Dabigatran Versus Warfarin for Atrial Fibrillation in Real-World Clinical Practice
A Systematic Review and Meta-Analysis

Robert J. Romanelli, PhD, MPH; Laura Nolting, BS; Marina Dolginsky, BS; Eunice Kym, PharmD; Kathleen B. Orrico, PharmD

Background—Trial data for the benefits and risks of dabigatran versus warfarin in the treatment of nonvalvular atrial fibrillation are lacking. We sought to review real-world observational evidence for the comparative effectiveness and safety of these agents.

Methods and Results—A systematic search of multiple databases was conducted from first available date to March 10, 2015 for longitudinal, observational studies comparing dabigatran with warfarin. Two reviewers evaluated studies for eligibility and extracted hazard ratios for ischemic stroke and gastrointestinal and intracranial bleeding. Hazard ratios were pooled using random-effects meta-analysis. Metaregression was performed to assess treatment-effect heterogeneity. We identified 232 unique citations. Seven retrospective cohort studies met study eligibility criteria, with 348,750 patients and a mean follow-up of 2.2 years. In pooled analyses, dabigatran-150 mg was not superior to warfarin in preventing stroke (hazard ratio, 0.92; 95% confidence interval, 0.84–1.01; P=0.066), but had a significantly lower hazard of intracranial bleeding (0.44; 0.34–0.59; P<0.001). Dabigatran-150 mg had a significantly greater hazard of gastrointestinal bleeding than warfarin (1.23; 1.01–1.50; P=0.041), which was potentiated in studies of older (elderly) versus younger populations (median/mean age, >75 versus <75 years; β=1.53; 95% confidence interval, 1.10–2.14; P=0.020).

Conclusions—In real-world clinical practice, dabigatran is comparable with warfarin in preventing ischemic stroke among patients with nonvalvular atrial fibrillation. However, dabigatran is associated with a lower risk for intracranial bleeding relative to warfarin, but—particularly among the elderly—a greater risk for gastrointestinal bleeding. Bleeding outcomes from observational studies are consistent with those from the pivotal Randomized Evaluation of Long-Term Anticoagulation Therapy trial. (Circ Cardiovasc Qual Outcomes. 2016;9:126-134. DOI: 10.1161/CIRCOUTCOMES.115.002369.)

Key Words: atrial fibrillation ▪ dabigatran ▪ evidence-based medicine ▪ meta-analysis ▪ stroke

Atrial fibrillation (AF) is associated with a 5-fold increase in the risk of stroke. For more than a half century, the vitamin K antagonist warfarin has been used for stroke prevention in patients with nonvalvular AF (NVAF). The safe use of warfarin requires frequent blood testing and dose adjustments to maintain therapeutic anticoagulation and prevent bleeding events, as well as dietary and other lifestyle restrictions. In 2010, the first novel oral anticoagulant (NOAC), dabigatran, a direct thrombin inhibitor, became available for the treatment of NVAF. Dabigatran challenged the mainstay of treatment, as use of this agent is a more convenient treatment option, which does not require frequent blood testing or lifestyle restrictions.

The Randomized Evaluation of Long-Term Anticoagulation Therapy (RE-LY) was a large-scale multicenter randomized clinical trial (RCT), evaluating 2 fixed doses of dabigatran (110 mg or 150 mg, twice daily) versus adjusted-dose warfarin for >2 years. The study was conducted in 44 countries and included 18,113 patients with NVAF. The trial showed that dabigatran-150 mg was superior to warfarin in preventing ischemic stroke, and had a lower rate of intracranial bleeding, but a higher rate of gastrointestinal bleeding. Dabigatran-110 mg was noninferior to warfarin in preventing ischemic stroke and had a similar rate of gastrointestinal bleeding, but a lower rate of intracranial bleeding. On the basis of the findings from this trial, dabigatran-110 mg and dabigatran-150 mg both became available for the treatment of NVAF in most countries, except in the United States, where the 150 mg, but not the 110 mg, dose was approved.

RCTs provide the strongest evidence for drug safety and efficacy. The synthesis of data from multiple and similarly designed RCTs, for a given drug and particular condition, is an important tool in comparative effectiveness research. Such techniques add to the extant body of knowledge by overcoming some of the limitations of individual trials, namely by increasing sample size and improving the precision of effect estimates. Moreover, these techniques allow for the evaluation
WHAT IS KNOWN
• Evidence for the safety and efficacy of dabigatran versus warfarin for stroke prevention in patients with nonvalvular atrial fibrillation (NVAF) is derived from a single, large randomized controlled trial.
• Observational studies can complement what is known about the risks and benefits of these therapies in clinical practice.

