Point of Care Comparative Effectiveness Trial Design for Stroke Prevention in Atrial Fibrillation

Tze Leung Lai[1,2], Mei-Chiung Shih[2,3] and Mintu Turakhia[4]

[1]*Department of Statistics, Stanford University*
[2]*Department of Health Research and Policy, Stanford University*
[3]*VA Palo Alto Cooperative Studies Program Coordinating Center*
[4]*Department of Medicine, Stanford University*

Abstract

The goal of comparative effectiveness trials is to support the evidence-based choice of treatments already in use. We describe the design of a stroke prevention trial that compares the effectiveness of three target-specific oral anticoagulants: dabigatran, rivaroxaban, and apixaban, for the treatment of nonvalvular atrial fibrillation and atrial flutter. In the last four years, three randomized trials have been published that compare these oral anticoagulants that are administered once or twice daily to warfarin. Although the primary outcome of time-to-stroke or systemic embolism was nearly identical in the three trials, there were considerable differences in inclusion and exclusion criteria and there remains considerable uncertainty as to the optimal first-line strategy. There are also increasing concerns about the effectiveness and safety of warfarin and the new oral anticoagulants outside of the trials, particularly in vulnerable patients served by the Veterans Health Administration. These considerations have led to the need for a pragmatic, point-of-care, multi-site clinical trial to compare the effectiveness of the three oral anticoagulants. We describe herein the design of one such trial in the VA Cooperative Studies Program and discuss enhancements of the design using (a) midcourse modification of sample size and early stopping for futility, (b) group sequential non-inferiority testing, (c) adaptive randomization that allows elimination and addition of treatments, and (d) personalized treatment choice based on baseline patient characteristics.

Key words and phrases: Anticoagulation, Atrial fibrillation, Comparative effectiveness research, Point of care clinical trials, Stroke prevention

1 Introduction

Atrial fibrillation and atrial flutter (AF, collectively) are abnormal heart rhythms of the upper chamber of the heart. In the United States, AF is the second most common cardiovascular diagnosis after hypertension [1], affecting 2.3 million Americans. The prevalence

of AF is 6% in Medicare-eligible Americans over age 65 [1], 10% over age 80 [2], and 5.0-8.3% in veterans [3]. Due to rising rates of obesity and hypertension in young adults, the age-adjusted incidence of AF is increasing [4] and is therefore expected to affect more veterans at younger ages.

AF increases the risk of stroke. The loss of coordinated electromechanical atrial activity predisposes to impaired atrial emptying, stasis of blood, and a prothrombotic state [5]. These factors cause blood clots to form in the heart and embolize systemically to cause stroke or other organ failure [4, 6], which is the major cause of AF-related morbidity and mortality. Among AF patients, the annual incidence of stroke is 4.5%. The estimated annual direct and indirect cost of stroke in the US is $57.9 billion. Although there are numerous treatment options for stroke reduction available to AF patients both inside and outside the Veterans Health Administration (VHA), there is wide variation in the use of any treatment in at-risk patients [7]. Among veterans receiving stroke prevention therapy, there is also wide variation in treatment selection, in large part due to the uncertainty surrounding the comparative effectiveness and safety of these agents. Until recently, warfarin has remained the mainstay of anticoagulation therapy. Safety and effectiveness of warfarin is directly attributable to quality of anticoagulation, as determined by the time within the therapeutic INR range of 2.0-3.0 (TTR). There is wide site-level variation in TTR across the VA and INR monitoring frequency accounted for most of the variation, with the highest variation in rural regions [8, 9].

In the last four years, three new target-specific oral anticoagulants (TSOACs) have received approval in the United States and in many other countries for patients with AF: dabigatran [10, the RE-LY trial], rivaroxaban [11, the ROCKET-AF trial], and apixaban [12, the ARISTOTLE trial]. These fixed-dose agents, which inhibit thrombin (Factor II) or Factor Xa, are administered once or twice daily and unlike warfarin, do not require routine laboratory testing. Although the primary outcome of time to stroke or systemic embolism was nearly identical in the pivotal trials for these three agents, there were considerable differences in inclusion and exclusion criteria. For example, patients in the ROCKET-AF trial [11] required prior stroke or higher stroke risk (CHADS2 score ≥ 3) for enrollment. All three pivotal trials showed at least noninferiority compared to warfarin for prevention of stroke and systemic embolism. All studies were superior to warfarin for reduction in intracranial hemorrhage and fatal bleeding.

