

Comparative Efficacy and Safety of New Oral Anticoagulants in Patients With Atrial Fibrillation

Sebastian Schneeweiss, MD, ScD; Joshua J. Gagne, PharmD, ScD; Amanda R. Patrick, MS; Niteesh K. Choudhry, MD, PhD; Jerry Avorn, MD

Background—Dabigatran, an oral thrombin inhibitor, and rivaroxaban and apixaban, oral factor Xa inhibitors, have been found to be safe and effective in reducing stroke risk in patients with atrial fibrillation. We sought to compare the efficacy and safety of the 3 new agents based on data from their published warfarin-controlled randomized trials, using the method of adjusted indirect comparisons.

Methods and Results—We included findings from 44535 patients enrolled in 3 trials of the efficacy of dabigatran (RE-LY), apixaban (ARISTOTLE), and rivaroxaban (ROCKET-AF), each compared with warfarin. The primary efficacy end point was stroke or systemic embolism; the safety end point we studied was major hemorrhage. To address a lack of comparability between trial populations caused by the restriction of ROCKET-AF to high-risk patients, we conducted a subgroup analysis in patients with a CHADS₂ score ≥ 3 . We found no statistically significant efficacy differences among the 3 drugs, although apixaban and dabigatran were numerically superior to rivaroxaban. Apixaban produced significantly fewer major hemorrhages than dabigatran and rivaroxaban.

Conclusion—An indirect comparison of new anticoagulants based on existing trial data indicates that in patients with a CHADS₂ score ≥ 3 dabigatran 150mg, apixaban 5mg, and rivaroxaban 20mg resulted in statistically similar rates of stroke and systemic embolism, but apixaban had a lower risk of major hemorrhage compared with dabigatran and rivaroxaban. Until head-to-head trials or large-scale observational studies that reflect routine use of these agents are available, such adjusted indirect comparisons based on trial data are one tool to guide initial therapeutic choices. (*Circ Cardiovasc Qual Outcomes*. 2012;5:00-00.)

Key Words: indirect comparison ■ anticoagulation ■ dabigatran ■ rivaroxaban ■ apixaban ■ warfarin ■ randomized controlled trial

Optimal clinical practice and coverage decision making require assessment of the effectiveness and safety of new medications relative to existing treatments,¹ but comparative information is often unavailable at the time of marketing authorization and initial use.² Recently, phase III trials have been completed comparing each of the 3 new oral anticoagulants to warfarin in the treatment of atrial fibrillation (AF).

Dabigatran, an oral direct thrombin inhibitor, was more efficacious than warfarin in reducing the risk of stroke when given at a dose of 150 mg BID to patients with nonvalvular AF (hazard ratio [HR], 0.66; 95% CI, 0.53–0.82).³ Rivaroxaban 20 mg QD and apixaban 5 mg BID, both oral factor Xa inhibitors, were found to be noninferior (rivaroxaban: HR, 0.88; 95% CI, 0.74–1.03) and superior (apixaban: HR, 0.79; 95% CI, 0.66–0.95) to warfarin, respectively, in intention-to-treat analyses.^{4,5} Warfarin's narrow therapeutic index and risk of interactions, both of which contribute to the need for frequent monitoring, make alternative agents appealing. These new

agents also offer a welcome therapeutic option for patients ineligible for warfarin.⁶

In the absence of data from direct comparisons of the new anticoagulants, clinicians and payors attempting to engage in evidence-based decision making are likely to turn to qualitative comparisons of the results from the treatment arms of the 3 trials. Such comparisons may be helpful or they may be inconclusive or misleading because of the differences in trial design and patient populations.⁷ The formal method of adjusted indirect comparison, which compares treatment effects versus a common comparator, has been proposed as an alternative to simple, naive indirect comparisons.⁸ Two recent meta-analyses have compared results of adjusted indirect comparisons, such as ours, with subsequent direct comparisons. These meta-analyses have found the results of adjusted indirect and direct comparisons to be generally consistent, with statistically similar findings in 86%⁹ to 93%⁷ of cases.

Received January 30, 2012; accepted May 21, 2012.

From the Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA.

The online-only Data Supplement is available at <http://circoutcomes.ahajournals.org/lookup/suppl/doi:10.1161/CIRCOUTCOMES.112.965988/-/DC1>.

Correspondence to Sebastian Schneeweiss, MD, ScD, Division of Pharmacoepidemiology and Pharmacoeconomics, 1 Brigham Circle, Suite 3030, Boston, MA 02120. E-mail schneeweiss@post.harvard.edu

© 2012 American Heart Association, Inc.

Circ Cardiovasc Qual Outcomes is available at <http://circoutcomes.ahajournals.org>

DOI: 10.1161/CIRCOUTCOMES.112.965988

WHAT IS KNOWN?