WHAT THE STUDY ADDS
• There are currently a limited number of well-designed retrospective observational studies that have evaluated the comparative safety and effectiveness of dabigatran and warfarin in stroke prevention for patients with nonvalvular atrial fibrillation.
• Pooled analysis of these studies shows that dabigatran is comparable with warfarin in stroke prevention, but it is associated with a lower risk of intracranial bleeding and an increased risk of gastrointestinal bleeding.
• The association between dabigatran use and gastrointestinal bleeding is more evident in populations of older (mean/median age, at least 75 years) versus younger patients (mean/median age, <75 years).

of heterogeneity of treatment effects across studies and the development of new hypotheses for future studies.
RCTs evaluating dabigatran in stroke prevention for patients with NVAF are lacking—we are aware of only the RE-LY trial. Consequently, we are dependent on evidence from observational studies if we wish to further examine dabigatran beyond what we have learned from RE-LY. Although observational studies do not provide the same robust level of evidence as RCTs because they are more prone to selection bias, well-designed observational studies may provide better generalizability to real-world practice because they often have less restrictive inclusion criteria and treatment occurs in a clinical setting under usual care. As such, high-quality observational studies can importantly contribute to the totality of evidence for the benefits and risks of an intervention, especially when there are gaps in evidence from RCTs.5-7 Under this framework, we sought to review real-world observational evidence for the comparative effectiveness and safety of dabigatran and warfarin among patients with NVAF.

Methods
We conducted this systematic review and structured the report of its findings based on the Meta-analysis of Observational Studies in Epidemiology (MOOSE) guidelines.

Data Sources and Searches
We searched the literature for studies comparing dabigatran with warfarin in real-world clinical practice. Four databases were searched (PubMed, CINHAL Plus, DARE, and Web of Science) from first available date to March 10, 2015, for the following terms: (nonvalvular or atrial fibrillation or AF or NVAF) AND (warfarin or Coumadin) and (novel oral anticoagulant or NOAC or dabigatran or Pradaxa) and (real world or observational or cohort). Manual searches of reference lists and table of contents of relevant journals were also performed.

Study Selection
Two independent reviewers evaluated the titles and abstracts of citations. Among citations meeting initial eligibility criteria, or those with no available abstract, the 2 reviewers examined the full text. We included longitudinal, observational cohort studies that compared dabigatran with warfarin for stroke prevention in NVAF. We excluded studies that reported only crude hazard estimates (ie, no statistical adjustment for between-group differences). No language restrictions were made.

Data Extraction and Quality Assessment
For each citation, 1 reviewer extracted information into tables on study, intervention, and patient characteristics, and outcome measures. A second reviewer assessed the accuracy of the extraction. We contacted authors, when necessary, for information that was not available in citations or if clarification was needed. We successfully contacted the authors of 2 manuscripts to obtain information on follow-up and outcome measures.

Observational studies by nature have a higher risk of bias relative to RCTs. Thus, we did not evaluate study quality using tools designed for RCTs. Instead, we considered factors important in observational study design and methods used to mitigate bias when comparing outcomes in this setting. Low, moderate, or high risk of bias was assigned to each citation within the following domains: (1) use of matching or statistical adjustment to handle selection bias, (2) potential for residual confounding, (3) use of methods to handle time-varying covariates and differential/informative censoring, and (4) detailed reporting of baseline characteristics and outcome measures (eg, numerators, denominators, and univariate and multivariable statistics).

Data Synthesis and Analysis
We extracted data on patient demographics, characteristics, and comorbidities. The primary effectiveness outcome of these studies was ischemic stroke and the primary safety outcomes were intracranial or gastrointestinal bleeding events. Other outcomes such as myocardial infarction, all-cause mortality, and other bleeding events were inconsistently reported across studies, and they were not included in this review. We summarized adjusted hazard ratios (HRs) and 95% confidence intervals (CIs) for primary outcomes from Cox proportional hazard models comparing dabigatran with warfarin (referred). When 95% CIs excluded the null, we considered 1 treatment superior to the other. When 95% CIs included the null, we considered treatments comparable (or noninferior). When appropriate, we pooled effect sizes (log-transformed HRs) from similar studies using random-effects meta-analysis with DerSimonian–Laird weighting.9 Pooled HRs (pHRs) and 95% CIs were generated. Heterogeneity across studies was assessed with the I2 statistic, and it was considered minimal, moderate, important, and substantial with an I2 of <25%, 25% to 50%, 50% to 75%, and >75%, respectively. We performed sensitivity analyses to assess the robustness of outcomes to a priori assumptions and to identify the source of between-study heterogeneity. We used meta-regressions to further explore heterogeneity of treatment effects by patient characteristics and study quality. For all test statistics, a P value <0.05 was considered statistically significant. Analyses were performed in Stata MP 13 (StataCorp; College Station, TX).

