE performance. The measure could be counterproductive if a hospital performs poorly on the VTE outcome metric, expends efforts to improve VTE prophylaxis resulting in increased awareness and vigilance in looking for VTE, and then finds more VTEs and becomes an even worse performer on the VTE measure.

A quality metric affected by surveillance bias may have unintended consequences. Clinicians may change practice by using VTE surveillance imaging less frequently, ordering chemoprophylaxis in low-VTE-risk populations, and increasing preoperative inferior vena cava filter use.^{6,27} Alternatively, hospitals with higher VTE imaging rates may be overusing VTE imaging and detecting otherwise asymptomatic VTE, which leads to treatment with anticoagulants or IVC filters and could cause harm. Surveillance bias likely influences other quality metrics,^{6,28-30} and the possibility that it might perversely encourage the wrong practices should be carefully considered when quality metrics are developed.

Conclusion

Although there is considerable attention given to measuring and improving hospital VTE rates, our findings suggest that a surveillance bias influences the validity of VTE measurement. The publicly reported PSI-12 VTE outcome measure reflects the intensity of VTE imaging rather than actual quality of care and should likely not be used for accountability purposes in quality measurement.

ARTICLE INFORMATION

Published Online: October 7, 2013.
doi:10.1001/jama.2013.280048.

Author Affiliations: Surgical Outcomes and Quality Improvement Center, Department of Surgery and Center for Healthcare Studies, Feinberg School of Medicine, Northwestern University and Northwestern Memorial Hospital, Chicago, Illinois (Bilimoria, Chung, Ju, Bentrem); Departments of Surgery, Anesthesiology and Critical Care Medicine, and Emergency Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland (Haut);

Department of Surgery, Jesse Brown VA Medical Center, Chicago, Illinois (Bentrem); Department of Surgery, University of California at Los Angeles, Los Angeles, California (Ko); Division of General Internal Medicine and Geriatrics, Department of Medicine, and Institute for Public Health and Medicine, Northwestern University, Chicago, Illinois (Baker).

Author Contributions: Drs Bilimoria and Chung had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Chung, Ju, Bilimoria.

Acquisition of data: Bentrem, Bilimoria.
Analysis and interpretation of data: All authors.
Drafting of the manuscript: Chung, Ju, Baker, Bilimoria.
Critical revision of the manuscript for important intellectual content: All authors.
Statistical analysis: Chung, Ju, Bilimoria.
Obtained funding: Bentrem, Bilimoria.
Administrative, technical, or material support: Bentrem, Bilimoria.
Study supervision: Bilimoria.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr. Bilimoria is the primary investigator of R21HS021857 from the Agency for Healthcare Research and Quality (AHRQ); reports receiving support from the American Cancer Society, National Comprehensive Cancer Network, Northwestern University, the Robert H. Lurie Comprehensive Cancer Center, Northwestern Memorial Foundation, and Northwestern Memorial Hospital; and has received honoraria from hospitals, professional societies, and continuing medical education companies for clinical care and quality improvement research presentations. Dr. Haut is the primary investigator of the Mentored Clinician Scientist Development Award KO81K08HS017952-01 from the AHRQ, receives royalties from Lippincott Williams & Wilkins for a book he coauthored, and has given expert witness testimony in various medical malpractice cases. Dr. Bentrem is supported by a career development award from the Health Services Research and Development Service of the Department of Veterans Affairs. Dr. Baker has received honoraria for research presentations from Abbott Laboratories and Merck. Dr. Baker currently receives grant funding from the AHRQ and the National Institutes of Health.

Funding/Support: This study was supported by AHRQ R21HS021857 and a Center Development Award from Northwestern University and Northwestern Memorial Hospital to Dr. Bilimoria. Dr. Ju is supported by National Institutes of Health 5T32HL094293. Dr. Haut is supported by a career development award 1K08HS017952-01 from the AHRQ.

Role of the Sponsor: The funding sponsors played no part in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript. The sponsors had no access to the data and did not perform any of the study analysis.

Disclaimer: The views expressed in this article do not necessarily represent the views of the US government.

Previous Presentation: This study was presented in part at the 2013 Academy Health Annual Research Meeting on June 24, 2013, in Baltimore, Maryland.

Additional Contributions: We acknowledge Min-Woong Sohn, PhD, and Elissa Oh, MS (Surgical Outcomes and Quality Improvement Center and Center for Healthcare Studies, Northwestern University), for assistance with data management and thank Jay Anderson, MBA, Gary Noskin, MD, Scott Greene, MD, and Cindy Barnard, MBA, MSJS (Northwestern Memorial Hospital), for their input on this study. None of those acknowledged received compensation for their contributions.

