Risk-Adjusted Percent Time in Therapeutic Range as a Quality Indicator for Outpatient Oral Anticoagulation

Results of the Veterans Affairs Study To Improve Anticoagulation (VARIA)

Adam J. Rose, MD, MSc; Elaine M. Hylek, MD, MPH; Al Ozonoff, PhD; Arlene S. Ash, PhD; Joel I. Reisman, AB; Dan R. Berlowitz, MD, MPH

Background—Oral anticoagulation is safer and more effective when patients receive high-quality care. However, there have been no prior efforts to measure quality of oral anticoagulation care or to risk adjust it to ensure credible comparisons. Our objective was to profile site performance in the Veterans Health Administration (VA) using risk-adjusted percent time in therapeutic range (TTR).

Methods and Results—We included 124 551 patients who received outpatient oral anticoagulation from 100 VA sites of care for indications other than valvular heart disease from October 1, 2006, to September 30, 2008. We calculated TTR for each patient and mean TTR for each site of care. Expected TTR was calculated for each patient and each site based on the variables in the risk adjustment model, which included demographics, comorbid conditions, medications, and hospitalizations. Mean TTR for the entire sample was 58%. Site-observed TTR varied from 38% to 69% or from poor to excellent. Site-expected TTR varied from 54% to 62%. Site risk-adjusted performance ranged from 18% below expected to 12% above expected. Risk adjustment did not alter performance rankings for many sites, but for other sites, it made an important difference. For example, the site ranked 27th of 100 before risk adjustment was one of the best (risk-adjusted rank, 7). Risk-adjusted site rankings were consistent from year to year (correlation between years, 0.89).

Conclusions—Risk-adjusted TTR can be used to profile the quality of outpatient oral anticoagulation in a large, integrated health system. This measure can serve as the basis for quality measurement and quality improvement efforts. **(*Circ Cardiovasc Qual Outcomes*. 2011;4:22-29.)**

Key Words: anticoagulants ■ quality of health care ■ ambulatory care ■ risk adjustment ■ patients ■ safety

Oral anticoagulation is a highly effective but potentially dangerous therapy.[1,2] The level of anticoagulation control is a critical determinant of benefit from warfarin[3–7]; indeed, patients with atrial fibrillation may not benefit from anticoagulation unless they achieve a certain level of control.[3] However, warfarin management is difficult, and achieving good control requires much effort and skill on the part of both the patient and the clinician.[1] Because many patients do not achieve excellent control,[8] there is great potential to improve outcomes for patients by improving quality of care in oral anticoagulation.[9] To improve quality of care, we first must be able to measure it.[10,11] Previous efforts to measure quality of care in oral anticoagulation therapy have focused disproportionately on the failure to provide anticoagulation to as many ideal candidates as possible.[12] However, receipt of anticoagulation is only a first step toward improving outcomes; we also need to measure the quality of oral anticoagulation management to ensure that the benefits of anticoagulation are maximized and the harms minimized.[9]

An ideal quality indicator for outpatient oral anticoagulation would have several characteristics: it would be easy to abstract, calculate, and understand; it would vary among providers or sites of care; improvement would be possible; and there would be strong evidence linking it to important outcomes, such as stroke, venous thromboembolism, and major hemorrhage. Percent time in therapeutic range (TTR) has many of these characteristics. It can be calculated from automated data, it can be improved,[8] and it has been linked to important outcomes.[3–7] Although it might be possible to consider using definitive outcomes themselves as quality indicators, reliance on these rare events would preclude quality measurement at all but the largest sites of care.[9,13]

Received May 24, 2010; accepted October 6, 2010.

From the Center for Health Quality, Outcomes, and Economic Research (A.J.R., A.O., J.I.R., D.R.B.), Bedford VA Medical Center, Bedford, Mass; Department of Medicine (A.J.R., E.M.H., A.S.A., D.R.B.), Section of General Internal Medicine, Boston University School of Medicine, Boston, Mass; Biostatistics Section (A.O.), Boston Children's Hospital, Boston, Mass; Department of Quantitative Health Sciences (A.S.A.), Division of Biostatistics and Health Services Research, University of Massachusetts School of Medicine, Worcester, Mass; and Department of Health Policy and Management (D.R.B.), Boston University School of Public Health, Boston, Mass.

The opinions expressed in this manuscript do not necessarily represent the official views of the Department of Veterans Affairs.

The online-only Data Supplement is available at http://circoutcomes.ahajournals.org/cgi/content/full/CIRCOUTCOMES.110.957738/DC1.

Correspondence to Adam J. Rose, MD, MSc, Center for Health Quality, Outcomes, and Economic Research, Bedford VA Medical Center, 200 Springs Rd, Building 70, Bedford, MA 01730. E-mail adamrose@bu.edu

Circ Cardiovasc Qual Outcomes **is available at http://circoutcomes.ahajournals.org** **DOI: 10.1161/CIRCOUTCOMES.110.957738**

One reason why TTR has not been used previously as a quality indicator may be the absence of a risk adjustment model for TTR. Risk adjustment can enhance the credibility of performance comparisons between providers or sites by ensuring that sites are being compared regarding quality of care rather than regarding merely differences in case mix.[14] Risk adjustment can increase the acceptance of quality measures by poorly performing sites who might otherwise protest that their performance is poor because their patients are sicker. Our group recently has derived and validated a risk adjustment model for TTR that should allow for fair site-site comparisons on this quality indicator.[15] In deriving and validating our model, we confirmed the widely held belief that some patients are indeed much harder to keep within the target range than others, suggesting that risk adjusting TTR is necessary.