Results
Search Results and Study Characteristics
We identified 232 unique citations, 7 of which describing retrospective, observational cohort studies met study eligibility criteria (Figure 1). Eligible studies were new-user analyses of dabigatran versus warfarin, conducted on nationally representative insurance claims data or registry data. Studies were published between 2012 and 2015, and all were written in English.
or heart value repair, and required a wash-out period for previous anticoagulation, ranging from 6 to 12 months. In 5 US studies, dabigatran daily dose was assumed to be 150 mg twice daily, which is the recommended, on-label dosage. One US study specified that ≈16% of patients with potential renal failure used the lower 75 mg dose twice daily; 1 is another study, the authors report that 9.6% of patients used this lower dose. For our analyses, we categorized all dabigatran use in US-based studies as 150 mg. Two primary non-US studies reported use of dabigatran-150 mg and dabigatran-110 mg twice daily; 1 study was conducted in Denmark and the other in Canada. The Canadian study excluded 18 patients, of whom 60000 in total, who received dabigatran-75 mg. We pooled estimates for each dabigatran dose versus warfarin across studies when there were at least 3-point estimates for comparison. One study exclusively reported outcomes dichotomized by age (<75/≥75 years of age); for this study we included both point estimates into pooled analyses for each outcome.

All studies used propensity-score techniques to match dabigatran to warfarin patients or used statistical adjustment or propensity-score weighting of baseline covariates in Cox proportional hazard models to obtain adjusted hazard estimates. Patients were followed to event or were censored at the end of continuous enrollment or the administrative period. We note that causal-inference methods (ie, handing of time-varying covariates) were not used in the included analyses. For quality assessment of each article please see Table 2.

Patient Characteristics

The studies included in this review comprised 348750 patients (warfarin, n=197348 [56.6%]; dabigatran-150 mg, n=197348

<table>
<thead>
<tr>
<th>Author</th>
<th>Country (Funding)</th>
<th>Data Source</th>
<th>Major Eligibility Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tsadok et al</td>
<td>Canada (Federal)</td>
<td>Administrative Claims</td>
<td>Patients receiving dabigatran (110 mg or 150 mg) or warfarin discharge from hospitalization with a primary diagnosis of AF.</td>
</tr>
<tr>
<td>Lauffenburger et al</td>
<td>United States (Federal)</td>
<td>National Commercial Claims</td>
<td>Aged ≥18 y, 1 inpatient or 2 outpatient claims for AF and ≥1 Rx for dabigatran or warfarin, continuously enrolled for 12 mo before the first fill date; excluded patients with anticoagulation in the 12 mo before, or with valvular or transient AF.</td>
</tr>
<tr>
<td>Abraham et al</td>
<td>United States (Federal)</td>
<td>National Commercial Claims</td>
<td>Aged ≥18 y, claim for AF within 12 mo before Rx for dabigatran or warfarin; 12 mo continuous enrollment and no anticoagulation in previous 12 mo; excluded patients with mechanical heart failure, mitral stenosis, chronic hemodialysis or peritoneal dialysis, or kidney transplant, and nursing facility patients.</td>
</tr>
<tr>
<td>Graham et al</td>
<td>United States (Federal)</td>
<td>National Medicare Claims</td>
<td>Aged ≥65 y, any inpatient or outpatient claim for AF and Rx for dabigatran or warfarin, ≥6 m continuous enrollment, and no anticoagulation in previous 6 mo; excluded patients in nursing facility or hospice, or with dialysis, kidney transplant, mitral value disease, heart value repair or replacement, DVT, PE, or joint replacement surgery.</td>
</tr>
<tr>
<td>Hernandez et al</td>
<td>United States (Federal)</td>
<td>National Medicare Claims</td>
<td>One inpatient or 2 outpatient claims for AF and Rx for dabigatran or warfarin within 2 mo of first diagnosis; no major exclusions reported.</td>
</tr>
<tr>
<td>Larsen et al</td>
<td>Denmark (none)</td>
<td>National Registry</td>
<td>Treatment with dabigatran (110 mg or 150 mg) or warfarin for AF; excluded patients with PE, DVT, or mitral stenosis.</td>
</tr>
<tr>
<td>Thelus et al</td>
<td>United States (none)</td>
<td>Military Healthcare Claims</td>
<td>Incident users of dabigatran or warfarin for AF; at least 12 mo continuous enrollment and ≥12 mo follow-up.</td>
</tr>
</tbody>
</table>