Despite these findings, which are described further in Section 2.3, there exists considerable uncertainty as to optimal first-line strategy in veteran and non-veteran patients, and this has led to reluctance and restraint in national VA policy to direct anticoagulation. For example, when dabigatran was first introduced on VA National Formulary, its use was restricted to patients who failed warfarin therapy, despite the results of the RE-LY trial [10]. There is also considerable uncertainty from a cost standpoint. Drug cost escalation remains a major issue for VA Pharmacy Benefits and Management, and their strategies are generally guided by curbing prescription costs when possible. On the other hand, pill cost reduction may not be net cost saving if the effectiveness of the lower-priced therapy

is worse. We have shown that when compared to warfarin, dabigatran is cost-effective in non-valvular AF [13]; the results are consistent with other cost-effectiveness analysis for US and non-US health care systems [14].

There are increasing concerns about the effectiveness and safety of warfarin and the new oral anticoagulants outside of the trials, particularly in vulnerable patients served by the VHA [15]. It has recently been demonstrated that, compared to white and other races, black patients with AF and VA-managed warfarin have lower TTR and poorer warfarin adherence, even though they tend to live closer to VA anticoagulation clinics (M Turakhia, American Heart Association Scientific Sessions, 2013; manuscript in review). There is also high variation in new oral anticoagulant adherence in the VA system. For example, the Palo Alto VA data indicate that among patients initiated on dabigatran, adherence is improved when initiation is managed for the first three months by the specialized anticoagulation clinic (mean medication possession ratio of 0.93 vs. 0.88 among ACC managed and non-ACC managed patients). Although this ratio may seem small, this would translate to an increase in the hazard of stroke in the VA of 6.5% [16]. The VA does nationally mandate a minimum of three months of anticoagulation clinic management for dabigatran initiation. Even with such measures in place, there is still very wide variation in TSOAC adherence across the VA. In summary, the comparative effectiveness and safety of these three TSOACs are critically important to optimizing patient care in VA. This type of comparative effectiveness trial is unlikely to be performed by any pharmaceutical company.

In Section 2, we describe the design of a pragmatic point-of-care clinical trial of the three TSOACs which leverages the existing AF and anticoagulation care infrastructure in the VA. Section 3 gives further discussion and describes some modifications, which are currently under consideration, of the design to make it more efficient.

2 A POC comparative effectiveness trial for stroke prevention in AF patients

In this section, we describe the design of a comparative effectiveness trial for stroke prevention in AF patients, currently being planned at the VA. One important design consideration is to mimic clinical practice as closely as possible in order to inform evidence-based treatment choices at the VA. Since the required sample size to detect a small but clinically meaningful difference in the primary outcome is very large, as shown in Section 2.4, feasibility and cost are two other important considerations in the trial design. The care structure provided by the VHA can meet these design requirements because of: (1) a very large number of patients with AF who meet eligibility; (2) high willingness of eligible patients to participate; (3) established unified, national anticoagulation clinic care structure with experience with TSOAC management; (4) availability of all three treatments on national formulary (with no premium in copayments for veteran patients for any of these drugs compared to warfarin or other drugs); (5) experience from the VA Cooperative Studies

Program (CSP) in designing and conducting anticoagulation multi-site randomized clinical trials.

This large-scale pragmatic clinical trial is being planned under the framework of point-of-care (POC) clinical trials [17], also referred to as clinically-integrated randomized trials [18], with the following main features: (1) Subjects are identified and randomized at health care encounter, therefore the study can include a broad range of patients and providers and reflect the range and distribution of patients seen in clinical practice. (2) Once the patients are enrolled and randomized in the study, there is minimal interference with the usual care, in the sense that patients will continue to be treated and followed by their providers and there is no or little research-only data collection. (3) The baseline and outcome data are captured through electronic medical records (EMR). This eliminates the need to support designated study personnel at each participating site to collect study data, which is a major cost for traditional randomized clinical trials. The VHA has experience from the VA Point-of-Care Program in planning and implementing pragmatic POC clinical trials.