We therefore set out to address 3 related questions, using a database of 100 anticoagulation clinics and >100 000 patients from the Veterans Health Administration (VA). First, does mean TTR differ among sites of care? Unless meaningful differences exist, profiling is unlikely to spur quality improvement. Second, does risk adjusting TTR meaningfully alter site rankings? Risk adjustment may enhance credibility,[14] but it requires effort, so it is important to know whether it matters. Finally, would risk-adjusted site rankings be relatively constant from year to year, suggesting that risk-adjusted TTR is measuring quality of care (a stable attribute of a site) rather than measuring mere statistical variation? Our overarching objective was to examine the suitability of risk-adjusted TTR as a potential quality indicator.

WHAT IS KNOWN

- The safety and effectiveness of oral anticoagulation can be improved by better control, (ie, more time in therapeutic range [TTR]).
- Although oral anticoagulation is prescribed for millions of patients each year, there has been no organized approach to measuring or improving the quality of oral anticoagulation.

WHAT THE STUDY ADDS

- We used clinic-level risk-adjusted TTR to profile the performance of 100 anticoagulation clinics in an integrated system of care (the Veterans Health Administration).
- We propose the use of risk-adjusted TTR as a quality indicator to measure and track the quality of oral anticoagulation in the Veterans Health Administration and other integrated health systems.
- Quality measurement in oral anticoagulation is a necessary prerequisite to quality improvement, which holds the promise of preventing adverse events due to inadequate or excessive anticoagulation.

Methods

Patients

The database for this study also has been described elsewhere.[15] The VA is the largest integrated health system in the United States, and for many years, has collected comprehensive data regarding the care delivered to its patients, including inpatient care, outpatient care, and pharmacy records. The Veterans Affairs Study to Improve Anticoagulation included all patients deemed to be receiving oral anticoagulation therapy from the VA between October 1, 2006, and September 30, 2008, based on the criteria described later. The study was approved by the Institutional Review Board of the Bedford VA Medical Center.

A flowchart of study inclusion criteria is shown in Figure 1. We included all patients who received warfarin from the VA during the 2-year study period (ie, at least 30 days' worth dispensed by the pharmacy) and who had at least 2 valid intervals for calculating TTR.[16] For this purpose, a valid interval consists of 2 international normalized ratio (INR) values separated by ≤56 days without an intervening hospitalization.

There were 2 levels of exclusions: individual patients and entire sites of care. On the patient level, we excluded patients whose primary indication to receive warfarin was valvular heart disease. Many such patients have a target INR range of 2.5 to 3.5 rather than the more standard 2 to 3, but it is not possible to determine with certainty which patients have the higher target range. Without specific knowledge of the target range, we cannot calculate TTR. We also excluded patients whose only recorded INR values were ≤1.2, reasoning that most such patients received INR tests for reasons unrelated to warfarin management (eg, frequent emergency department visits).

Sites of Care

There are 128 sites of care within the VA, each of which includes a hospital, an outpatient care center, and several outlying community-based clinics. Each site has a specialized anticoagulation clinic, which usually is run by clinical pharmacists under the supervision of a medical director.[17] Therefore, essentially all patients whose anticoagulation is managed in the VA are managed by specialized anticoagulation clinics. Self-testing devices are not covered by the VA; therefore, few patients have or use them.

We excluded 28 sites from our study and several months of data from an additional 14 sites because our data-checking procedures revealed possible problems with data completeness at those sites. The problem with data completeness relates to the laboratory data only. Although accurate data are collected about which laboratory tests are drawn (because something akin to a billing code is generated), the data regarding laboratory results must be checked carefully. Specifically, the name given to each laboratory test by the local facility is not uniform throughout the system, and these names may change unexpectedly; after this happens, there may be a period of several months where the local laboratory results are not captured by the national database until the name change is noted. We identified which sites had this issue by dividing the data into 3-month periods; problematic sites had few or no INR results in certain periods, whereas the number of INR tests drawn remained constant over time. In contrast, there were 86 sites that had complete data for the entire 2-year study period and 14 sites that began to have complete data during the period (usually early in the study) and continued to have it through to the end. Thus, 28 sites were excluded because of incomplete data, and 14 sites were partially included. The 14 partially included sites performed similarly to the 86 sites with complete data, suggesting that they differed only in data collection rather than in performance (results not shown). Most patients only visited 1 site of care, and their INR data were assigned to that site. If a patient visited >1 site (3% of patients), we partitioned his or her data by site.