- Dabigatran, an oral thrombin inhibitor, and rivaroxaban and apixaban, oral factor Xa inhibitors, have been found to be safe and effective in reducing stroke risk in patients with atrial fibrillation compared with warfarin.
- In the absence of data from direct comparisons of these new anticoagulants, adjusted indirect comparison can be used to compare treatment effects using warfarin as a common comparator group.

WHAT THIS STUDY ADDS

- Patients in 3 major warfarin-controlled randomized trials were comparable when limited to those with a CHADS₂ score of ≥ 3 .
- We found no statistically significant differences among the 3 drugs in their efficacy in preventing stroke and systemic embolism, although apixaban and dabigatran were numerically superior to rivaroxaban.
- Apixaban produced statistically significant fewer major hemorrhages than dabigatran and rivaroxaban.

In a recent indirect comparison of new oral anticoagulants in patients with AF,¹⁰ the authors failed to assure comparability of trials, as ROCKET-AF required a CHADS₂ score ≥ 2 , whereas the other trials also included patients with scores of 0 and 1. Because CHADS₂ is a strong predictor of the primary end point in all trials, the findings of such an indirect comparison are likely confounded. They further violated the comparability of trials by contrasting the major bleed rates from an ITT analysis of RE-LY to the results from the on-treatment analyses of ARISTOTLE and ROCKET-AF. We conducted an adjusted indirect comparison using the clinical trial data available on the 3 new anticoagulants, with warfarin as a common reference treatment. To overcome limitations of the previous study, we abstracted additional data so as to compare more similar analyses and patient populations.

Methods

Search Strategy

We searched PubMed and clinicaltrials.gov in October 2011 for all phase III randomized controlled trials (RCTs) of patients receiving apixaban, rivaroxaban, or dabigatran versus warfarin (the only studied common comparison treatment) for the prevention of thrombotic events in AF. We restricted our analysis to studies reporting event rates or containing sufficient data to calculate event rates. We also searched FDA advisory committee briefing materials and approval packages for any relevant supplemental data on dabigatran and rivaroxaban; apixaban has not yet been reviewed by FDA for the indication of prophylaxis in AF.

Data Abstraction

Data were abstracted independently by 2 authors for a primary efficacy end point of stroke or systemic embolism, a secondary efficacy end point of all-cause mortality, and the safety outcome of major hemorrhage. We identified intention-to-treat as the appropriate analytical strategy because this approach preserves the random treatment assignment within an RCT. In addition to end points, we abstracted

data on inclusion and exclusion criteria, end point definitions, risk factors for stroke in the warfarin control groups, and study design and analysis to assess the comparability of the trials.

Analysis

We performed comparisons across the trials using Bucher's method,^{8,11} an adjusted indirect comparison approach that analyzes the magnitude of relative treatment effects against a common comparator rather than absolute event rates in individual study arms to estimate the comparative safety and efficacy of ≥ 2 treatments. We used this approach to compare relative event rates among patients treated with apixaban, dabigatran (150 mg), and rivaroxaban versus warfarin. We focused on dabigatran 150 mg, the formulation approved in the United States, but also assessed the comparative safety and efficacy of dabigatran 110 mg in a secondary analysis, because this formulation may be of interest outside the United States. Bucher's method rests on the assumption that the trials or subgroups within trials are sufficiently similar with respect to potential clinical and methodological modifiers of relative treatment effects, such as patient characteristics, intervention characteristics, follow-up time, outcome definitions and ascertainment (clinical moderators), and randomization and blinding (methodological moderators). We assessed the validity of this assumption by evaluating the level of clinical and methodological similarity among the clinical trials and by comparing outcome event rates among warfarin-treated control patients across the different trials to assess treatment effectiveness in the common referent group. Because the majority of patients in ROCKET-AF had CHADS₂ scores ≥ 3 , we compared rivaroxaban with apixaban and dabigatran in an analysis restricted to this population.¹²

Results

We included findings from 44 535 patients with AF enrolled in 3 randomized trials, comparing the efficacy of dabigatran 150 mg BID (RE-LY),^{4,13} rivaroxaban 20 mg QD (ROCKET-AF),⁵ and apixaban 5 mg BID (ARISTOTLE)⁶ with warfarin. Inclusion and exclusion criteria were generally similar across trials, except that ROCKET-AF participants were required to have a CHADS₂ score ≥ 2 ; this was reflected in the higher proportions of patients with stroke risk factors in both the rivaroxaban and the warfarin control groups of ROCKET-AF (Table 1). The trial populations were similar in age, sex distribution, systolic blood pressure, and prevalence of a history of myocardial infarction at baseline. A higher proportion of warfarin-treated patients in RE-LY were classified as having paroxysmal AF (33.8%) than those in ARISTOTLE (15.5%) or ROCKET-AF (17.8%). Prior vitamin K antagonist use was more common in warfarin users in ROCKET-AF (62.5%) than in those in ARISTOTLE (57.2%) or RE-LY (48.6%).