AF indicates atrial fibrillation; DVT, deep vein thrombosis; PE, pulmonary embolism; and Rx, prescription.

*Conference abstract.
by guest on October 14, 2017 http://circoutcomes.ahajournals.org/ Downloaded from

1.10–1.26; \( \beta \approx 50\% \) increased hazard

Using meta-regression, we found an \( I^2 = 0.514; \) \( P \approx 0.020 \). One study found a reduced risk \( 15 \) and another found a similar risk \( 10 \) of gastrointestinal bleeding for dabigatran-110 mg compared with warfarin in patients <75 years of age; the latter study showed an increased risk of \( \approx 50\% \) for dabigatran-150 mg versus warfarin in patients ≥75 years. \( 10,13,16 \) Dabigatran was potentially driven by age. Stratified by study age distribution, there was no evidence for an increased risk of gastrointestinal bleeding with dabigatran in studies of younger populations, but there was an increased risk of \( \approx 50\% \) for dabigatran-150 mg versus warfarin in studies of older populations (mean/median age, ≥75 years).

Pooled analyses showed no benefit of dabigatran-150 mg (pHR, 0.92; 95% CI, 0.84–1.01; \( P = 0.066; I^2 = 6.5\% \)) or dabigatran-110 mg (pHR, 0.92; 95% CI, 0.72–1.18; \( P = 0.514; I^2 = 60.5\% \)) in ischemic stroke prevention relative to warfarin (Figure 2). However, both dabigatran-150 mg (pHR, 0.44; 95% CI, 0.34–0.59; \( P < 0.001; I^2 = 63.6\% \)) and dabigatran-110 mg (pHR, 0.49; 95% CI, 0.34–0.72; \( P < 0.001; I^2 = 23.5\% \)) had lower hazards of intracranial bleeding compared with warfarin. The hazard of gastrointestinal bleeding was significantly greater for dabigatran-150 mg versus warfarin (pHR, 1.23; 95% CI, 1.01–1.50; \( P = 0.041; I^2 = 89.8\% \)), but not for dabigatran-110 mg (pHR, 0.91; 95% CI, 0.55–1.51; \( P = 0.705; I^2 = 83.4\% \)). We note substantial heterogeneity in pooled estimates for gastrointestinal bleeding.

### Discussion

In this systematic review, we identified 7 well-designed retrospective, observational cohort studies evaluating dabigatran compared with warfarin for stroke prevention in patients with NVAF, which together, included 348,750 patients with an average follow-up of 2.2 years. Pooled analyses showed no benefit of either dabigatran dose over warfarin in preventing ischemic stroke; however, dabigatran-150 mg and dabigatran-110 mg, respectively, were associated with a 56% and 51% lower hazard for intracranial bleeding compared with warfarin. The hazard of gastrointestinal bleeding was 23% greater with dabigatran-150 mg relative to warfarin; however, we found that the excess risk of gastrointestinal bleeding with dabigatran was potentially driven by age. Stratified by study age distribution, there was no evidence for an increased risk of gastrointestinal bleeding with dabigatran in studies of younger populations, but there was an increased risk of \( \approx 50\% \) for dabigatran-150 mg versus warfarin in studies of older populations (mean/median age, ≥75 years).