Laboratory Values and Calculation of TTR

We included INR values within the VA system when patients were on warfarin, that is, when a patient was either (1) in possession of warfarin or (2) having INR tests at least every 42 days. We defined the period of warfarin possession as the duration of the most recent VA prescription for warfarin plus 30 days. Because patients may be instructed to take half-doses of warfarin, we recognize that going >30 days beyond the end of a prescription does not necessarily

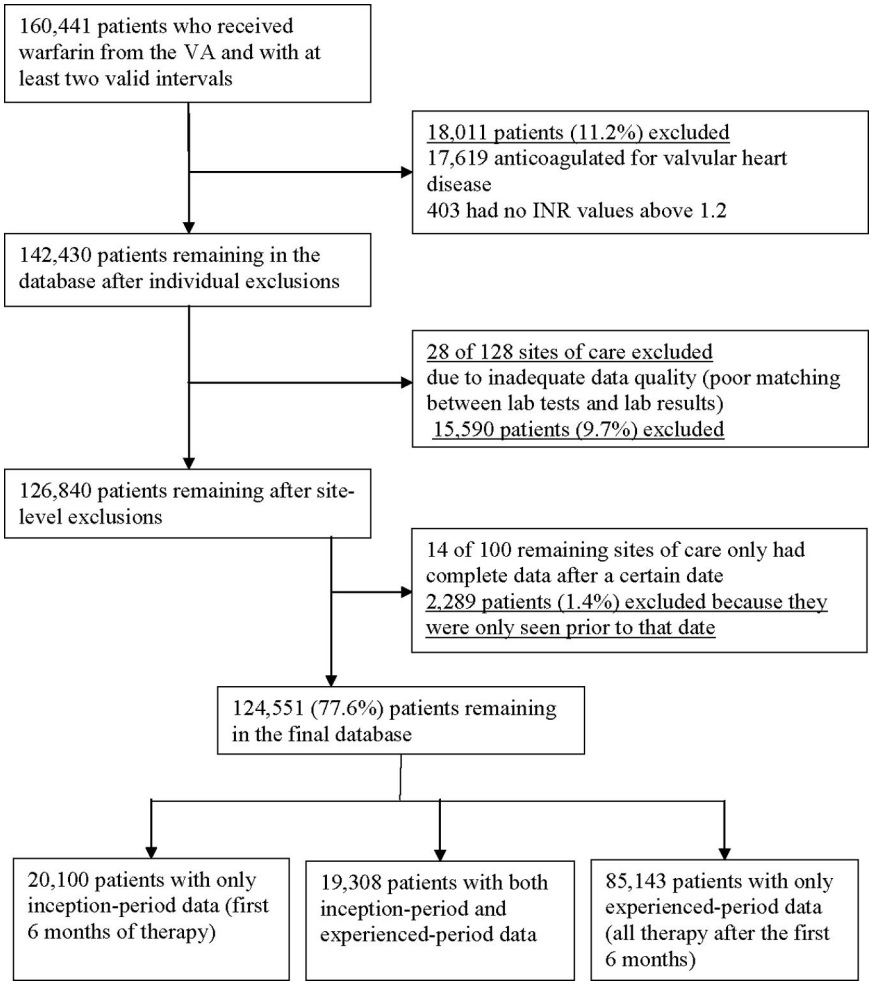

Figure 1. Flowchart of included and excluded patients.

indicate that warfarin therapy has stopped. Therefore, we also allowed a consistent pattern of INR measurements (ie, every 42 days or fewer) to indicate that a patient was still being managed. A similar approach was used to define time on warfarin in a previous study.[18]

We excluded INR tests measured while the patient was hospitalized within the VA system. Patients who are hospitalized may receive temporary parenteral anticoagulation (eg, with heparin) or no anticoagulation, so out-of-range INR values while hospitalized may be intentional and do not necessarily reflect poor quality of care.

We calculated TTR using the Rosendaal method,[16] which uses linear interpolation to assign an INR value to each day between successive observed INR values. Gaps of ≥56 days between INR values are not interpolated. After interpolation, the percentage of time during which the interpolated INR values lie between 2.0 and 3.0 (from 0% to 100%) is calculated.[16]

Risk Adjustment Model

We have previously described the derivation and validation of our risk adjustment model for TTR.[15] We considered many potential variables that we believed were likely to affect TTR, including demographics, area-level poverty, driving distance to care, physical health conditions, mental health conditions, number of medications, and number of hospitalizations. Most variables were retained within the model, with the exception of several comorbid conditions that did not have appreciable effect sizes. The model was derived and validated according to customary procedures, which included considerations of maximizing predictive ability, clinical credibility, and ease of use and understanding.[15] The Table contains all the variables that were retained in the final model.

Inception and Experienced Management of Warfarin

We calculated TTR and built separate risk adjustment models for 2 time periods—the first 6 months of therapy (inception period) and all care after 6 months (experienced period)—because these 2 periods are qualitatively different with regard to anticoagulation management. We defined each patient's date of warfarin inception, looking back as far as October 1, 2005. Inception was defined as the first INR value >1.2 or the first outpatient warfarin fill, whichever came first. It would be extremely unusual for a patient to record an INR value >1.2 unless he or she had taken warfarin. We then divided the sample into inception time (the first 6 months of warfarin therapy for each patient) and experienced time (any time thereafter). A single patient might contribute only to the inception data set (if he or she had <6 months of therapy), only to the experienced data set (if he or she began warfarin >6 months before the inception of our study), or both. Coefficients for the 2 models differed substantially,[15] confirming that the 2 periods should be risk adjusted separately.