2.1 Study objectives

The primary objective of this trial is to determine the comparative effectiveness of dabigatran, rivaroxaban and apixaban in veterans with VA-managed anticoagulation. The primary effectiveness outcome is the composite of stroke and systemic embolism, intracranial hemorrhage and death. We hypothesize that rivaroxaban is more effective than dabigatran and apixaban in reducing the combined risk of stroke and systemic embolism, intracranial hemorrhage, and death. Secondary effectiveness outcomes include ischemic stroke or systemic embolism, death, and net clinical benefit, which is a weighted composite outcome incorporating stroke or systemic embolism, intracranial hemorrhage, major hemorrhage, and death [19, 20].

The secondary objectives are to determine the comparative safety and cost effectiveness of the three anticoagulants. The primary safety outcome is fatal bleeding. We hypothesize that dabigatran is at increased risk of fatal bleeding compared to rivaroxaban and to apixaban. Secondary safety outcomes include intracranial hemorrhage, major non-fatal hemorrhage, fatal hemorrhage, and minor bleeding. Cost effectiveness evaluates the economic cost and budget impact in the VHA against the comparative effectiveness of the three TSOACs.

2.2 Study population

The study population will consist of VA patients age 18 years or older with atrial fibrillation or atrial flutter, with an indication for anticoagulation in the opinion of the patients' provider(s), and with greater than minimal risk of stroke. Patients who are currently or previously treated with warfarin (warfarin-experienced) will be allowed to enter into study, as well as patients who have not previously been treated with warfarin (warfarin-naive).

Table 1: Design and patient characteristics of the pivotal trials of three anticoagulants.

	RE-LY (Dabigatran)	ROCKET-AF (Rivaroxaban)	ARISTOTLE (Apixaban)
Sample size	18,113 (3 arms)	14,264	18,201
Trial design	open label	double blind, double dummy	double blind, double dummy
Mean age (yrs)	71.5	73	70
Male ratio	63.6%	60.1%	65.3%
% Previous CVA	20%	54.7%	18.9%
CHADS score	2.1	3.5	2.1
%0–1	32%	0%	34%
%≥ 3	33%	87%	30%

2.3 Study treatments and previous clinical findings

Study participants will be randomized in equal allocation to receiving one of three TSOACs, using a permuted block randomization scheme that is stratified by site. Dosing of the TSOACs is based on FDA-approved dosing in the United States, which was tested in the three pivotal trials of each of these drugs compared to warfarin: (a) dabigatran 150 mg orally, twice daily, (b) rivaroxaban 20 mg orally, once daily (or 15 mg orally, once daily, in patients with creatine clearance 15-50 ml/min), (c) apixaban 5 mg orally, twice daily (or 2.5 mg orally, twice daily, in patients with at least two of the following characteristics: age 80 years, body weight 60 kg, or serum creatinine 1.5 mg/dL). Duration of treatment is long-term in accordance with routine practice. The design and patient characteristics of the three pivotal trials (see the third paragraph of Section 1) are summarized in Table 1, in which CVA refers to cerebrovascular accident and CHADS refers to Congestive heart failure, Hypertension, Age, Diabetes, and prior Stroke or transient ischemic attack. The three arms in the RE-LY are dabigatran 110mg and 150mg, and warfarin. The clinical findings are summarized in Table 2, in which HR refers to hazard ratio of the anticoagulant to warfarin for the composite event of stroke or systemic embolism or death, and ICH refers to intracerebral hemorrhage.

2.4 Sample size

The primary comparisons are the three pairwise comparisons among the three TSOACs. The primary endpoint is the composite of stroke and systemic embolism, intracranial hemorrhage, and all-cause mortality. Based on results from the pivotal studies and the VA's EMR, we anticipate the two-year primary event rate is 15% for dabigatran and apixaban and 13% for rivaroxaban. The corresponding number of patients needed to treat per year

Table 2: Clinical outcomes of the pivotal trials of three anticoagulants.

	RE-LY (Dabigatran, 150mg)	ROCKET-AF (Rivaroxaban)	ARISTOTLE (Apixaban)
Event rate %	1.11	1.70	1.27
(vs warfarin)	(1.69)	(2.20)	1.60
HR	0.66 ($p < .001$)	0.79 ($p < .001$)	0.79 ($p < .001$)
ICH %	0.30	0.49	0.24
(vs warfarin)	(0.34)	(0.74)	(0.47)
Major bleeding %	3.37	3.60	2.13
(vs warfarin)	(3.57)	(3.45)	(3.09)
All-cause mortality	3.64	4.52	3.52
%/year (vs warfarin)	(4.13)	(4.91)	(3.94)
Conclusion, compared to warfarin	Superior efficacy. Similar bleeding. Less ICH.	Non-inferior in efficacy & safety.	Superior efficacy. Less bleeding & ICH. Lower mortality.

to prevent 1 primary event is about 100, which is lower than the typical cutoff of 100-150, and therefore this 2% difference is considered clinically meaningful.