Pooled safety outcomes from the included observational studies were comparable with those reported in the RE-LY trial. Among 18,113 patients in RE-LY, there was a 60% and 69% lower risk of intracranial bleeding with dabigatran-150 mg and dabigatran-110 mg, respectively, versus warfarin. \( 3 \) Furthermore, patients who received dabigatran-150 mg had a 50% increased risk of gastrointestinal bleeding relative to warfarin, but no increased risk was observed for patients who received dabigatran-110 mg. \( 3 \) In a subanalysis of the RE-LY trial, study authors also found a trend for the effect of age on gastrointestinal bleeding risk. \( 19 \) In RE-LY, risk of gastrointestinal bleeding was greater with dabigatran-150 mg relative to warfarin in patients ≥75 years of age (relative risk, 1.79; 95% CI, 1.35–2.37) but not in patients <75 years of age (relative risk, 1.19; 95% CI, 0.87–1.63), yet the interaction was not statistically significant (\( P = 0.06 \)). \( 19 \) This trend was corroborated in at least 3 of the observational studies, included in our review, which conducted similar stratified analyses.

In a Medicare population, Graham et al \( 13 \) found the expected hazard of gastrointestinal bleeding with dabigatran-150 mg (HR, 2.49; 1.61–3.83) exceeded that with warfarin (HR, 1.62; 95% CI, 1.02–2.58) in patients >75 versus ≤75 years of age. Finally, in a Canadian population Tsadok et al \( 10 \) found that

### Table 2. Assessment of Observational Study Quality*

<table>
<thead>
<tr>
<th>Selection bias</th>
<th>Tsadok et al(^{10} )</th>
<th>Lauffenburger et al(^{11} )</th>
<th>Abraham et al(^{12} )</th>
<th>Graham et al(^{13} )</th>
<th>Hernandez et al(^{14} )</th>
<th>Larsen et al(^{15} )</th>
<th>Thelus et al(^{16} )†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias because of residual confounding</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Mod</td>
</tr>
<tr>
<td>Bias because of time-varying covariates/information censoring</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Bias because of selective reporting of study measures and outcomes</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Mod</td>
</tr>
<tr>
<td>Bias because of selective reporting of study measures and outcomes</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
</tr>
</tbody>
</table>

*Low, moderate (mod), or high risk of bias assigned to each of the 4 domains.
†Conference abstract.

[40.2%]; and dabigatran-110 mg, \( n=11,305 \) [3.2%]) with an overall mean follow-up of 794 days (=2.2 years; Table 3). Cohorts were heterogeneous with regard to the inclusion of females, and a history of stroke, hypertension, or diabetes mellitus. There was limited information available on patient characteristics in the one included conference abstract.

### Heterogeneity of Treatment Effects

We further explored effect modification of treatment outcomes by patient characteristics. The most notable difference between studies was age distribution. Three of the seven studies evaluated younger populations (mean/median age, <75 years) \( 11,12,15 \) and 4 studies evaluated older (elderly) populations (mean/median age, ≥75 years). \( 10,13,14,16 \) Dabigatran-150 mg was associated with an increased hazard for gastrointestinal bleeds relative to warfarin in studies of older populations (pHR, 1.51; 95% CI, 1.20–1.90; \( P < 0.001; I^2 = 85.0\% \)) but not in those of younger populations (pHR, 0.99; 95% CI, 0.83–1.18; \( P = 0.292; I^2 = 53.6\% \); Figure 3). Using meta-regression, we found an \( \approx 50\% \) increased hazard of gastrointestinal bleeding with dabigatran-150 mg in studies of older versus younger populations (\( \beta = 1.53; 95\% CI, 1.10–1.26; P = 0.020 \)). One study found a reduced risk \( 15 \) and another found a similar risk \( 10 \) of gastrointestinal bleeding for dabigatran-110 mg compared with warfarin in patients <75 years of age; the latter study showed an increased risk of bleeding with dabigatran-110 mg versus warfarin in patients ≥75 years of age (Figure 3). We found no other differences in treatment effect by patient characteristics or by study quality (data not shown).
dabigatran-150 mg was associated with an increased risk of gastrointestinal bleeding in patients ≥75 years of age compared with warfarin (HR, 1.35; 95% CI, 1.01–1.82), but not in patients <75 years of age (HR, 0.96; 95% CI, 0.74–1.51).

