



Comparative Effectiveness Review Disposition of Comments Report

Research Review Title:

Stroke Prevention in Patients With Atrial Fibrillation: A Systematic Review Update

Draft review available for public comment from February 5, 2018, to March 22, 2018.

Research Review Citation: Sanders GD, Lowenstern A, Borre E, Chatterjee R, Goode A, Sharan L, Allen LaPointe NM, Raitz G, Shah B, Yapa R, Davis JK, Lallinger K, Schmidt R, Kosinski A, Al-Khatib S. Stroke Prevention in Patients With Atrial Fibrillation: A Systematic Review Update. Comparative Effectiveness Review No. 214. (Prepared by the Duke Evidence-based Practice Center under Contract No. 290-2015-00004-I for AHRQ and PCORI.) AHRQ Publication No. 18(19)-EHC018-EF. PCORI Publication No. 2018-SR-04. Rockville, MD: Agency for Healthcare Research and Quality; October 2018. Posted final reports are located on the Effective Health Care Program [search page](#). DOI: <https://doi.org/10.23970/AHRQEPCCER214>.

Comments to Research Review

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each research review is posted to the EHC Program Web site or AHRQ Web site in draft form for public comment for a 45 day period. Comments can be submitted via the Web site, mail or E-mail. At the conclusion of the public comment period, authors use the commentators' submissions and comments to revise the draft research review.

Comments on draft reviews and the authors' responses to the comments are posted for public viewing on the Web site approximately 3 months after the final research review is published. Comments are not edited for spelling, grammar, or other content errors. Each comment is listed with the name and affiliation of the commentator, if this information is provided. Commentators are not required to provide their names or affiliations in order to submit suggestions or comments.

The tables below include the responses by the authors of the review to each comment that was submitted for this draft review. The responses to comments in this disposition report are those of the authors, who are responsible for its contents, and do not necessarily represent the views of the Agency for Healthcare Research and Quality.

Archived: This report is greater than 3 years old. Findings may be used for research purposes, but should not be considered current.



#	Commentator & Affiliation	Section	Comment	Response
1.	Peer Reviewer #1	Quality of Report	Superior	Thank you. No response needed
2.	Peer Reviewer #2	Quality of Report	Good	Thank you. No response needed
3.	Peer Reviewer #3	Quality of Report	Good	Thank you. No response needed
4.	Peer Reviewer #4	Quality of Report	Superior	Thank you. No response needed
5.	Peer Reviewer #5	Quality of Report	Good	Thank you. No response needed
6.	Peer Reviewer #6	Quality of Report	Good	Thank you. No response needed
7.	Peer Reviewer #1	General	This is an outstanding comprehensive tour of the landscape of risk analysis in AFIB. I appreciate that the authors highlight that ASA is no longer a recommends stroke prediction strategy for AFIB in the European guidelines and that AHA ACC have not yet been updated to match emphasis on OACs superiority for AFIB is important.	Thank you. No response needed
8.	Peer Reviewer #2	General	This appears to me to be carefully done, using appropriate methods to address well-defined and relevant key questions.	Thank you. No response needed

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
9.	Peer Reviewer #2	General	<p>For Key Questions 1 and 2, my main concern about clinical relevance is the emphasis on c-statistics and the corresponding lack of focus on synthesizing evidence about actual event rates. I believe that a risk scale will only be useful to patients (or their providers) if it gives them information about how likely they are to experience the event to be predicted. For example, if I score 3 on the CHA2DS2-VASc, then I want to know what my risk is, and it does not help me to know this scale's c-statistic. If the published rates at each given score cannot be synthesized because of heterogeneous populations, event definitions, ascertainment accuracy, etc., then I think the clinical utility of the risk score is questionable, even if it has high c-statistics within each individual study. I would like to see this report more clearly show and synthesize the evidence about event rates, because predicting these will drive the clinical utility.</p>	<p>We have modified the tables in KQ1 and KQ2 to better highlight the event rates for the various outcomes given the different risk scores and to aid the reader in comparing rates across studies.</p>

