Quality of Stroke Prevention Care in Atrial Fibrillation

Many Moving Targets

Mintu P. Turakhia, MD, MAS

In atrial fibrillation (AF), the loss of coordinated electromechanical atrial activity predisposes to impaired atrial emptying, stasis of blood, and a prothrombotic state. These factors cause blood clots to form in the heart and embolize systemically to cause stroke or other organ failure. AF is responsible for 15% of 700,000 strokes in the United States each year, with an estimated annual direct and indirect cost of $57.9 billion in the United States alone.

To prevent stroke or systemic embolism in AF, patients and clinicians have long had only 2 choices: (1) warfarin anticoagulation or (2) aspirin or other antiplatelet regimens. Because of warfarin’s narrow therapeutic window and substantial risk of hemorrhage with overanticoagulation or in vulnerable patients, a large clinical challenge has been the identification of patients in whom the benefits of anticoagulation outweigh the risks.

To simplify and standardize clinical decision-making, risk prediction rules such as CHADS2 were developed to create a stroke risk score for AF (1 point each for heart failure, hypertension, age ≥75, diabetes; 2 points for prior stroke or transient ischemic attack). Subsequently, society care guidelines established recommendations for warfarin anticoagulation based on these scores, generally advocating warfarin with 2 or more risk factors, and recommending aspirin with zero or 1 risk factor. More elaborate risk prediction schemes that improve discrimination for stroke prediction, estimate risk of hemorrhage, and assess “net clinical benefit” have been developed and endorsed. However, their complexity may limit use in clinical practice. Warfarin prescription remains underused even in moderate-to-high-risk patients, which may indicate barriers to acceptance and diffusion of practice guidelines.

Most notably, warfarin decision-making has been based strictly on patient-level factors, despite the fact that therapy itself requires ongoing care, monitoring, and intervention from a health care center or provider. For major cardiovascular procedures such as coronary artery bypass grafting, patients and referring providers may scrutinize operator or facility complication rates, many of which are publicly reported, but few would consider the anticoagulation performance of a facility, even though the lifetime risk of adverse events on warfarin exceeds complication rates of most cardiovascular procedures.

Risk-Adjusted Quality of Warfarin Anticoagulation

In this issue of Circulation: Cardiovascular Quality and Outcomes, Rose et al challenge the conventional decision-making framework. In an elegant study using national Veterans Health Administration (VA) data, the investigators profiled 100 VA facilities caring for 125,000 patients on warfarin for AF and other nonvalvular indications. They measured quality of anticoagulation care using a process measure to assess whether warfarin is managed within the therapeutic window. For each patient, the percentage of time in international normalized ratio (INR) therapeutic range (TTR) of 2.0 to 3.0 was calculated and aggregated to the facility level to create a facility average. TTRs for patients and facilities were calculated separately for the first 6 months after new warfarin initiation and for chronic therapy.

There are several notable findings. First, there is remarkably wide variation in TTR across facilities, even in a health care system that has implemented a system-wide network of specialized anticoagulation clinics. Although TTR >60% is a generally accepted threshold above which warfarin confers significant benefit compared with antiplatelet therapy, facility TTRs varied from 38% to 69%. Such variation is not unique to VA or single-care systems. Randomized trials, which typically represent ideal care in motivated patients, have shown wide site variation of TTR. With 18,113 patients from 951 international sites, the RE-LY study, comparing warfarin with dabigatran, is one of the largest warfarin trials to date. Still, facility TTR averages varied from 57% to 72%, resulting in a gradient of risk of stroke and hemorrhage across facility TTR quartiles. However, the RE-LY sub-study modeled only at the facility level, and therefore variability of TTR variation explained by patient-level factors was not evaluated.

The Rose report addresses this critical knowledge gap by calculating risk-adjusted TTR with covariates including patient-level demographics, distance, health conditions, medications, and hospitalizations. After adjustment, site-expected TTR improved and variation decreased (54% to 62%). Risk adjustment only slightly affected rankings but appeared to reclassify sites at the extremes of case mix. Notably, site-level differences accounted for only 2.9% of variability of
TTR, as the variation of case mix was much smaller than observed patient TTR variation.