Stroke or systemic embolism was the primary efficacy end point in all 3 trials. Major bleeding was defined consistently across trials as clinically overt bleeding that occurred at a critical site or bleeding that was accompanied by a drop in hemoglobin of at least 2 g/dL or transfusion of at least 2 units of packed red cells, or death. Major bleeding was the primary safety outcome in ARISTOTLE and RE-LY and was a component of the primary safety outcome of major or nonmajor clinically relevant bleeding in ROCKET-AF. Primary efficacy data were analyzed according to an intention-to-treat principle in all trials, although the ITT analysis was reported as a secondary analysis in ROCKET-AF. Only RE-LY reported results from an ITT analysis as the primary safety analysis; as-treated results were available from all trials. Two-year treatment discontinuation rates in RE-LY were 20.7% for 110 mg dabigatran,

Table 1. Selected Patient Characteristics in the Warfarin Control Groups of the RE-LY (Dabigatran Versus Warfarin), ARISTOTLE (Apixaban Versus Warfarin), and ROCKET-AF (Rivaroxaban Versus Warfarin) Trials

	ARISTOTLE	RE-LY	ROCKET AF
	Warfarin group (N=9081)	Warfarin group (N=6022)	Warfarin group (N=7133)
Patient characteristics			
Age, median (IQR) or mean±SD	70 (63–76)	71.6±8.6	73 (65–78)
Female, %	35.0	36.7	39.7
Systolic blood pressure, median (IQR) or mean±SD	130 (120–140)	131.2±17.4	130 (120–140)
Weight in kg, median (IQR) or mean±SD	82 (70–95)	82.7±19.7	NR
Prior myocardial infarction, %	13.9	16.1	18.0
Prior spontaneous or clinically relevant bleeding, %*	16.7	NR	NR
Type of atrial fibrillation, %			
Newly diagnosed	NR	NR	1.4
Paroxysmal	15.5	33.8	17.8
Persistent/permanent	84.4	66.1	80.8
Prior use of vitamin K antagonist, %†	57.2	48.6	62.5
Prior aspirin use, %	30.5	40.6	36.7
Stroke risk factors			
Prior stroke/TIA/systemic embolism, %	19.7	19.8	54.6
Heart failure or reduced ejection fraction, %	35.4	31.9	62.3
Diabetes mellitus, %	24.9	23.4	39.5
Hypertension, %	87.6‡	78.9	90.8
CHADS ₂ score			
Mean±SD	2.1±1.1	2.1±1.1	3.5±1.0
≤1, %	34.0	30.9	0.0
2, %	35.8	37.0	13.1
≥3, %	30.2	32.1	86.9
Target INR for warfarin group	2–3	2–3	2–3
Mean time in target INR range	62%	64%	55%
Warfarin management, blinding	Blinded dose adjustment algorithm	Unblinded dose adjustment with advice on INR control	Blinded INR values provided
Median follow-up, years	1.8	2.0	1.9
Double blinded	Yes	No§	Yes

*RE-LY and ROCKET AF excluded patients with recent gastrointestinal bleeding or a history of intracranial, intraocular, spinal, retroperitoneal or atraumatic intra-arterial bleeding.

†In ARISTOTLE, prior vitamin K antagonist use was defined as use >30 consecutive days; in RE-LY it was defined as lifetime use of ≥61 days; the definition was not reported in the ROCKET AF publication.

‡Defined as hypertension requiring treatment.

§Warfarin was given in an unblinded manner.

IQR indicates interquartile range; NR, not reported.

21.2% for 150 mg dabigatran, and 16.6% for warfarin; 2-year discontinuation rates in ROCKET-AF were 34.7% for rivaroxaban and 33.5% for warfarin; and total discontinuation rates in ARISTOTLE were 25.3% for apixaban and 27.5% for warfarin.

The trials differed in blinding and warfarin dosing. ROCKET-AF investigators were blinded to anticoagulation group assignment and were provided with real (warfarin arm) or sham (rivaroxaban arm) INR values; ARISTOTLE investigators were similarly blinded to group assignment and were provided with a blinded dose adjustment algorithm¹⁴ and feedback. In contrast, RE-LY investigators were unblinded to warfarin-group status, regulating its dose with advice on INR control. During the trials, INR control was better in the warfarin arms of ARISTOTLE (62% mean time in therapeutic

range) and RE-LY (64% mean time in therapeutic range) than in ROCKET-AF (55% mean time in therapeutic range). The stroke and embolism event rate was substantially higher in the warfarin arm of ROCKET-AF (2.42 per 100 patient-years; estimated 95% CI, 2.16–2.71) compared with the arm in RE-LY (1.71 per 100 patient-years; estimated 95% CI, 1.49–1.96) or ARISTOTLE (1.60 per 100 patient-years; estimated 95% CI, 1.42–1.80) (Table 2). These differences were not present when patients with CHADS₂ scores ≥3 were compared (Table 3).