REFERENCES

- Gould MK, Garcia DA, Wren SM, et al; American College of Chest Physicians. Prevention of VTE in nonorthopedic surgical patients: *Antithrombotic Therapy and Prevention of Thrombosis*, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest*. 2012;141(2)(suppl):e227S-e277S.
- Agency for Healthcare Research and Quality. Postoperative pulmonary embolism or deep vein thrombosis rate. <http://www.qualityindicators.ahrq.gov/Downloads/Modules/PSI/V44/TechSpecs/PSI%2012%20Postoperative%20PE%20or%20DVT%20Rate.pdf>. Accessed March 1, 2013.
- Centers for Medicare & Medicaid Services. Hospital value-based purchasing. <http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/hospital-value-based-purchasing/index.html>. Accessed September 1, 2013.
- Sackett DL. Bias in analytic research. *J Chronic Dis*. 1979;32(1-2):51-63.
- Pierce CA, Haut ER, Kardooni S, et al. Surveillance bias and deep vein thrombosis in the National Trauma Data Bank. *J Trauma*. 2008;64(4):932-936; discussion 936-937.
- Haut ER, Pronovost PJ. Surveillance bias in outcomes reporting. *JAMA*. 2011;305(23):2462-2463.
- Haut ER, Schneider EB, Patel A, et al. Duplex ultrasound screening for deep vein thrombosis in asymptomatic trauma patients: a survey of individual trauma surgeon opinions and current trauma center practices. *J Trauma*. 2011;70(1):27-33; discussion 33-34.
- Henwood PC, Kennedy TM, Thomson L, et al. The incidence of deep vein thrombosis detected by routine surveillance ultrasound in neurosurgery patients receiving dual modality prophylaxis. *J Thromb Thrombolysis*. 2011;32(2):209-214.
- Centers for Medicare & Medicaid Services. Centers for Medicare & Medicaid Services Hospital Compare website. <http://www.medicare.gov/hospitalcompare/>. Accessed February 27, 2013.
- Joint Commission. Specifications manual for national hospital inpatient quality measures. http://www.jointcommission.org/specifications_manual_for_national_hospital_inpatient_quality_measures.aspx. Accessed September 16, 2012.
- Agency for Healthcare Research and Quality. Patient safety indicators (PSI): risk adjustment coefficients for the PSI, version 4.4. <http://www.qualityindicators.ahrq.gov/Downloads/Modules/PSI/V44/Risk%20Adjustment%20Tables%20PSI%204.4.pdf>. Accessed August 7, 2012.
- Lehrman WG, Elliott MN, Goldstein E, et al. Characteristics of hospitals demonstrating superior performance in patient experience and clinical process measures of care. *Med Care Res Rev*. 2010;67(1):38-55.
- Schmaltz SP, Williams SC, Chassin MR, et al. Hospital performance trends on national quality measures and the association with Joint Commission accreditation. *J Hosp Med*. 2011;6(8):454-461.
- Brand CA, Barker AL, Morello RT, et al. A review of hospital characteristics associated with improved performance. *Int J Qual Health Care*. 2012;24(5):483-494.
- Friese CR, Earle CC, Silber JH, Aiken LH. Hospital characteristics, clinical severity, and outcomes for surgical oncology patients. *Surgery*. 2010;147(5):602-609.
- Bilimoria KY, Bentrem DJ, Stewart AK, Winchester DP, Ko CY. Comparison of commission on cancer-approved and -nonapproved hospitals in the United States: implications for studies that use the National Cancer Data Base. *J Clin Oncol*. 2009;27(25):4177-4181.
- Altom LK, Deierhoi RJ, Grams J, et al. Association between Surgical Care Improvement Program venous thromboembolism measures and postoperative events. *Am J Surg*. 2012;204(5):591-597.
- Vartak S, Ward MM, Vaughn TE. Do postoperative complications vary by hospital teaching status? *Med Care*. 2008;46(1):25-32.
- Cuzick J. A Wilcoxon-type test for trend. *Stat Med*. 1985;4(1):87-90.
- Nicholas LH, Osborne NH, Birkmeyer JD, Dimick JB. Hospital process compliance and surgical outcomes in Medicare beneficiaries. *Arch Surg*. 2010;145(10):999-1004.
- Kardooni S, Haut ER, Chang DC, et al. Hazards of benchmarking complications with the National Trauma Data Bank. *J Trauma*. 2008;64(2):273-277.
- Misra M, Roitberg B, Ebersole K, Charbel FT. Prevention of pulmonary embolism by combined modalities of thromboprophylaxis and intensive surveillance protocol. *Neurosurgery*. 2004;54(5):1099-1102; discussion 1102-1103.
- McAndrew CM, Fitzgerald SJ, Kraay MJ, Goldberg VM. Incidence of postthrombotic syndrome in patients undergoing primary total knee arthroplasty for osteoarthritis. *Clin Orthop Relat Res*. 2010;468(1):178-181.
- Wu EC, Barba CA. Current practices in the prophylaxis of venous thromboembolism in bariatric surgery. *Obes Surg*. 2000;10(1):7-13; discussion 14.
- Merkow RP, Bentrem DJ, Winchester DP, et al. Effect of including cancer-specific variables on risk-adjusted hospital surgical quality comparisons. *Ann Surg Oncol*. 2013;20(6):1766-1773.
- Merkow RP, Krniecik TE, Bentrem DJ, et al. Effect of including cancer-specific variables on models examining short-term outcomes. *Cancer*. 2013;119(7):1412-1419.
- Baker DW, Qaseem A; American College of Physicians' Performance Measurement Committee. Evidence-based performance measures: preventing unintended consequences of quality measurement. *Ann Intern Med*. 2011;155(9):638-640.
- Wu N, Mor V, Roy J. Resident, nursing home, and state factors affecting the reliability of Minimum Data Set quality measures. *Am J Med Qual*. 2009;24(3):229-240.
- Lin MY, Hota B, Khan YM, et al; CDC Prevention Epicenter Program. Quality of traditional surveillance for public reporting of nosocomial bloodstream infection rates. *JAMA*. 2010;304(18):2035-2041.
- Magill SS, Fridkin SK. Improving surveillance definitions for ventilator-associated pneumonia in an era of public reporting and performance measurement. *Clin Infect Dis*. 2012;54(3):378-380.