The study demonstrates that facility profiling based on risk-adjusted TTR rankings is feasible and could be used to profile performance. Certainly, this approach appears to offer more granularity than the current performance measure of a minimum frequency of INR monitoring (usually every 30 days)\textsuperscript{15}. INR frequencies may be sufficient and still fail to achieve adequate TTR. This could be important because process measures of frequency benchmarking may fail to capture quality of warfarin dose adjustment care or facility effects on patient compliance. However, our work using VA data has shown that facility mean INR frequencies account for most of the variation ($r^2=0.29$) of facility-level TTR.\textsuperscript{16} There also appears to be no volume-outcome relationship to TTR, which confirms our prior observations.\textsuperscript{17}

Several points should be considered when interpreting this study. First, although TTR does predict stroke and hemorrhage outcomes, the adjusted TTR model was developed using TTR as the outcome. Therefore, the magnitude of contribution of patient-level adjusted TTR to stroke, hemorrhage, or mortality outcomes is uncertain. Second, patients with infrequent INR monitoring (>56 days apart) were excluded from analysis because TTR could not be reliably estimated. This can lead to systematic exclusion of patients with poor monitoring, which may itself cause adverse events. Although the exclusion is necessary and appropriate for TTR benchmarking, cohort exclusions may lead to unevenly applied restriction across facilities or may fail to adequately capture important components of the structure-process-outcome relationship.\textsuperscript{18} Therefore, the relationship of unadjusted and adjusted TTR and INR frequencies to clinical outcomes should be rigorously reexamined to confirm construct validity. Finally, certain adjustment variables, such as distance to facility or zip code poverty, may not be available or easily analyzable by other care systems. Therefore, the impact of missingness of risk adjustment variables on model performance requires evaluation before widespread implementation.

**Clinical Decision-Making in the Context of New Stroke Prevention Therapies**

The decision to prescribe warfarin is most difficult in patients not at the margins of stroke or bleeding risk, but where both are matched. For example, in a 51-year-old patient with a CHADS2 score of 1, will warfarin prevent stroke more effectively than aspirin, or is aspirin sufficient and also less likely to cause hemorrhage? Should a 77-year-old patient with prior major gastrointestinal hemorrhage with an INR of 3.2 on warfarin and CHADS2 score of 4 have warfarin restarted? New drug therapies, which have started to enter the market, may offer unique safety and efficacy profiles more appropriate for these clinical scenarios.\textsuperscript{19} Moreover, differences in expected stroke rates on warfarin therapy due to quality (TTR), secular trends, or other factors could also influence treatment thresholds for warfarin versus other agents.

To explore these issues, Eckman et al created a Markov decision model to examine 4 stroke prevention strategies: warfarin, aspirin, no therapy, or dabigatran as a prototypical new agent.\textsuperscript{20} Using a base case of a 69-year-old man with AF and CHADS2 score of 2, the investigators found that warfarin had higher expected quality-adjusted survival than aspirin or no therapy. However, when stroke risk is less than 1.7% per year, aspirin resulted in greater quality-adjusted survival, largely the result of offsetting risk of hemorrhage with warfarin. However, when adding dabigatran or a “newer, safer agent” to the model, the new drug had better quality-adjusted survival at a relatively low stroke risk of more than 0.9% per year. These findings support the concept that agents with comparable efficacy to warfarin but lower bleeding risk are likely to yield greater benefits across a large spectrum of AF patients, from very low to high stroke risk.

These findings are similar to a recently published cost-effectiveness analysis from our group.\textsuperscript{21} With a slightly younger base case patient and CHADS2 score of 1 to 2, our calculated quality-adjusted life years were comparable to the Eckman study, with dabigatran edging out warfarin by a similar margin. However, warfarin had higher quality-adjusted survival when intracranial hemorrhage rates decreased to less than 0.6% per year. As the projected intracranial hemorrhage risk was increased, dabigatran dominated regardless of stroke rates on warfarin. We also found that compared with warfarin, dabigatran was likely to be cost-effective under a broad range of assumptions.