Because of increased baseline stroke risk among the participants of ROCKET-AF compared with ARISTOTLE and RE-LY, we conducted 2 analyses: (1) an indirect comparison of apixaban and dabigatran among all participants, and (2) a comparison of all 3 agents in participants with a CHADS₂

Table 2. Users of Dabigatran (RE-LY*), Apixaban (ARISTOTLE), and Rivaroxaban (ROCKET-AF) and Rates of Stroke and Systemic Embolism, Ischemic Stroke, Death of Any Cause, and Major Hemorrhage Compared With Warfarin

	New Anticoagulant			Warfarin			Hazard Ratio (95% CI)
	Subjects	Events	Event Rate per 100 Person-Years	Subjects	Events	Event Rate (95% CI) per 100 Person-Years†	
Primary efficacy end point: stroke or systemic embolism (intention-to-treat analysis)							
Apixaban (ARISTOTLE)	9120	212	1.27	9081	265	1.60 (1.42–1.80)	0.79 (0.66–0.95)
Dabigatran 110 mg (RE-LY)	6015	183	1.54	6022	202	1.71 (1.49–1.96)	0.90 (0.74–1.10)
Dabigatran 150 mg (RE-LY)	6076	134	1.11				0.66 (0.52–0.81)
Rivaroxaban (ROCKET-AF)‡	7081	269	2.12	7090	306	2.42 (2.16–2.71)	0.88 (0.75–1.03)
Death of any cause (intention-to-treat analysis)							
Apixaban (ARISTOTLE)	9120	603	3.52	9081	669	3.94 (3.65–4.25)	0.89 (0.80–1.00)
Dabigatran 110 mg (RE-LY)	6015	446	3.75	6022	487	4.13 (3.78–4.51)	0.91 (0.80–1.03)
Dabigatran 150 mg (RE-LY)	6076	438	3.64				0.88 (0.77–1.00)
Rivaroxaban (ROCKET-AF)‡	7081	582	4.5d	7090	632	4.9§ (4.53–5.29)	0.92 (0.82–1.03)
Major hemorrhage (on-treatment analysis)							
Apixaban (ARISTOTLE)	9088	327	2.13	9052	462	3.09 (2.82–3.38)	0.69 (0.60–0.80)
Dabigatran 110 mg (RE-LY)¶	NR	NR	NR	NR	NR	NR	0.83 (0.71–0.96)
Dabigatran 150 mg (RE-LY)¶	NR	NR	NR				0.98 (0.85–1.14)
Rivaroxaban (ROCKET-AF)¶, #	7111	395	3.60	7125	386	3.45 (3.12–3.81)	1.04 (0.90–1.20)

*Includes updated data published by Connolly et al in *N Engl J Med*, 2010.¹³

†95% CIs were estimated using the method for person-time data described in Ref. 15. We estimated person-time by dividing the number of events by the event rate.

‡The primary analysis in ROCKET-AF was per-protocol, restricted to subjects who received at least 1 study dose followed through 2 days after the last dose.

The intention-to-treat analyses of the primary efficacy end point and death presented here were published as secondary analyses.

§Mortality rates in ROCKET-AF were published to only 1 decimal point.

¶The published RE-LY safety analysis was intention-to-treat (hazard ratio, 0.93; 95% CI, 0.81–1.07 for dabigatran 150 mg and hazard ratio, 0.80; 95% CI, 0.69–0.93 for dabigatran 110 mg). Hazard ratios and 95% CIs for the on-treatment analysis were presented in the FDA Advisory Committee briefing documents.¹⁶

¶¶In ROCKET-AF major bleeding was a component of the primary safety outcome: major and clinically relevant nonmajor bleeding.

#For the ROCKET-AF trial, the number of subjects is greater in the on-treatment safety analysis than in the intention-to-treat efficacy analysis because the latter excluded data on 93 patients (50 in the rivaroxaban group and 43 in the warfarin group) due to violations in Good Clinical Practice guidelines at 1 participating site.

NR indicates not reported.

score ≥ 3 . In adjusted indirect comparisons of the overall trial populations (Figure), we found no significant differences in the rate of stroke or systemic embolism between dabigatran and apixaban. Apixaban was associated with a 30% lower rate of major hemorrhage than dabigatran (HR, 0.70; 95% CI, 0.57–0.86) (Figure). There were no significant differences among the agents in all-cause mortality (Table 2).

In subgroups of patients with CHADS₂ scores ≥ 3 , which ensured fairer and less confounded comparisons among rivaroxaban, apixaban, and dabigatran concerning baseline stroke risks, both dabigatran and apixaban reduced the risk of stroke and embolism by about 20% compared with rivaroxaban, without reaching statistical significance (Figure). In the same subgroup, the risk of major hemorrhage was again lowest for apixaban compared with both other drugs (Figure).

A secondary analysis comparing dabigatran 110 mg with apixaban among all subjects and with rivaroxaban among the subgroup of patients with CHADS₂ scores ≥ 3 found no significant differences between agents in the efficacy or risk of major hemorrhage (online-only Data Supplement).