Observational evidence for the comparative effectiveness of dabigatran versus warfarin was not necessarily consistent with findings from the RE-LY trial. In RE-LY, dabigatran-150 mg was superior to warfarin in stroke prevention, with a 26% decreased risk for ischemic stroke.3 Dabigatran-110 mg, however, was noninferior to warfarin in stroke prevention. Although we found a trend toward a reduced risk of ischemic stroke with dabigatran-150 mg versus warfarin, the effect size was small and did not reach statistical significance (for dabigatran-150 mg: PHR, 0.92; 95% CI, 0.84–1.01; P=0.066). We note that only 1 of the 6 observational studies that evaluated ischemic stroke as an outcome measure reported a benefit for dabigatran-150 mg over warfarin (adjusted HR, 0.80; 95% CI, 0.67–0.96).13

Limitations
The findings of this systematic review should be interpreted in the context of several limitations. Observational studies have a greater risk for bias than RCTs, and confounding because of unmeasured factors cannot be ruled out. Most studies were conducted on administrative claims databases, for which information on over-the-counter medications, such as aspirin, is not available. Aspirin use was unmeasured in most observational studies, and outcomes may be biased if aspirin was used at different rates among patients receiving dabigatran and warfarin. Furthermore, clinical or laboratory information was not available in these studies; thus, time in therapeutic range for warfarin could not be assessed or controlled for. Suboptimal time in therapeutic range in warfarin users may have biased results favoring dabigatran. Although observational studies used robust methods to handle selection bias, none of the included studies used appropriate methods to handle time-varying covariates and informative censoring (ie, marginal-structural models). In the absence of these methods, findings from observational studies do not allow for causal inferences.

In this review, we report outcomes for ischemic stroke and intracranial and gastrointestinal bleeding events because they were the most commonly and consistently reported across observational studies. Other outcomes, such as myocardial infarction, any major/minor bleeding, and all-cause mortality, were reported in only a few studies but were inconsistently measured. Thus, we did not review these outcomes. Gray literature was not searched and publication bias cannot be ruled out. An assessment of publication bias could not be performed because of the inclusion of relatively few studies.

Data presented in this review reflect relative risk, which is not always clinically meaningful. It is important to bear in mind that event rates for the outcome of interest are low under standard treatment. In the RE-LY trial, rates of ischemic stroke, intracranial bleeding, and gastrointestinal bleeding were 1.2%, 0.74%, and 1.02%, respectively, per year in the warfarin-treated group.3 Unadjusted absolute event rates were similarly low among warfarin-treated patients in the observational studies included in this review, but varied across studies (data not show). Given these low event rates, even large reductions in relative risk in dabigatran-treated patients would yield small absolute reductions in risk.

The number needed to treat (NNT), represented as the reciprocal of the absolute difference in risk, is often more useful for clinical interpretation. The concept of NNT was developed for RCTs with binary outcomes measured at fixed time points.20,21 The observational studies included in our review appropriately used survival data, in which exposure time is right censored for

Table 3. Patient Demographics and Clinical Characteristics from Primary Observational Studies

<table>
<thead>
<tr>
<th></th>
<th>Tsadok et al10</th>
<th>Lauffenburger et al1</th>
<th>Abraham et al12</th>
<th>Graham et al13</th>
<th>Hernandez et al14</th>
<th>Larsen et al15</th>
<th>Thelus et al16*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total, n</td>
<td>63,110</td>
<td>64,935</td>
<td>15,498</td>
<td>134,414</td>
<td>9404</td>
<td>13,554</td>
<td>47,845</td>
</tr>
<tr>
<td>Dabigatran-150</td>
<td>6992</td>
<td>21,070</td>
<td>7749</td>
<td>67,207</td>
<td>1302</td>
<td>2239</td>
<td>33,548</td>
</tr>
<tr>
<td>Dabigatran-110</td>
<td>8992</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>2379</td>
<td>...</td>
</tr>
<tr>
<td>Warfarin</td>
<td>47,192</td>
<td>43,865</td>
<td>7749</td>
<td>67,207</td>
<td>8102</td>
<td>89,36</td>
<td>14,297</td>
</tr>
<tr>
<td>Mean follow-up, d</td>
<td>210</td>
<td>358</td>
<td>202</td>
<td>1307</td>
<td>221</td>
<td>424</td>
<td>1078</td>
</tr>
<tr>
<td>Mean age, y</td>
<td>78</td>
<td>70</td>
<td>67</td>
<td>...</td>
<td>75</td>
<td>71</td>
<td>75</td>
</tr>
<tr>
<td>75 y, %</td>
<td>67</td>
<td>41</td>
<td>27</td>
<td>59</td>
<td>...</td>
<td>37</td>
<td>...</td>
</tr>
<tr>
<td>Female, %</td>
<td>50</td>
<td>40</td>
<td>36</td>
<td>52</td>
<td>59</td>
<td>42</td>
<td>...</td>
</tr>
<tr>
<td>Comorbidities, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heart failure</td>
<td>35</td>
<td>25</td>
<td>25</td>
<td>18</td>
<td>51</td>
<td>8</td>
<td>...</td>
</tr>
<tr>
<td>Hypertension</td>
<td>76</td>
<td>72</td>
<td>...</td>
<td>87</td>
<td>88</td>
<td>20</td>
<td>...</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>30</td>
<td>30</td>
<td>29</td>
<td>34</td>
<td>44</td>
<td>12</td>
<td>...</td>
</tr>
<tr>
<td>Previous stroke</td>
<td>11</td>
<td>10</td>
<td>...</td>
<td>4</td>
<td>22</td>
<td>17</td>
<td>...</td>
</tr>
<tr>
<td>Renal disease</td>
<td>28</td>
<td>10</td>
<td>8</td>
<td>18</td>
<td>33</td>
<td>3</td>
<td>...</td>
</tr>
<tr>
<td>Previous bleed</td>
<td>10</td>
<td>12</td>
<td>...</td>
<td>4</td>
<td>11</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