Statistical Analyses

First, we calculated TTR for each patient during the inception and experienced periods. We then calculated a mean observed TTR (O) for each site and period. We applied our risk adjustment models to calculate the expected (E) TTR first for each patient and then for each site on the basis of the patient population at that site. We then calculated an observed minus expected (O-E) score for each site for the 2-year study period. Each site had 2 O-E scores: 1 for inception management and 1 for experienced management. Site O-E scores for the 2 time periods were highly correlated, so we combined them into

Table. Baseline Sample Characteristics for the Overall Sample, the Inception Cohort (ie, First 6 Months of Anticoagulation Therapy), and the Experienced Cohort (All Management Thereafter)

Variable	Overall Sample (n=124 551)	Inception Cohort (n=39 408)	Experienced Cohort (n=104 451)
Female sex	2589 (2.1)	1046 (2.7)	1983 (1.9)
Age, y	72 (61–79)	66 (59–76)	72 (62–79)
Race/ethnicity			
Non-Hispanic white	95 312 (76.5)	29 106 (73.9)	80 682 (77.2)
Non-Hispanic black	11 240 (9.0)	4455 (11.3)	8847 (8.5)
Hispanic	3570 (2.9)	1148 (2.9)	2976 (2.8)
Asian	368 (0.3)	134 (0.3)	302 (0.3)
Native American	360 (0.3)	148 (0.4)	279 (0.3)
Other/unknown	13 701 (11.0)	4417 (11.2)	11 365 (10.9)
Poverty in zip code of residence, %	10.7 (6.6–16.0)	10.9 (6.7–16.4)	10.7 (6.6–15.9)
Distance from nearest VA facility, miles	7.8 (3.7–16.5)	7.6 (3.6–15.9)	7.8 (3.7–16.5)
Primary indication for warfarin*			
Atrial fibrillation	76 894 (61.7)	21 568 (54.7)	67 045 (64.2)
Venous thromboembolism	36 402 (29.2)	13 933 (35.4)	28 567 (27.3)
All others combined	11 255 (9.0)	3907 (9.9)	8839 (8.5)
Physical comorbid conditions			
Cancer (newly diagnosed)	9236 (7.4)	3938 (10.0)	7091 (6.8)
Chronic kidney disease	17 333 (13.9)	5224 (13.3)	14 804 (14.2)
Chronic liver disease	1584 (1.3)	565 (1.4)	1253 (1.2)
Chronic lung disease	36 138 (29.0)	11 006 (27.9)	30 666 (29.4)
Diabetes	48 874 (39.2)	14 420 (36.6)	41 842 (40.1)
Epilepsy	3485 (2.8)	1090 (2.8)	2926 (2.8)
Heart failure	38 860 (31.2)	10 129 (25.7)	34 208 (32.8)
Hyperlipidemia	91 844 (73.7)	26 957 (68.4)	78 711 (75.4)
Hypertension	103 300 (82.9)	31 343 (79.5)	87 733 (84.0)
Mental comorbid conditions			
Alcohol abuse	12 695 (10.2)	5338 (13.5)	9719 (9.3)
Bipolar disorder	3109 (2.5)	1278 (3.2)	2382 (2.3)
Dementia	6356 (5.1)	1723 (4.4)	5511 (5.3)
Major depression	27 432 (22.0)	9380 (23.8)	22 569 (21.6)
Substance abuse (nonalcohol)	5827 (4.7)	2673 (6.8)	4228 (4.0)
No. of medications	8 (6–12)	…	…
Hospitalized at least once	34 328 (27.6)	…	…

Data are presented as no. (%) or median (interquartile range).

*Patients whose main indication for anticoagulation was valvular heart disease or prosthetic heart valve were excluded from this study.

a single O-E score for each site, weighting by the number of patients in each period.

A site was considered a high or low outlier performer if its combined O-E score deviated from 0 by ≥5% and if the O-E score differed from 0 at the 0.01 level of significance (using *Z* scores). We used the 0.01 level of significance as opposed to the traditional 0.05 to address issues of multiple testing. A 5% absolute difference in TTR has been cited by other studies as constituting a meaningful difference in performance and is probably the agreed-on standard for a clinically important difference.[19]

We also investigated whether empirical Bayesian methods would change determinations of outlier status.[20] This technique has been shown to alter results regarding outlier status, particularly when sample sizes are small.[21] Typically, empirical Bayesian estimation shrinks estimates for the observed site mean TTR toward the overall mean, and fewer sites are outliers when analyzed by such methods. However, we observed almost no shrinkage presumably because of the large sample sizes even at the smaller sites of care. Therefore, we concluded that Bayesian methods were not necessary for this situation and relied on *Z* scores to determine outliers.