Discussion

Formal adjusted indirect comparisons of randomized trials can be a useful tool to estimate the comparative efficacy of newly marketed medications before any head-to-head trials or

pharmacoepidemiological studies become available, if certain study characteristics are met. We compared all subjects in ARISTOTLE and RE-LY anticoagulant trials, as well as patients with an elevated risk for stroke (CHADS₂ ≥ 3), to make the results of 1 trial (ROCKET-AF) more comparable. Among all participants, apixaban had a lower risk of major bleeding than dabigatran, with no significant differences in efficacy. Among patients with CHADS₂ ≥ 3 , dabigatran and apixaban trended toward greater efficacy than rivaroxaban, but the differences were not significant. Major bleeding risk on apixaban was significantly less than that on rivaroxaban or dabigatran.

We used intention-to-treat results for efficacy end points and on-treatment analyses for safety end points. Intention-to-treat is arguably the preferred approach for analyzing data from an RCT because it preserves random treatment assignment. The lack of complete intention-to-treat safety data from ARISTOTLE and RE-LY is a limitation. The use of an on-treatment approach with a short risk window of only 2 days after the last dose anticoagulant for the safety analysis raises the possibility that bleeding events attributable to anticoagulant use that manifest after 2 days would have been missed. In RE-LY, where an ITT analysis was reported as the primary safety analysis, the ITT HR for major bleeding for 110 mg dabigatran compared with warfarin was 0.80 (95% CI, 0.69–0.93) versus 0.83 (95% CI, 0.71–0.96) for an

Table 3. Primary Efficacy and Safety in Subgroups of Trial Patients With CHADS₂ Score ≥ 3

	New Anticoagulant			Warfarin			Hazard Ratio (95% CI)
	Subjects	Events	Event Rate per 100 Person-Years	Subjects	Events	Event Rate (95% CI) per 100 Person-Years*	
Primary efficacy end point: stroke or systemic embolism (intention-to-treat analysis)							
Apixaban (ARISTOTLE)	2758	94	1.9†	2744	132	2.8† (2.35–3.31)	0.68 (0.52–0.88)
Dabigatran 110 mg (RE-LY)‡	1968	82‡	2.12	1933	101§	2.68 (2.19–3.24)	0.79 (0.59, 1.05)¶
Dabigatran 150 mg (RE-LY)‡	1981	74‡	1.88	1933	101§	2.68 (2.19–3.24)	0.70 (0.52, 0.95)¶
Rivaroxaban (ROCKET-AF)¶	6156	239	2.25	6155	270	2.56 (2.27–2.88)	0.88 (0.74, 1.05)
Major Hemorrhage (on-treatment analysis)							
Apixaban (ARISTOTLE)	NR#	126	2.9†	NR#	173	4.2† (3.61–4.86)	0.69 (0.55–0.87)
Dabigatran 110 mg (RE-LY) approximated**	1966	147	3.80	1931	172	4.61 (3.96–5.34)	0.82 (0.66–1.03)¶
Dabigatran 150 mg (RE-LY) approximated**	1979	188	4.86	1931	172	4.61 (3.96–5.34)	1.05 (0.86–1.30)¶
Rivaroxaban (ROCKET-AF)††	6187	337	3.64	6191	337	3.60 (3.23–4.00)	1.01 (0.87–1.18)

*95% CIs were estimated using the method for person-time data described in Ref. 15. We estimated person-time by dividing the number of events by the event rate.

†Event rates in ARISTOTLE were published to only 1 decimal point.

‡This analysis does not include the updated data published by Connolly et al in *N Engl J Med*, 2010, as subgroup-specific data were not reported.¹³

§The number of events in this subgroup was not reported in RE-LY; we estimated these by multiplying the number of subjects by the median follow-up time (2.0 years) and then multiplying this by the event rate.

¶ICIs around the hazard ratio for RE-LY were estimated using the mid-*P*-value function described in Ref. 15. We estimated person-time by dividing the number of events by the event rate.

¶¶The primary analysis in ROCKET-AF was per-protocol, restricted to subjects who received at least 1 study dose followed through 2 days after the last dose. The intention-to-treat analysis of the primary efficacy end point presented here was published as a secondary analysis.

#The number of subjects included in the on-treatment analysis was not presented in the publication.

**An on-treatment analysis was not reported for subgroups within RE-LY. The numbers presented here are from the intention-to-treat analysis presented in the FDA Advisory Committee briefing documents.¹⁶

††For the ROCKET-AF trial, the number of subjects is greater in the on-treatment safety analysis than in the intention-to-treat efficacy analysis because the latter excluded data on 93 patients (50 in the rivaroxaban group and 43 in the warfarin group) due to violations in Good Clinical Practice guidelines at 1 participating site.

NR indicates not reported.

on-treatment analysis; for 150 mg dabigatran, the ITT HR was 0.93 (0.81–1.07) versus 0.98 (0.85–1.14) for an on-treatment analysis. ARISTOTLE reported a 27% reduction in major bleeding from a secondary ITT analysis compared with a 31% reduction in an on-treatment analysis.