*Conference abstract.
Figure 2. Forest plots comparing dabigatran to warfarin for effectiveness and safety outcomes. Hazard ratios (HR) and 95% confidence intervals (CI) are shown for dabigatran 150 mg or 110 mg relative to warfarin (referent). Pooled estimates (displayed as blue diamonds for dabigatran-150 mg and green diamonds for dabigatran-110 mg) were calculated by random-effects meta-analysis with DerSimonian-Laird weighting. The I² statistic denotes between-study heterogeneity for pooled analyses. Level of significance was P<0.05.

both clinical and administrative reasons. Because the number of patients under observation differs over time, so to would the NNT. Currently, translation of methods to calculate NNT using survival data are not well established and in practice have produced varied and questionable results. With these methods, the assumption of constant hazards would need to be met, which cannot be tested with aggregated data. Issues with survival data aside, deriving NNT from meta-analyses has been cautioned against because of potential differences in follow-up length, clinical settings, and baseline risk across studies. In this regard, both survival data and pooled estimates from meta-analysis are likely not amenable to estimating NNT.

Strengths

This systematic review and meta-analysis has several merits. To our knowledge, it is the first to evaluate real-world evidence
for the effectiveness and safety of dabigatran versus warfarin for stroke prevention in NVAF. Meta-analyses of RCTs have been previously performed by evaluating all NOACs, combined, relative to warfarin.\textsuperscript{27-30} Although these meta-analyses included well-designed trials, there is the potential for heterogeneous effects across NOACs and, thus, generalizability of outcomes to the use of each agent is questionable. Indeed, an adjusted indirect comparison meta-analysis of RCTs in NVAF comparing dabigatran, rivaroxaban, or apixaban each with warfarin showed that there are differences in efficacy and safety outcomes between these agents.\textsuperscript{31} The RE-LY trial was the only included RCT of dabigatran versus warfarin for stroke prevention in the aforementioned meta-analyses. In fact, we are unaware of other RCTs in NVAF of dabigatran versus warfarin.

We evaluated large, nationally representative cohort studies that used robust methods to handle selection bias. Inclusion criteria in these studies were less stringent than those in RE-LY, and may provide better generalizability to clinical practice. Unlike RE-LY, observational cohort studies did not exclude patients with severe renal impairment (creatinine clearance ≤30 mL/min), active liver disease, or conditions associated with an increased risk of bleeding. Furthermore, patients receiving dabigatran-110 mg in clinical practice, tended to be older than patients in the RE-LY trial who were randomized to this dose. Such discrepancies in eligibility criteria and treatment selection bias may explain some of the observed differences in findings between the RE-LY trial and observational studies, particularly when evaluating the lower dose of dabigatran. Combined, the 7 primary observational studies had a nearly 20-fold larger patient population than RE-LY, and a similar mean duration of follow-up. Taken together, this review represents an important contribution to the growing body of evidence for the comparative effectiveness and safety of dabigatran and warfarin from real-world clinical practice, and it has identified several gaps in knowledge for future investigation.

**Future Studies**

There are a limited number of well-designed studies evaluating dabigatran relative to warfarin in real-world clinical practice. Although follow-up in observational studies was on average ≥2 years, additional studies are needed to assess the long-term benefits and risks of dabigatran compared with warfarin. We note that there are few studies that evaluated dabigatran-110 mg, which is available only outside the United States. In our review, 2 of the 7 studies evaluated dabigatran-110 mg,\textsuperscript{30,31} and patients on this lower dose represented only 3% of the total evaluable population. The Food and Drug Administration’s decision not to approve dabigatran-110 mg in the United States was controversial because some clinicians think that this lower dose would allow for individualized treatment options in patients at greater risk for bleeding.\textsuperscript{31} It remains to be seen how subanalyses of the RE-LY trial and accumulating observational evidence may impact this decision.