To have 80% power to detect a 2% difference in the 2-year event rate (15% vs 13%, hazard ratio 0.86) between two treatment groups at a significance level 5%/3 (due to Bonferroni correction for 3 tests), 1763 events are needed. Assuming 3 years of recruitment and additional 1.5 years of follow-up after the last randomized participant (total study duration 4.5 years, median follow-up 3 years), the number of patients needed is 4381 per group. Rounding this number up to 4400, the study plans to enroll 4400×3=13,200 patients over 3 years. Table 3 gives the required number of participants per group for a range of power and study duration values.

Table 3: Sample size needed to have 80% to 90% power to detect the difference of 15% vs. 13% in the two-year primary event rate.

Power	# Events	Recruitment (yr) + additional follow up (yr)					
		2+2	2.5+1.5	2.5+2	3+1	**3+1.5**	3+2
80%	1763	4360	4726	4068	5172	**4381**	3818
85%	1982	4900	5311	4572	5812	4923	4291
90%	2275	5624	6096	5247	6671	5651	4925

2.5 Recruitment and data collection methods

All patients with atrial fibrillation or atrial flutter with an indication for anticoagulation are potential study participants. Leveraging the VA's anticoagulation clinic care structure, potential participants will be identified through electronic medical records (EMR) either at the entry of new patients or at outpatient visits of patients already on oral anticoagulation. Patients fulfilling all eligibility criteria and having no exclusion criterion will be invited to participate in the study, via a centralized recruitment and consent process.

Data will be accessed at predetermined intervals from administrative data sources such as VA's EMR and chart information, using algorithms for screening and identifying primary and secondary clinical outcomes from VA's EMR and other VA and non-VA databases. Discharge summaries and other pertinent medical records will also be obtained from both VA and non-VA facilities to determine the primary cause of all hospitalizations. For patients with Medicare benefits, linked Medicare data will also be obtained from the VA Information Resource Center (VIReC).

3 Discussion

3.1 Midcourse modification of sample size and early stopping for futility

The sample size calculations in Section 2.4 are based on certain assumptions on the event rates and effect sizes. These assumptions may be found to differ substantially from the observed data during the course of the trial. A related example is the TASTE (Thrombus Aspiration during ST-segment Elevation) myocardial infarction trial [21]. On the basis of mortality data from patients with ST-segment elevation myocardial infarction who underwent percutaneous coronary intervention (PCI) in Sweden between 2008 and 2009, the trial design assumed that the one-month mortality with PCI would be 6.3%. Under this assumption it was calculated that 456 events would have to occur for the trial to have 80% power to detect a hazard ratio of at least 1.3 with PCI as compared with PCI plus thrombus aspiration, leading to a sample size of 5000. When enrollment approached 5000 patients, the 30-day mortality (estimated without knowledge of treatment assignments) was observed to be substantially lower (2.9%) in the study cohort. This led the steering committee of the trial to amend the protocol by increasing the sample size to 7138 patients and to adopt a group sequential design for which interim analysis was conducted by a data monitoring committee.

This experience of TASTE suggests that although the sample size calculation in Table 3 is reasonable at the design stage before data are actually collected, the protocol should allow the possibility of mid-course modification of the sample size. In view of its large sample size, the trial would represent a large waste of resources and effort if it falls short of achieving statistical significance because the sample size calculation is based on assumptions that differ substantially from reality, which one can learn during the course of the trial by using

a group sequential or adaptive design; see Chapters 4, 6, and 8 of [22]. Clearly there are limits on the maximum sample size because of time and resource constraints. An advantage of a group sequential or adaptive design is that it allows early stopping for futility if an interim analysis shows little chance of a significant result when the trial is continued to its maximum sample size.