Bucher's method of adjusted indirect comparisons assumes that the trials are sufficiently similar to one another in terms of potential modifiers of treatment effects, such that the relative efficacy of a given treatment compared with a common comparator would be homogeneous across each of the study populations. This assumption was likely violated for ROCKET-AF, which was restricted to patients with a CHADS₂ score ≥ 2 , and therefore restricted our comparisons of rivaroxaban to subgroups with CHADS₂ scores ≥ 3 . Reassuringly, event rates in the warfarin comparison groups of all 3 trials among this subgroup were similar. However, it is still possible that other factors, such as the poorer INR control in the warfarin arm of ROCKET-AF, could threaten the validity of our analysis. Although primary outcome event rates were similar among warfarin-treated patients across trials, the major hemorrhage rate was slightly lower among warfarin-treated patients in the ROCKET-AF trial compared with other trials, raising the possibility that another unobserved factor could influence our results. Data on the baseline characteristics of the subgroups with CHADS₂ scores ≥ 3 were not available, precluding further evaluation of differences in baseline risk.

In addition to clinical similarity, Bucher's method requires methodological similarity. Biases inherent to each of the trials may cause indirect comparisons to be biased, unless trials are

all biased in the same direction and to the same extent. It is not possible to predict the direction and magnitude of bias in specific instances. Although blinding generally reduces bias, the lack of blinding in RE-LY may have resulted in its greater mean time in therapeutic range on warfarin, thus increasing the apparent relative effectiveness of warfarin compared with dabigatran. Lack of complete blinding can have a smaller effect on end point ascertainment if objectively measured end points are of interest, as in our study.¹⁷

Our analysis was limited by the available RCT data. Apixaban has not yet been reviewed by FDA for the indication of thromboembolic event prophylaxis in AF, and thus advisory committee briefing materials are not yet available. Rivaroxaban has not been evaluated in lower-risk patients, and although it was possible to compare rivaroxaban with apixaban and dabigatran in patients with CHADS₂ ≥ 3 , this restriction resulted in reduced statistical power.

Adjusted indirect treatment comparisons such as ours generally produce reliable estimates. A recent study found that results of indirect and head-to-head comparisons were statistically similar in 86% of 112 analyses, which covered a wide range of clinical indications, and in only 1 case (<1%) the indirect comparison results were statistically significant and in the opposite direction of the corresponding head-to-head comparison results.⁹ Statistically inconsistent findings were more common when fewer trials were available, outcomes were subjective, and a statically significant treatment effect was found by either the direct or indirect comparison. Indirect comparisons were less likely to find statistically significant