Model-based approaches have several limitations. Analyses rely on a series of trial and observational studies to derive model inputs for event rates. These inputs may carry significant uncertainty, especially in the absence of direct comparisons, such as dabigatran to antiplatelet therapy. Relatively short trial follow-up of several years’ duration must be extended to longer time horizons, which requires assumptions on the durability of efficacy. Acceptance of model structure and core assumptions may lead to structural uncertainty, which could dramatically affect results even if a model appears conceptually sound. For example, it is not apparent if the increased risk of myocardial infarction with dabigatran compared with warfarin observed in RE-LY was incorporated into the model.\textsuperscript{22} Our work has shown that compared with warfarin, high-dose dabigatran averted 1000 intracranial hemorrhages and 600 ischemic strokes but resulted in 400 additional myocardial infarctions in a simulated Markov cohort of 10 000 patients.\textsuperscript{21} The benefits and harms of adding antiplatelet to dabigatran therapy are still not well described.

Still, the consistent findings across these 2 studies with independently derived models demonstrate reproducibility and stability of results. There are, however, a few points worth noting. Only high-dose dabigatran was modeled in the Eckman study. Although RE-LY studied low- and high-dose formulations,\textsuperscript{12} the Food and Drug Administration only approved the high-dose form in the United States. At low rates of stroke but average to high rates of intracranial hemorrhage, a low-dose formulation (110 mg twice daily) of dabigatran has better quality-adjusted survival than high-dose dabigatran because of lower bleeding risk and comparable stroke prevention in a low-risk patient.\textsuperscript{21} Finally, off-warfarin stroke rates calculated from observational data are highly vulnerable to confounding by indication, in which covariates can influ-
ence both the decision to not prescribe Coumadin and the outcome. Differences in this confounding or propensity to receive warfarin may vary across cohorts, time, or health care systems, which may affect model calibration and explain vastly different observed stroke rates across different cohorts.

Implications
How might these findings inform policy and practice? First, the work by Rose et al should help advance the argument to mandate greater facility accountability for the quality and outcomes of warfarin anticoagulation services. Because of issues of computation, calibration, and need for outcomes validation of risk-adjusted TTR models, it is unlikely that minimum TTR value (adjusted or unadjusted) could be used to establish a minimum standard of care. However, we should remember that whereas the goal of warfarin is to prevent stroke, the goal of an anticoagulation practice is to maintain an INR of acceptable range (2.0 to 3.0). Therefore, ranking facilities based on how well they achieve this goal, adjusted or not, would serve as a potent motivator for quality improvement. The impetus would be strongest if patients and providers had access to quality-based rankings and the ability to decide where they seek anticoagulation services.

Eckman et al22 explore the emerging paradigm of newer and potentially safer therapies using a model-based approach. Their work may challenge the use of CHADS2-based classification for warfarin alternatives because CHADS2-based guidelines implicitly assume a tradeoff between stroke events and hemorrhage in defining treatment thresholds. Device-based therapy, such as percutaneous occlusion of the left atrial appendage, is available in Europe and is in a pivotal phase III trial in the United States. After the device is endothelialized and the appendage is excluded from the systemic circulation, no anticoagulation is required. Therefore, with these devices, the CHADS2 score may be less relevant for decision-making than a bleeding risk score.

In a few years’ time, we can expect to have a multitude of outpatient oral anticoagulation. In AF. Competition is likely to improve efficacy and safety while lowering cost. Since the prevalence of AF is expected to double by 2030,23 the development and diffusion of safer and more effective therapies is likely to have a striking benefit on the public health.

Disclosures
Dr Turakhia was supported by the following grants: Veterans Health Services Research and Development Career Development Award (CDA09027-1) and American Heart Association National Scientist Development Grant (09SDG2250647). Dr Turakhia also provides consultative services on comparative effectiveness research in atrial fibrillation to Precision Health Economics, LLC (amount >$10,000). The content and opinions expressed are solely the responsibility of the authors and do not necessarily represent the views or policies of the Department of Veterans Affairs.

References
Clinical Performance Measures for Atrial Fibrillation) developed in collaboration with the Heart Rhythm Society. 


18. Turakhia MP, Heidenreich PA. Using restriction to minimize bias in observational studies. 


**KEY WORDS:** Editorials • atrial fibrillation • stroke • warfarin • anticoagulation • dabigatran • quality • outcomes • time in therapeutic range