3.2 Testing non-inferiority in a group sequential design

Traditional designs of comparative effectiveness trials emphasize superiority, and the design in Section 2 follows this tradition. However, even for the case of three treatments as in that trial, such superiority testing often requires large sample sizes because of Bonferroni-type adjustments for multiple comparisons and the typically small to moderate effect sizes between treatments, as illustrated by Table 3. In addition, the actual effect sizes are unknown at the design stage and some of pairwise differences may be close to 0. Such information becomes increasingly available during the course of the trial and can be used by a group sequential design to make the trial more efficient. Instead of testing for superiority to identify the best treatment, Shih and Lavori [23] propose to select a treatment that is within a prescribed distance (called the "non-inferiority margin") from the unknown best treatment. Moreover, instead of the traditional design that analyzes the data at the end of the trial with a fixed sample size, they advocate a group sequential design that carries out interim analyses based on the data collected up to the time of the interim analysis. In particular, at each interim analysis, the design allows early stopping for "futile" treatments in comparison with the currently observed best treatment; generalized likelihood ratio statistics (or partial likelihood ratio statistics for censored time-to-event data) are used for significance testing. The study is stopped if all treatments other than the currently leading one are eliminated. Otherwise continue the study and randomize the next group of patients to the surviving treatments. At the last interim analysis, if there remain more than two treatments, advance the best two treatments to the final analysis at which they are compared by using the generalized likelihood ratio statistic, selecting the leader if the statistic exceeds some critical value, otherwise selecting both treatments. The critical values for significance testing at the interim and final analyses are chosen such that the overall type I error is maintained at a prescribed level. Details are given by Shih and Lavori [23, pp. 1780–1781, 1787–1788]; see also Bartroff et al. [22, Sects. 4.2, 4.3, and 6.5].

3.3 Using baseline patient characteristics to personalize treatment choice

A possible reason for the relatively small treatment differences among the anticoagulants is the heterogeneity of the patient population. These differences may be magnified in certain patient classes. If that is indeed the case, treatment comparison should be carried out within each patient class, resulting in a personalized treatment choice depending on the patient's baseline characteristics that define the classes. Lai et al. [24] have recently intro-

duced a novel group sequential design to develop and test biomarker-guided personalized therapies involving approved cancer treatments. The design fulfills multiple objectives, which include (a) treating accrued patients in the trial with the best (yet unknown) available treatment, (b) developing a treatment strategy for future patients, and (c) demonstrating that the strategy developed indeed has a better treatment effect than that of the standard of care, or that of any of the approved treatments. In a group sequential trial, sequential decisions are made only at times of interim analysis. Equal randomization is applied to the treatments up to the first interim analysis. Then an adaptive randomization scheme is used, assigning the highest randomization probability to the leading treatment in each biomarker class. In addition, generalized likelihood ratio statistics are used for early elimination of significantly inferior treatments from a biomarker class, with the elimination threshold so chosen that there is a guaranteed probability of $1-\alpha$ that the best treatment for each biomarker class is not eliminated, where α corresponds to the type I error. To accomplish this, Lai et al. [24] use subset selection ideas from the selection and ranking literature, in which selecting a subset of treatments, with a guaranteed probability of at least $1-\alpha$ that it contains the best treatment, is an alternative to the indifference zone (or non-inferiority margin) approach, which guarantees the probability of correctly selecting the best treatment to be at least $1-\alpha$ when the largest mean effect differs from the second largest by at least δ, as used by Shih and Lavori [23].

For the stroke prevention trial in AF patients, in place of the biomarkers for cancer patients considered in [24], there are two important patient characteristics that may affect treatment choice. The first is presence or absence of renal dysfunction, as some of the TSOACs are eliminated mainly through the renal system and may therefore accumulate more and be less safe in patients with chronic kidney disease. The second is risk of stroke: low to moderate risk ($1 \leq$ CHADS2 score ≤ 2) versus moderate to high risk (CHADS2 score ≥ 3). In this connection, it should be noted that the pivotal trial of rivaroxaban only enrolled patients with moderate to high risk but the anticoagulant's labeling does not have this restriction.

3.4 Adaptive randomization with arm elimination and addition

The adaptive randomization scheme in [24], which allows arm elimination that is tantamount to assigning zero randomization probability to an eliminated arm, can also be modified to allow inclusion of a new arm in the trial after its initiation. In the context of the comparative effectiveness trial for approved TSOACs, there is a fourth TSOAC, edoxaban, for which a pivotal trial has recently been completed [25, the ENGAGE AF-TIMI 48 trial] but which has not yet received FDA approval. At the time of some interim analysis of the trial, edoxaban may have gained approval and the VHA may want to consider it alongside the other three TSOACs for optimizing patient care in the VA. The modified adaptive randomization scheme assigns to this additional treatment the same randomization probability as that of the leading treatment until the next interim analysis, when this

treatment has been applied to a reasonable number of patients to assess its efficacy in comparison with other treatments.