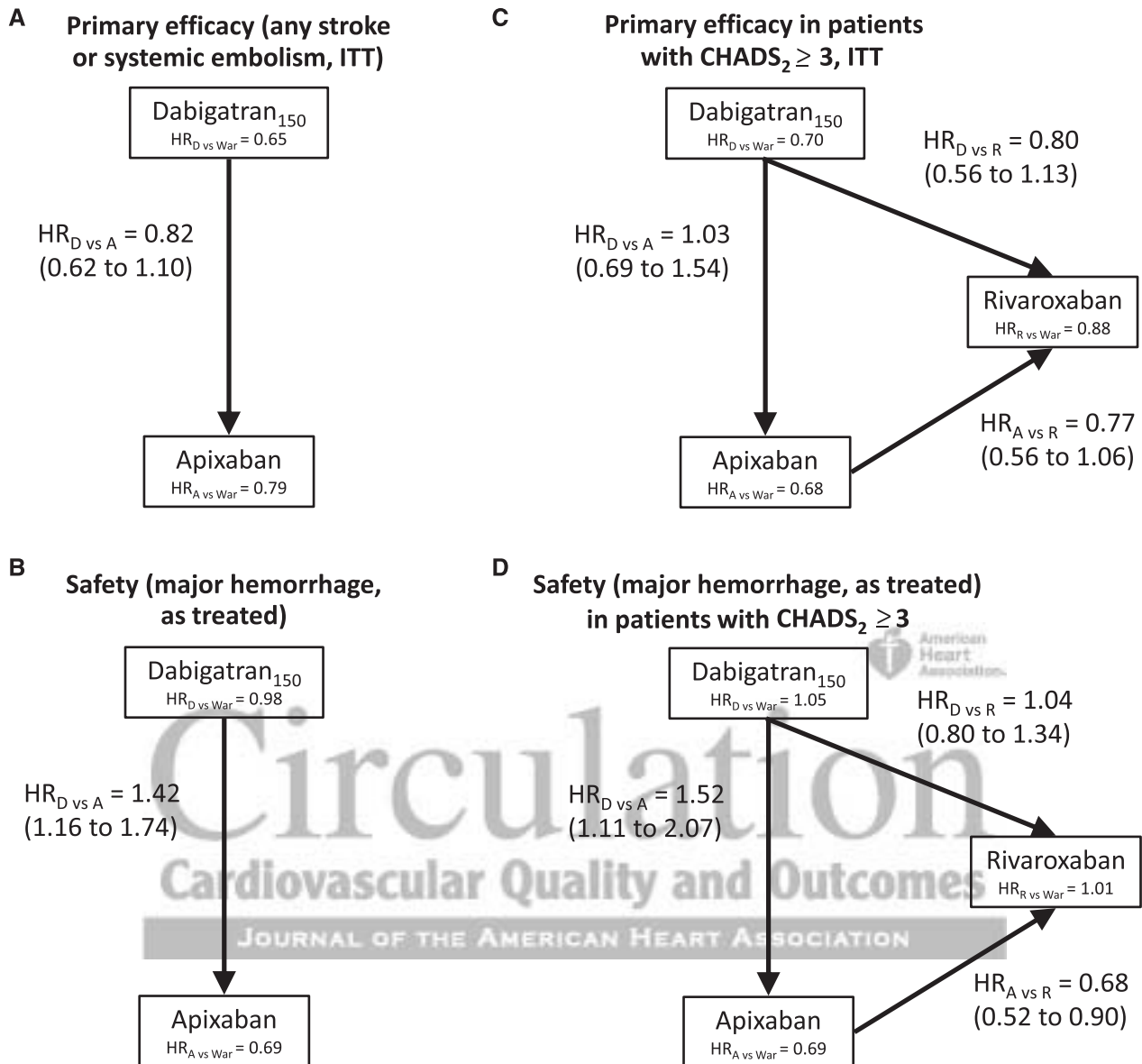


Figure. The relative efficacy and safety of dabigatran (150 mg), apixaban, and rivaroxaban from adjusted indirect comparisons of 3 randomized trials. Hazard ratios with 95% CIs. The agent at the head of the arrow is the referent agent (ie, the denominator term).

differences between treatments than direct comparisons, which are 4× as precise for the same sample size.⁹ Power was further limited in our analysis because we focused on a high-risk patient subgroup to enhance comparability across trials. These findings underscore the importance of evaluating clinical and methodological similarity when conducting indirect comparisons, as we have done,¹⁸ and following up initial preapproval trials with rigorously conducted observational studies, which can provide direct comparisons in large populations.¹⁹

Optimal clinical practice and policy making require assessing the comparative effectiveness of newly approved medications as soon as possible after their market authorization. The 3 new oral anticoagulants that we compared have all been found to be superior or at least noninferior to warfarin in large randomized trials of stroke prevention in patients with AF, and they may be particularly good options for patients ineligible

for warfarin or for whom INR monitoring cannot be performed adequately. The differences we found among the new anticoagulants, and between each of them and warfarin, are modest in scale compared with the substantial differences in outcome between no anticoagulant and any anticoagulant in AF patients.

In the absence of head-to-head comparisons from randomized trials or effectiveness studies using routine care data, adjusted indirect comparisons using data from RCTs can provide useful approximations for clinical decision makers, information that can be generated even before a drug is authorized for marketing for a given indication, as with apixaban. By supplementing data published in the medical literature with data from FDA advisory committee briefing documents for dabigatran and rivaroxaban, examining results from secondary and subgroup analyses to increase comparability across trials in study designs and populations, performing a

quantitative evaluation, and highlighting differences between the trials, our analysis provides an improvement over the qualitative comparisons that decision makers may undertake, which may be inconclusive or misleading because of the differences in trial design. Failure to insure comparable populations may bias results.¹⁰ However, the fact that apixaban has not yet been reviewed by FDA for the indication of stroke prophylaxis in AF limits the availability of supplemental data for this particular agent. In the case of the new anticoagulants, our understanding of the relative effectiveness and safety of these drugs may well change in the context of routine care depending on actual adherence patterns, baseline risks of patients who are prescribed these drugs outside of trial criteria, and other real-world differences that may not be predicted by trial results. Follow-up active observational studies that monitor the comparative safety and effectiveness of newly marketed drugs using health care claims data can then provide such follow-up information in a timely manner.¹ Because such studies can themselves take some time to organize and conduct, rigorous comparisons of existing trial data—if feasible and responsibly conducted—may offer the best basis for evidence-based decision making when new drugs are first brought to market.

Sources of Funding

This work was funded by the National Library of Medicine (RO1-LM010213), the National Center for Research Resources (RC1-RR028231), and the National Heart Lung and Blood Institute (RC4-HL106373).

Disclosures

Dr Schneeweiss is a principal investigator of the Brigham and Women's Hospital DEcIDE Center on Comparative Effectiveness Research and the DEcIDE Methods Center, both funded by AHRQ and of the Harvard-Brigham Drug Safety and Risk Management Research Center funded by FDA. Dr Schneeweiss is a paid consultant to the consulting firms WHISCON and Booz & Co. Dr Schneeweiss is a principal investigator of investigator-initiated grants to the Brigham and Women's Hospital Division of Pharmacoepidemiology and Pharmacoeconomics from Novartis, Pfizer, and Boehringer-Ingelheim on unrelated topics but not on any medications discussed in this article. The other authors report no conflicts.