We compared site rankings before and after risk adjustment to determine its effect. We calculated the proportion of the variance in TTR that was explained by patient characteristics, site-level variability, and both together. We also divided the study period in half (ie, fiscal year 2007 [FY07] versus FY08) to evaluate the stability of site O-E scores and rankings between the 2 years. Analyses were conducted using SAS version 9.1 and R version 2.8 statistical software. Dr Rose had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Results

Patient Population

We studied 124 551 unique patients who received anticoagulation from 100 sites of care (Figure 1). Baseline characteristics for the sample are described in the Table. The sample was mostly men (98%) with an average age of 72 years. Most (62%) patients were anticoagulated for atrial fibrillation, with the remainder anticoagulated for venous thromboembolism (29%) or other indications (9%) (eg, mural thrombus, cardiomyopathy, pulmonary hypertension). The population had a substantial burden of comorbidity, for example, 39% had diabetes, 31% had heart failure, 14% had chronic kidney disease, and 7% had received a new diagnosis of (nonskin) cancer during the study period. The burden of mental illness and substance abuse also was considerable: 22% had major depression, 10% had received a diagnosis of alcohol abuse, and 5% had dementia. As would be expected with a population of sicker patients, they received many medications (median of 8), and 28% were hospitalized at least once during the 2-year study period.

Site Performance

The median number of patients per site was 1584 (range, 103 to 5103; interquartile range, 1043 to 2345). Mean site-expected TTR during inception was based on patient population at each site and varied from 44% to 50%. Mean site-expected TTR during the experienced period varied from 58% to 65%. We examined case mix at the 5 sites with the lowest predicted TTR (difficult case mix) and the 5 sites with the highest predicted TTR (easy case mix). The difficult sites treated more minority patients and more patients living in high-poverty zip codes. Patients at the difficult sites had higher rates of almost every comorbid condition that is associated with lower TTR in our database,[15] often several-fold higher. Finally, patients at the difficult sites received more medications and were more likely to be hospitalized. In short, there was no single factor that explained these differences in case mix; all the variables appeared to play a role. These between-site differences in case mix will be addressed more fully in a separate article.

During the inception period, the mean observed TTR was 48%; site mean observed TTR varied from 30% to 59% during inception. Site O-E scores for inception, which measured the difference between observed and expected TTR, ranged from −17% to +13%. Of the 100 sites of care, during the inception period, there were 12 high outliers, indicating performance at least 5% better than expected and O-E different from 0 at the 0.01 level of significance. There were 12 low outliers, indicating performance at least 5% worse than expected and O-E different from 0 at the 0.01 level of significance. During the experienced period, the mean observed TTR was 61%; site mean observed TTR varied from 41% to 72%. Site O-E scores for the experienced period ranged from −19% to +12%. Of the 100 sites of care, during the experienced period, there were 14 high outliers and 10 low outliers.

The correlation between site O-E scores for the inception and experienced periods was high ($r=0.89$; $P<0.001$) (Figure 2), suggesting that they measure a similar construct and may be combined. Mean combined TTR for the entire sample was 58%, and sites varied from 38% to 69%. Site-expected TTR ranged from 54% to 62%, and O-E scores ranged from −18% to +12%. Of the 100 sites of care, there were 13 high outliers and 10 low outliers. The 100 sites of care in our study are shown in Figure 3; a tabular form of these results appears in the online-only data supplement. The overall performance of our risk adjustment model was as follows: Patient-level characteristics alone explained 13.3% of the variability in TTR, site-level variability alone explained 2.9%, and both together explained 15.9%.

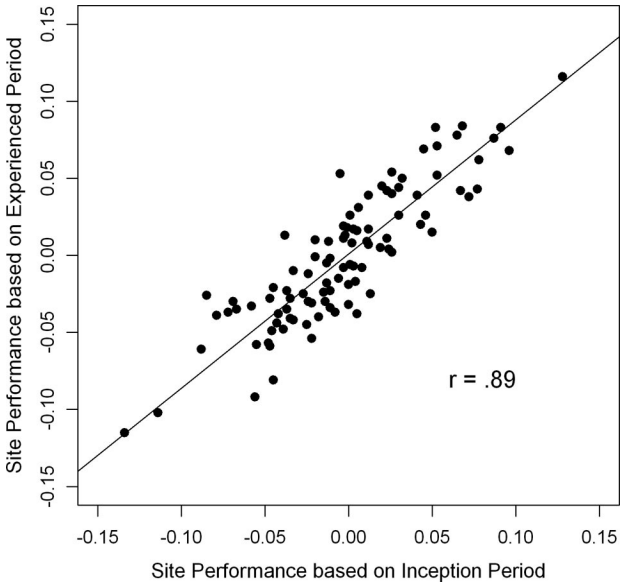

Figure 2. Comparison of site performance between the inception period (first 6 months of therapy) and the experienced period (all therapy thereafter). Performance is measured by the O-E score, which compares the TTR achieved by each site to the TTR that would be expected based on its patient population. Possible O-E scores range from −1 to +1, with positive scores indicating better-than-expected performance. Performance in the 2 periods was highly correlated ($r=0.89$; $P<0.001$).

Risk adjustment changed site rankings only slightly; there was a high degree of correlation between site rankings before and after risk adjustment ($r=0.93$; $P<0.001$). However, risk adjustment made an important difference for several sites of care. For example, site FC was ranked 27th of 100 before risk adjustment (observed TTR, 60.6%, or 2.7% above average). However, site FC had a challenging patient population (expected TTR, 53.5%, or 4.4% below average). Therefore, after risk adjustment, site FC was revealed to be one of the best performers (O-E, +7.0%; adjusted rank, 7th) as well as a high outlier. In contrast, site GI was ranked 95th before risk adjustment (observed TTR, 50.2%, or 7.7% below average) and would have been a low outlier using a definition based on unadjusted results. However, this site also had a challenging patient population (expected TTR, 53.9%, or 4% below average); although the risk-adjusted performance was poor (O-E, −3.7%; adjusted rank, 80th), it did not meet our definition for an outlier site. Not all sites were helped by risk adjustment, of course; for some sites, it invited a harsher view of their performance in light of their relatively easy case mix.