Acknowledgments

This research was supported in part by the Clinical and Translational Science Award 1UL1 RR025744 for the Stanford Center for Clinical and Translational Education and Research (Spectrum) from the National Center for Research Resources, National Institutes of Health, and by the National Science Foundation (T.L. Lai) and the US Department of Veterans Affairs Cooperative Studies Program (M.-C. Shih).

References

[1] WM Feinberg, JL Blackshear, A Laupacis, R Kronmal, and RG Hart. Prevalence, age distribution, and gender of patients with atrial fibrillation: Analysis and implications. *Archives of Internal Medicine*, 155(5):469–473, MAR 13 1995. ISSN 0003-9926. doi: 10.1001/archinte.155.5.469.

[2] AS Go, EM Hylek, KA Phillips, YC Chang, LE Henault, JV Selby, and DE Singer. Prevalence of diagnosed atrial fibrillation in adults - National implications for rhythm management and stroke prevention: The AnTicoagulation and Risk Factors in Atrial Fibrillation (ATRIA) study. *Journal of the American Medical Association*, 285(18): 2370–2375, MAY 9 2001. ISSN 0098-7484. doi: 10.1001/jama.285.18.2370.

[3] Swee Y. Tan, Jonathan Myers, and Victor F. Froelicher. Relationship between stature and the prevalence of atrial fibrillation and cardiovascular mortality. *Circulation*, 114(18):433, OCT 31 2006. ISSN 0009-7322. 79th Annual Scientific Session of the American Heart Association, Chicago, IL, November 12–15, 2006.

[4] PA Wolf, EJ Benjamin, AJ Belanger, WB Kannel, D Levy, and RB D'Agostino. Secular trends in the prevalence of atrial fibrillation: The Framingham study. *American Heart Journal*, 131(4):790–795, APR 1996. ISSN 0002-8703. doi: 10.1016/S0002-8703(96)90288-4.

[5] SG Rockson and GW Albers. Comparing the guidelines: Anticoagulation therapy to optimize stroke prevention in patients with atrial fibrillation. *Journal of the American College of Cardiology*, 43(6):929–935, MAR 17 2004. ISSN 0735-1097. doi: 10.1016/j.jacc.2003.11.028.

[6] RG Hart, DG Sherman, JD Easton, and JA Cairns. Prevention of stroke in patients with nonvalvular atrial fibrillation. *Neurology*, 51(3):674–681, SEP 1998. ISSN 0028-3878.

[7] Mintu P. Turakhia, Donald D. Hoang, Xiangyan Xu, Susan Frayne, Susan Schmitt, Felix Yang, Ciaran S. Phibbs, Claire T. Than, Paul J. Wang, and Paul A. Heidenreich. Differences and trends in stroke prevention anticoagulation in primary care vs cardiology specialty management of new atrial fibrillation: The Retrospective Evaluation and Assessment of Therapies in AF (TREAT-AF) study. *American Heart Journal*, 165(1):93–101, JAN 2013. ISSN 0002-8703. doi: 10.1016/j.ahj.2012.10.010.

[8] Mintu P. Turakhia. Quality of stroke prevention care in atrial fibrillation: Many moving targets. *Circulation-Cardiovascular Quality and Outcomes*, 4(1):5–8, JAN 2011. ISSN 1941-7713. doi: 10.1161/CIRCOUTCOMES.110.959890.

[9] Adam J. Rose, Dan R. Berlowitz, Arlene S. Ash, Al Ozonoff, Elaine M. Hylek, and Jeremy D. Goldhaber-Fiebert. The business case for quality improvement oral anticoagulation for atrial fibrillation. *Circulation-Cardiovascular Quality and Outcomes*, 4 (4):416–424, JUL 2011. ISSN 1941-7713. doi: 10.1161/CIRCOUTCOMES.111.960591.

[10] Stuart J. Connolly, Michael D. Ezekowitz, Salim Yusuf, John Eikelboom, Jonas Oldgren, Amit Parekh, Janice Pogue, Paul A. Reilly, Ellison Themeles, Jeanne Varrone, Susan Wang, Marco Alings, Denis Xavier, Jun Zhu, Rafael Diaz, Basil S. Lewis, Harald Darius, Hans-Christoph Diener, Campbell D. Joyner, Lars Wallentin, and RE-LY Steering Comm & Investigator. Dabigatran versus warfarin in patients with atrial fibrillation. *New England Journal of Medicine*, 361(12):1139–1151, SEP 17 2009. ISSN 0028-4793. doi: 10.1056/NEJMoa0905561.