There is little known about the safety and effectiveness of dabigatran-75 mg, which is available for patients with renal impairment in the United States and in Canada. This dose is increasingly used in elderly patients with a potentially high risk for gastrointestinal bleeding in the United States.\textsuperscript{18} Despite the availability of dabigatran-75 mg, it was not tested in RE-LY, and its approval was based not on efficacy or safety data, but on pharmacodynamic/pharmacokinetic modeling.\textsuperscript{32} Accordingly, more real-world evidence is needed to evaluate the benefits of lower dabigatran doses (110 and 75 mg), particularly in high-risk elderly patients.

Registry data will provide much needed information on real-world outcomes with dabigatran. The Global
Anticoagulant Registry in the FIELD (GARFIELD) study is an international longitudinal prospective registry of patients with AF receiving anticoagulation for stroke prevention. GARFIELD plans to enroll 55,000 patients at >1,000 clinical centers across 50 countries. Results from this study are eagerly awaited.

In addition to dabigatran, there are currently 3 other NOACs available in the United States for stroke prevention in NVAF, rivaroxaban, apixaban, and edoxaban, which were approved by the Food and Drug Administration in 2011, 2012, and 2015, respectively. Observational evidence for these agents is limited. Future studies ought to be designed to evaluate real-world experience with newer agents relative to dabigatran and warfarin.

**Role of Observational Studies in Comparative Effectiveness Research**

The synthesis of results from multiple and similarly designed RCTs is an important tool in comparative effectiveness research. Although RCTs provide the strongest level of evidence for drug safety and efficacy, multiple trials are often not conducted for a given drug and particular condition. In the absence of such evidence, well-designed observational studies serve an important role, and the synthesis of data from these studies can prove valuable, in evidence appraisal. High-quality observational studies, unlike many RCTs, may have better generalizability to real-world practice as they often have less stringent inclusion criteria and patients are treated in a clinical setting under usual care. We conducted this review within the framework that observational studies are an important contribution to the totality of evidence to understand the benefits and risks of an intervention, especially when there are gaps in evidence from RCTs. The goals of this review were achieved in that we successfully: (1) identified relevant and well-designed observational studies that evaluated dabigatran in stroke prevention, (2) assessed the quality of this evidence, (3) synthesized data and noted when heterogeneity in treatment effect was present, (4) put findings from observational studies in the context of evidence from RCTs, and (5) identified gaps in knowledge for future study.

The conduct of observational studies are inevitable, and results from such studies should complement, not take the place of, evidence derived from RCTs. The merits of observational evidence should be judged by both the quality of the evidence and the current gaps in knowledge from RCTs. Future observational studies that are published for dabigatran in stroke prevention should undergo similar scrutiny.

**Conclusions**

In this systematic review, real-world evidence shows that dabigatran is comparable with warfarin in preventing ischemic stroke in patients with NVAF. However, dabigatran is associated with a lower risk for intracranial bleeding relative to warfarin, but—particularly in the elderly—a greater risk for gastrointestinal bleeding. Bleeding outcomes from observational studies are consistent with those from the pivotal RE-LY trial. Future observational studies are needed to confirm findings from this review.

**Disclosures**

None.

**References**


Dabigatran Versus Warfarin for Atrial Fibrillation in Real-World Clinical Practice: A Systematic Review and Meta-Analysis
Robert J. Romanelli, Laura Nolting, Marina Dolginsky, Eunice Kym and Kathleen B. Orrico

Circ Cardiovasc Qual Outcomes. 2016;9:126-134; originally published online January 26, 2016; doi: 10.1161/CIRCOUTCOMES.115.002369
Circulation: Cardiovascular Quality and Outcomes is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2016 American Heart Association, Inc. All rights reserved.
Print ISSN: 1941-7705. Online ISSN: 1941-7713

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circoutcomes.ahajournals.org/content/9/2/126

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation: Cardiovascular Quality and Outcomes can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation: Cardiovascular Quality and Outcomes is online at:
http://circoutcomes.ahajournals.org//subscriptions/