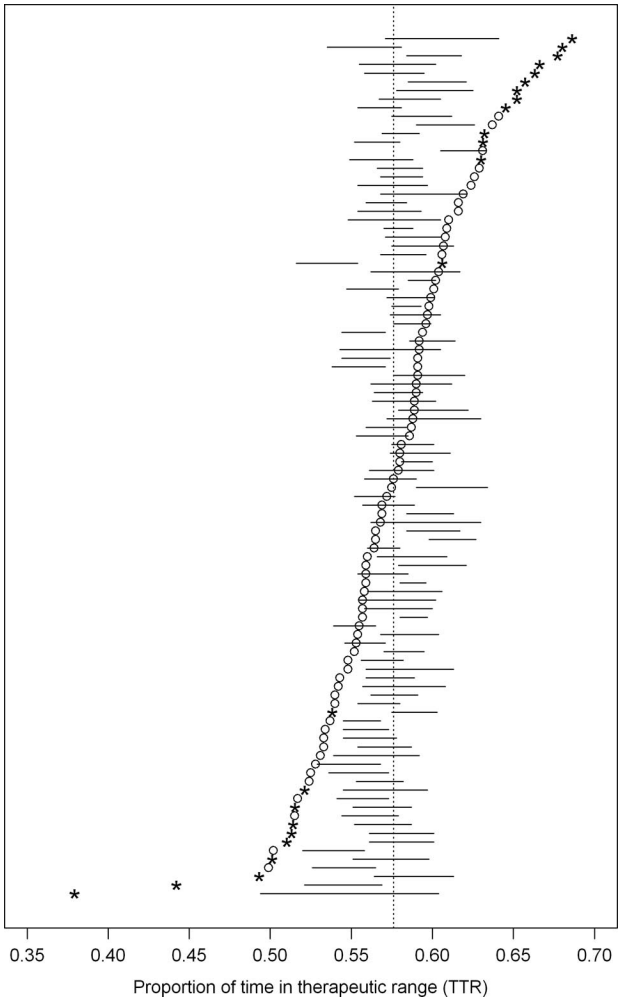

Figure 3. Overall site performance as measured by observed mean site TTR minus expected mean site TTR. The vertical line indicates the overall mean TTR (≈58% time in range). Each horizontal bar indicates the 99% CI around a site's expected mean TTR, whereas each ○ indicates the observed mean site TTR. For outlier sites (observed TTR differs from expected by ≥5%), the ○ is replaced by an *.

For example, site JC was ranked 79th before risk adjustment (observed TTR, 53.8%, or 4.1% below average), but because of its relatively easy case mix (expected TTR, 58.9%), the site was an outlier (O-E, −5.0%; adjusted rank, 92nd).

Stability of Rankings Between Years

We also divided our study period in half and compared the 2 years to each other. Most (62%) patients had INR values in both years, but 18% were only followed in FY07 and 20% only in FY08. We examined the stability of site performance between the 2 years. There was a high correlation between site performance in FY07 and FY08 both by O-E score and by risk-adjusted site rankings (r=0.89 and 0.88, respectively; P<0.001 for both) (Figure 4). Additionally, outliers were relatively consistent between years. For example, of the 9 high outliers in FY07 (year 1 of the study), 8 sites were also high outliers in FY08, whereas of the 8 sites that were low outliers in FY07, 6 were also low outliers in FY08.

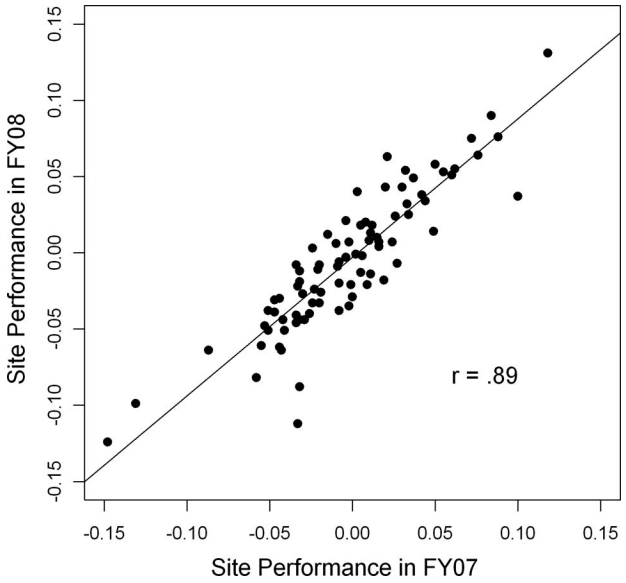

Figure 4. Comparison of site performance between FY07 and FY08. Performance is measured by the O-E score, which compares the TTR achieved by each site to the TTR that would be expected based on its patient population. Possible O-E scores range from −1 to +1, with positive scores indicating better-than-expected performance. Performance in the 2 years was highly correlated (r=0.89; P<0.001).