[11] Manesh R. Patel, Kenneth W. Mahaffey, Jyotsna Garg, Guohua Pan, Daniel E. Singer, Werner Hacke, Guenter Breithardt, Jonathan L. Halperin, Graeme J. Hankey, Jonathan P. Piccini, Richard C. Becker, Christopher C. Nessel, John F. Paolini, Scott D. Berkowitz, Keith A. A. Fox, Robert M. Califf, and ROCKET AF Investigators. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. *New England Journal of Medicine*, 365(10):883–891, SEP 8 2011. ISSN 0028-4793.

[12] Christopher B. Granger, John H. Alexander, John J. V. McMurray, Renato D. Lopes, Elaine M. Hylek, Michael Hanna, Hussein R. Al-Khalidi, Jack Ansell, Dan Atar, Alvaro Avezum, M. Cecilia Bahit, Rafael Diaz, J. Donald Easton, Justin A. Ezekowitz, Greg Flaker, David Garcia, Margarida Geraldes, Bernard J. Gersh, Sergey Golitsyn, Shinya Goto, Antonio G. Hermosillo, Stefan H. Hohnloser, John Horowitz, Puneet Mohan, Petr Jansky, Basil S. Lewis, Jose Luis Lopez-Sendon, Prem Pais, Alexander Parkhomenko, Freek W. A. Verheugt, Jun Zhu, Lars Wallentin, and ARISTOTLE Comm Investigators. Apixaban versus warfarin in patients with atrial fibrillation. *New England Journal of Medicine*, 365(11):981–992, SEP 15 2011. ISSN 0028-4793. doi: 10.1056/NEJMoa1107039.

[13] James V. Freeman, Ruo P. Zhu, Douglas K. Owens, Alan M. Garber, David W. Hutton, Alan S. Go, Paul J. Wang, and Mintu P. Turakhia. Cost-effectiveness of

dabigatran compared with warfarin for stroke prevention in atrial fibrillation. *Annals of Internal Medicine*, 154(1):1–11, JAN 4 2011. ISSN 0003-4819.

[14] Matthew D. Solomon, Aditya J. Ullal, Donald D. Hoang, James V. Freeman, Paul Heidenreich, and Mintu P. Turakhia. Cost-effectiveness of pharmacologic and invasive therapies for stroke prophylaxis in atrial fibrillation. *Journal of Cardiovascular Medicine*, 13(2):86–96, FEB 2012. ISSN 1558-2027. doi: 10.2459/JCM.0b013e32834f23cf.

[15] Supriya Shore, Evan P. Carey, Mintu P. Turakhia, Cynthia A. Jackevicius, Fran Cunningham, Louise Pilote, Steven M. Bradley, Thomas M. Maddox, Gary K. Grunwald, Anna E. Barn, John S. Rumsfeld, Paul D. Varosy, Preston M. Schneider, Lucas N. Marzec, and P. Michael Ho. Adherence to dabigatran therapy and longitudinal patient outcomes: Insights from the Veterans Health Administration. *American Heart Journal*, 167(6):810–817, 2014. ISSN 0002-8703. doi: 10.1016/j.ahj.2014.03.023.

[16] Supriya Shore, Evan P. Carey, Mintu P. Turakhia, Cynthia A. Jackevicius, Fran Cunningham, Louise Pilote, Steven M. Bradley, Thomas M. Maddox, Gary K. Grunwald, Anna E. Baron, John S. Rumsfeld, Paul D. Varosy, Preston M. Schneider, Lucas N. Marzec, and P. Michael Ho. Adherence to dabigatran therapy and longitudinal patient outcomes: Insights from the Veterans Health Administration. *American Heart Journal*, 167(6):810–817, JUN 2014. ISSN 0002-8703. doi: 10.1016/j.ahj.2014.03.023.