References

- Schneeweiss S, Gagne JJ, Glynn RJ, Ruhl M, Rassen JA. Assessing the comparative effectiveness of newly marketed medications: methodological challenges and implications for drug development. *Clin Pharmacol Ther.* 2011;90:777–790.
- Goldberg NH, Schneeweiss S, Kowal MK, Gagne JJ. Availability of comparative efficacy data at the time of drug approval in the United States. *JAMA.* 2011;305:1786–1789.
- Connolly SJ, Ezekowitz MD, Yusuf S, Eikelboom J, Oldgren J, Parekh A, Pogue J, Reilly PA, Themeles E, Varrone J, Wang S, Alings M, Xavier D, Zhu J, Diaz R, Lewis BS, Darius H, Diener HC, Joyner CD, Wallentin L; RE-LY Steering Committee and Investigators. Dabigatran versus warfarin in patients with atrial fibrillation. *N Engl J Med.* 2009;361:1139–1151.
- Patel MR, Mahaffey KW, Garg J, Pan G, Singer DE, Hacke W, Breithardt G, Halperin JL, Hankey GJ, Piccini JP, Becker RC, Nessel CC, Paolini JF, Berkowitz SD, Fox KA, Califf RM; ROCKET AF Investigators. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. *N Engl J Med.* 2011;365:883–891.
- Granger CB, Alexander JH, McMurray JJ, Lopes RD, Hylek EM, Hanna M, Al-Khalidi HR, Ansell J, Atar D, Avezum A, Bahit MC, Diaz R, Easton JD, Ezekowitz JA, Flaker G, Garcia D, Gerasides M, Gersh BJ, Golitsyn S, Goto S, Hermosillo AG, Hohnloser SH, Horowitz J, Mohan P, Jansky P, Lewis BS, Lopez-Sendon JL, Pais P, Parkhomenko A, Verheugt FW, Zhu J, Wallentin L; ARISTOTLE Committees and Investigators. Apixaban versus warfarin in patients with atrial fibrillation. *N Engl J Med.* 2011;365:981–992.
- Wittkowsky AK. Effective anticoagulation therapy: defining the gap between clinical studies and clinical practice. *Am J Manag Care.* 2004;10(10 Suppl):S297–306; discussion S312.
- Song F, Altman DG, Glenny AM, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ.* 2003;326:472.
- Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol.* 1997;50:683–691.
- Song F, Xiong T, Parekh-Bhurke S, Loke YK, Sutton AJ, Eastwood AJ, Holland R, Chen YF, Glenny AM, Deeks JJ, Altman DG. Inconsistency between direct and indirect comparisons of competing interventions: meta-epidemiological study. *BMJ.* 2011;343:d4909.
- Lip GY, Larsen TB, Skjøth F, Rasmussen LH. Indirect Comparisons of New Oral Anticoagulant Drugs for Efficacy and Safety When Used for Stroke Prevention in Atrial Fibrillation. *J Am Coll Cardiol.* 2012.
- Wells GA, Sultan SA, Chen L, Khan M, Coyle D. Indirect Evidence: Indirect Treatment Comparisons in Meta-Analysis. Ottawa, Canada: Canadian Agency for Drugs and Technologies in Health; 2009.
- Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of clinical classification schemes for predicting stroke: results from the National Registry of Atrial Fibrillation. *JAMA.* 2001;285:2864–2870.
- Connolly SJ, Ezekowitz MD, Yusuf S, Reilly PA, Wallentin L; Randomized Evaluation of Long-Term Anticoagulation Therapy Investigators. Newly identified events in the RE-LY trial. *N Engl J Med.* 2010;363:1875–1876.
- ROCKET-AF Study Investigators. Rivaroxaban—once daily, oral, direct factor Xa inhibition compared with vitamin K antagonism for prevention of stroke and Embolism Trial in Atrial Fibrillation: rationale and design of the ROCKET AF study. *Am Heart J.* 2010;159:340–347.e1.
- Greenland S, Rothman KJ. Introduction to categorical statistics. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology.* 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008:253–254.
- Boehringer Ingelheim. Advisory Committee Briefing Document: Dabigatran Etxelate (DE), DE 150 BID and DE 110 BID; 2010.
- Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG, Gluud C, Martin RM, Wood AJ, Sterne JA. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ.* 2008;336:601–605.
- Song F, Harvey I, Lilford R. Adjusted indirect comparison may be less biased than direct comparison for evaluating new pharmaceutical interventions. *J Clin Epidemiol.* 2008;61:455–463.
- Avorn J. In defense of pharmacoepidemiology—embracing the yin and yang of drug research. *N Engl J Med.* 2007;357:2219–2221.