Discussion

In this study, we demonstrate that risk-adjusted TTR can be used as a quality indicator for oral anticoagulation care. Within the VA system, risk-adjusted TTR varies widely, with some sites performing as much as 18% below or 12% above expected. Risk-adjusted TTR is feasible to measure and is relatively consistent from year to year, suggesting that it is measuring an aspect of quality of care that is stable over time. This measure could be used by the VA or other integrated systems of care to profile annual performance and serve as an aid and impetus for quality improvement.

Risk adjustment is important for enhancing the credibility of site profiling; without risk adjustment, sites could claim that their poor performance was solely because of their case mix.[14] However, we found that the range of case mix was much smaller than the range of observed performance and that sites with very difficult or very easy patient populations were found equally among the best- and worst-performing sites (Figure 3), suggesting that although case mix varies among VA sites, the quality of oral anticoagulation care delivered to those patients varies even more widely.

We observed wide variations in performance within the VA system, from very-low TTR (38%) to very high (69%). However, after risk adjustment, site-site differences only accounted for 2.9% of the variability in TTR. Although this finding may sound like a small proportion of variability to explain, it is actually quite similar to the findings of previous studies. For example, using a population of patients with diabetes, Hofer et al[13] found that after risk adjustment, ≤4% (and as little as 1%) of the variability in hospitalization rates, visit rates, laboratory utilization rates, and glycemic control was attributable to differences between physicians. In discussing this result, these authors suggested that standardizing

processes of care among providers may not suffice to produce the quality improvement we seek; it may also be necessary to improve process of care on a system level. We therefore plan to use the present study as a basis on which to build a comprehensive program of quality measurement and quality improvement in anticoagulation care.[9] Our goal would be to improve the mean TTR in the VA to at least 70% (from the current 58%). Achieving this goal will require a focus on system-level approaches to improve processes of care. For example, our group[22] has shown that more judicious decisions regarding when to change the dose of warfarin can improve TTR considerably. Addressing this and other processes of care in a systematic way could greatly improve TTR in the VA; however, continually measuring risk-adjusted TTR will be a precondition to any program of quality improvement.[9]

Future studies also could expand this work outside the VA setting. A system of quality measurement, such as the one described in this article, would be relatively easy to implement in any integrated health system with a well-developed electronic medical record. However, the majority of patients are cared for in smaller community practices, not integrated health systems. The benefits of quality measurement could be extended to even more patients by incorporating risk-adjusted TTR into existing, commercially available anticoagulation management software, allowing sites to not only track their performance, but also potentially compare it to other sites after adjustment for case mix.

This study has important strengths, including the fact that we profiled performance among 100 sites of care within the nation's largest integrated healthcare system. In addition, the careful construction of the risk adjustment model and the considerable detail contained within its predictor variables improve our confidence that we have adjusted for case mix and that the residual variability in site performance is attributable to the quality of care.

As with any study, we acknowledge limitations. Our risk adjustment model did not include some factors that contribute to variability in TTR, especially diet and adherence to therapy. However, it could also be argued that dietary variation and medication nonadherence potentially are remediable, so adjusting for these factors could erase an important opportunity for sites to improve TTR. Second, our analyses combined patients with different indications for therapy, especially atrial fibrillation and venous thromboembolism. Although our risk adjustment model included a variable for indication, we were concerned that these groups might be too different to be combined in 1 model. To address this concern, we ran our analyses separately for patients with each indication; site performance by O-E scores did not change appreciably. Third, because of uncertainty regarding the target INR range, our study did not include patients anticoagulated for valvular heart disease, including prosthetic heart valves. However, we have every reason to expect that sites that provide superior care for patients with other indications for therapy also would provide superior care for patients with valvular heart disease. Fourth, because of strict requirements for completeness of data, we were unable to profile performance at some sites of care. Fifth, it is likely that some of the patients in our inception cohort had previously received warfarin outside the VA. If these patients were excluded from the inception cohort, we might have observed lower inception TTR; however, this is unlikely to affect site-site comparisons. Finally, although we examined TTR as a quality indicator, other summary measures of INR control might be used. For example, proportion of INR values in range could allow for more-frequent feedback about performance than TTR, which can only be calculated over a several-month period. Monthly proportion of values in range could therefore be used to provide real-time performance feedback, enhancing quality improvement efforts. In the future, we plan to compare TTR with other measures.

In summary, this article establishes that risk-adjusted TTR can be used to profile anticoagulation care in a large integrated healthcare system. Patients deserve to receive the very best care, particularly when the stakes are as high as they are with warfarin, but our report demonstrates that some patients are receiving much better care than others. It is time to shine the light of performance measurement on oral anticoagulation care.

Acknowledgments

We thank Barbara G. Bokhour, PhD, for her contributions to this manuscript.

Sources of Funding

This study was funded by the US Department of Veterans Affairs, Health Services Research and Development Service (Career Development Award to Dr Rose). The sponsor had no role in the design and conduct of the study; the collection, management, analysis, and interpretation of the data; or the preparation, review, and approval of the manuscript.