[17] Louis D. Fiore, Mary Brophy, Ryan E. Ferguson, Leonard D'Avolio, John A. Hermos, Robert A. Lew, Gheorghe Doros, Chester H. Conrad, Joseph A. (Gus) O'Neil, Jr., Thomas P. Sabin, James Kaufman, Stephen L. Swartz, Elizabeth Lawler, Matthew H. Liang, J. Michael Gaziano, and Philip W. Lavori. A point-of-care clinical trial comparing insulin administered using a sliding scale versus a weight-based regimen. *Clinical Trials*, 8(2):183–195, APR 2011. ISSN 1740-7745. doi: 10.1177/1740774511398368.

[18] Andrew J. Vickers and Peter T. Scardino. The clinically-integrated randomized trial: Proposed novel method for conducting large trials at low cost. *Trials*, 10, MAR 5 2009. ISSN 1745-6215. doi: 10.1186/1745-6215-10-14.

[19] Daniel E. Singer, Yuchiao Chang, Margaret C. Fang, Leila H. Borowsky, Niela K. Pomernacki, Natalla Udaltsova, and Alan S. Go. The net clinical benefit of warfarin anticoagulation in atrial fibrillation. *Annals of Internal Medicine*, 151(5):297–305, SEP 1 2009. ISSN 0003-4819.

[20] John W. Eikelboom, Stuart J. Connolly, Robert G. Hart, Lars Wallentin, Paul Reilly, Jonas Oldgren, Sean Yang, and Salim Yusuf. Balancing the benefits and risks of 2 doses of dabigatran compared with warfarin in atrial fibrillation. *Journal of the American College of Cardiology*, 62(10):900–908, SEP 3 2013. ISSN 0735-1097. doi: 10.1016/j.jacc.2013.05.042.

[21] Ole Frobert, Bo Lagerqvist, Goran K. Olivecrona, Elmir Omerovic, Thorarinn Gudnason, Michael Maeng, Mikael Aasa, Oskar Angeras, Fredrik Calais, Mikael Danielewicz, David Erlinge, Lars Hellsten, Ulf Jensen, Agneta C. Johansson, Amra Karegren, Johan Nilsson, Lotta Robertson, Lennart Sandhall, Iwar Sjogren, Ollie Ostlund, Jan Harnek, and Stefan K. James. Thrombus aspiration during ST-segment elevation myocardial infarction. *New England Journal of Medicine*, 369(17):1587–1597, OCT 24 2013. ISSN 0028-4793. doi: 10.1056/NEJMoa1308789.

[22] Jay Bartroff, Tze Leung Lai, and Mei-Chiung Shih. *Sequential Experimentation in Clinical Trials: Design and Analysis*, volume 298 of *Springer Series in Statistics*. Springer, New York, 2012. ISBN 10: 1461461138; 13: 978-1461461135.

[23] Mei-Chiung Shih and Philip W. Lavori. Sequential methods for comparative effectiveness experiments: Point of care clinical trials. *Statistica Sinica*, 23:1775–1791, OCT 2013. ISSN 1017-0405. doi: 10.5705/ss.2012.231s. Statistical genetics, change-point problems, random fields and sequential analysis; a special issue in honor of Professor David Siegmund's 70th Birthday.

[24] Tze Leung Lai, Olivia Yueh-Wen Liao, and Dong Woo Kim. Group sequential designs for developing and testing biomarker-guided personalized therapies in comparative effectiveness research. *Contemporary Clinical Trials*, 36(2):651–663, NOV 2013. ISSN 1551-7144. doi: 10.1016/j.cct.2013.08.007.

[25] Robert P. Giugliano, Christian T. Ruff, Eugene Braunwald, Sabina A. Murphy, Stephen D. Wiviott, Jonathan L. Halperin, Albert L. Waldo, Michael D. Ezekowitz, Jeffrey I. Weitz, Jindrich Spinar, Witold Ruzyllo, Mikhail Ruda, Yukihiro Koretsune, Joshua Betcher, Minggao Shi, Laura T. Grip, Shirali P. Patel, Indravadan Patel, James J. Hanyok, Michele Mercuri, Elliott M. Antman, and ENGAGE AF-TIMI 48 Investigators. Edoxaban versus warfarin in patients with atrial fibrillation. *New England Journal of Medicine*, 369(22):2093–2104, NOV 28 2013. ISSN 0028-4793. doi: 10.1056/NEJMoa1310907.

Department of Statistics, Stanford University
Email: lait@stanford.edu

VA Palo Alto Cooperative Studies Program Coordinating Center and
Department of Health Research and Policy, Stanford University
Email: meichiun@stanford.edu

Department of Medicine, Stanford University
Email: mintu@stanford.edu