Disclosures

Dr Rose is supported by a Career Development Award from the Veterans Affairs Health Services Research and Development Service. Dr Hylek has received honoraria from Bayer and Bristol Myers Squibb and has served on advisory boards for Boehringer Ingelheim, Bristol Myers Squibb, Merck, and Sanofi Aventis.

References

1. Ansell J, Hirsh J, Hylek E, Jacobson A, Crowther M, Palareti G. Pharmacology and management of the vitamin K antagonists: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th edition). *Chest.* 2008;133:160S–198S.
2. Budnitz DS, Shehab N, Kegler SR, Richards CL. Medication use leading to emergency department visits for adverse drug events in older adults. *Ann Intern Med*. 2007;147:755–765.
3. Connolly SJ, Pogue J, Eikelboom J, Flaker G, Commerford P, Franzosi MG, Healey JS, Yusuf S. Benefit of oral anticoagulant over antiplatelet therapy in atrial fibrillation depends on the quality of international normalized ratio control achieved by centers and countries as measured by time in therapeutic range. *Circulation*. 2008;118:2029–2037.
4. van Leeuwen Y, Rosendaal FR, Cannegieter SC. Prediction of hemorrhagic and thrombotic events in patients with mechanical heart valve prostheses treated with oral anticoagulants. *J Thromb Haemost*. 2008;6:451–456.
5. van Walraven C, Oake N, Wells PS, Forster AJ. Burden of potentially avoidable anticoagulant-associated hemorrhagic and thromboembolic events in the elderly. *Chest*. 2007;131:1508–1515.
6. White HD, Gruber M, Feyzi J, Kaatz S, Tse HF, Husted S, Albers GW. Comparison of outcomes among patients randomized to warfarin therapy according to anticoagulant control: results from SPORTIF III and V. *Arch Intern Med*. 2007;167:239–245.
7. Veeger NJ, Piersma-Wichers M, Tijssen JG, Hillege HL, van der Meer J. Individual time within target range in patients treated with vitamin K antagonists: main determinant of quality of anticoagulation and predictor

of clinical outcome. A retrospective study of 2300 consecutive patients with venous thromboembolism. *Br J Haematol*. 2005;128:513–519.
8. van Walraven C, Jennings A, Oake N, Fergusson D, Forster AJ. Effect of study setting on anticoagulation control: a systematic review and metaregression. *Chest*. 2006;129:1155–1166.
9. Rose AJ, Berlowitz DR, Frayne SM, Hylek EM. Measuring quality of oral anticoagulation care: extending quality measurement to a new field. *Jt Comm J Qual Patient Saf*. 2009;35:146–155.
10. Brook RH, McGlynn EA, Cleary PD. Quality of health care. Part 2: measuring quality of care. *N Engl J Med*. 1996;335:966–970.
11. Donabedian A. *Explorations in Quality Assessment and Monitoring.* Ann Arbor, MI: Health Administration Press; 1980.
12. Ogilvie IM, Newton N, Welner SA, Cowell W, Lip GY. Underuse of oral anticoagulants in atrial fibrillation: a systematic review. *Am J Med.* 2010;123:638–645, e634.
13. Hofer TP, Hayward RA, Greenfield S, Wagner EH, Kaplan SH, Manning WG. The unreliability of individual physician "report cards" for assessing the costs and quality of care of a chronic disease. *JAMA*. 1999;281: 2098–2105.
14. Iezzoni LI, ed. *Risk Adjustment for Measuring Health Care Outcomes.* 3rd ed. Chicago, IL: Health Administration Press; 2003.
15. Rose AJ, Hylek EM, Ozonoff A, Ash AS, Reisman JI, Berlowitz DR. Patient characteristics associated with oral anticoagulation control: results of the Veterans AffaiRs Study to Improve Anticoagulation (VARIA). *J Thromb Haemost*. 2010;8:2182–2191.
16. Rosendaal FR, Cannegieter SC, van der Meer FJ, Briet E. A method to determine the optimal intensity of oral anticoagulant therapy. *Thromb Haemost*. 1993;69:236–239.
17. Department of Veterans Affairs, Veterans Health Administration. VHA Directive 2009–003. "Anticoagulation Therapy Management." February 2, 2009.
18. Go AS, Hylek EM, Chang Y, Phillips KA, Henault LE, Capra AM, Jensvold NG, Selby JV, Singer DE. Anticoagulation therapy for stroke prevention in atrial fibrillation: how well do randomized trials translate into clinical practice? *JAMA*. 2003;290:2685–2692.
19. Samsa GP, Matchar DB. Relationship between test frequency and outcomes of anticoagulation: a literature review and commentary with implications for the design of randomized trials of patient self-management. *J Thromb Thrombolysis*. 2000;9:283–292.
20. Spiegelhalter DJ, Abrams KR, Myles JP. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation.* West Sussex, UK: Wiley; 2004.
21. Berlowitz DR, Christiansen CL, Brandeis GH, Ash AS, Kader B, Morris JN, Moskowitz MA. Profiling nursing homes using Bayesian hierarchical modeling. *J Am Geriatr Soc*. 2002;50:1126–1130.
22. Rose AJ, Ozonoff A, Berlowitz DR, Henault LE, Hylek EM. Warfarin dose management affects INR control. *J Thromb Haemost*. 2009;7:94–101.