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
10.	Peer Reviewer #2	General	An additional possibility for improving clinical utility would be to ensure for Key Question 3 that quantitative estimates of differences between treatments are consistently given and properly emphasized. Notably, the Key Points on page 83-84 give only directions of differences and lack any information on the magnitudes. Because the magnitude of any advantage of one therapy over another is relevant for clinical decisions, I believe that inclusion of some quantitative information about the sizes of the differences could make these summaries much more useful. In addition, several tables lack quantitative estimates for some or all rows, despite having a column heading for “Magnitude of effect”.	Given the large number of comparisons and key points we have not added in the quantitative information to the key points – but instead include that information in the SOE tables and abstract. We do include the hazard and risk ratio information for select key points/outcomes.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
11.	Peer Reviewer #2	General	<p>Finally, in a high-level review such as this, I believe that it is important to follow best practices and to carefully avoid questionable interpretations. A very widespread problem in the medical literature is the misinterpretation of $p > 0.05$ as providing evidence for “no difference”. Unfortunately, some interpretations in this review appear to lapse toward that misinterpretation. I recommend that all interpretations of “no difference” or “no association”, for both individual studies and synthesized results, be reviewed and revised to better reflect the point estimates and confidence intervals, rather than just whether or not $p < 0.05$. Some guidelines are available at www.CTSpedia.org/ResultsInterpretation. The discussion of misconceptions 5 and 6 in the paper available at https://urldefense.proofpoint.com/v2/url?u=http-3A__amstat.tandfonline.com_doi_suppl_10.1080_00031305.2016.1154108_suppl-5Ffile_utas-5Fa-5F1154108-5Fsm5368.pdf&d=DwlFaQ&c=imBPVzF25OnBgGmVOlcsiEgHoG1i6YHLR0Sj_gZ4adc&r=9HI35bRoP7Lbw6vdF6tBw_ijOOJWHA7dyYcqHbH849A&m=qPmmKnezC5gLht44cxEd8lwM82FWw7L4lxpua08jrs&s=9bKyi4JbG5JMYxEBqRgsmwB1MmspifTT0grAOCIFMHc&e= also provides some explanation of this issue.</p>	Throughout we have modified the text to say “there is no evidence of a difference” or “there is no evidence of a benefit” rather than “there is no difference” or “there is no benefit” respectively.
12.	Peer Reviewer #3	General	This is an exhaustive study of important clinical questions with patient oriented outcomes emphasized. Kudos for that.	Thank you. No response needed
13.	Peer Reviewer #3	General	The assessment of study quality and overall strength of evidence used standard methods.	Thank you. No response needed



#	Commentator & Affiliation	Section	Comment	Response
14.	Peer Reviewer #3	General	The summary of key points were helpful and concise.	Thank you. No response needed
15.	Peer Reviewer #3	General	The forest plots were helpful and well done.	Thank you. No response needed
16.	Peer Reviewer #3	General	I may have missed it but I did not see if there were limitations on the type of studies. For example, were cross sectional studies included? And, it is not clear if poor quality articles were included.	We do not limit the type of observational studies allowed in the report though do evaluate the quality of the studies and their potential risk of bias. Given the methodological limitations of observational studies which do not use propensity matched controls, we do not include these findings in our quantitative analyses. A list of excluded studies is included in Appendix D.
17.	Peer Reviewer #5	General	The review of evidence on the topic is exhaustive, scholarly, and clear. As such it provides an up-to-date summary of the evidence in this important clinical problem. The authors are commended for their painstaking review and presentation of a prodigious amount of data.	Thank you. No response needed

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
18.	Peer Reviewer #5	General	<p>I have two general areas of concern about the report that I believe if addressed would improve its usefulness to policy makers and clinicians. I should disclose that I am a geriatrician by training and consequently have reviewed the report through that lens. The population of patients I see with atrial fibrillation consequently tends to be older with multiple comorbidities.</p> <p>The 2 areas of concern are: (1) the interpretation and discussion of the CHADS2 versus CHA2DS2-VASc prediction instruments, (2) the relative lack of addressing falls and multiple comorbidities as a significant factor in decision-making regarding atrial fibrillation.</p> <p>1. CHADS2 versus CHA2DS2-VASc The interpretation and discussion of the CHADS2 and CHA2DS2-VASc instruments would be strengthened by a more overt and frank discussion of the equivalence of the 2 instruments. The analysis of information provides very comparable c-statistics regarding the 2 instruments, and in fact there is mention made that they are not statistically different from each other.</p>	As noted in the revised report, the CHADS2 and CHAD2DS2-VASc scores are comparable in their prediction abilities and we now state this throughout the report (key messages, introduction, and results).

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
19.	Peer Reviewer #5	General	Specifically am concerned about your statement on page 187 that "our findings generally support professional guidelines' recommendation to use the CHA2DS2-VASc score for assessment of stroke risk." I believe the report needs to more explicitly state that when all of the comparative studies are considered, these 2 predictive instruments appear to be equal, period. The report already states on page 57 and several other places that the CHADS2 and CHA2DS2-VASc instruments have "comparable prediction abilities for stroke risk". How do you justify the conclusion stated on p.187 with all of the other findings and statements of equal predictive ability across the two instruments?	We have removed this sentence and now emphasize the comparability of CHADS2, CHA2DS2-VASc, and the ABC risk scores.
20.	Peer Reviewer #5	General	In addition to a more consistent message of equivalence between the CHADS2 and CHA2DS2-VASc, it would be helpful to point out that this conclusion is different from the conclusion reported in the 2014 AHA/ACC/HRS Guidelines for the Management of Patients with Atrial Fibrillation. The AHA/ACC report concluded that the CHA2DS2-VASc was superior, particularly for patients at lower levels of stroke risk, and came out much more definitively in favor of using this instrument over the CHADS2.	We now consistently refer to the comparability of these scores (and the ABC score) given the available evidence. We also highlight the difference between our findings and those of the AHA/ACC/HRS guidelines in the discussion.



#	Commentator & Affiliation	Section	Comment	Response
21.	Peer Reviewer #5	General	There is one other area of this topic that, if addressed, would strengthen this review update: the difference in frequencies of older patients classified at high stroke risk (and thus indicated for anticoagulation) by the 2 instruments. The CHA2DS2-VASc classifies virtually all older adults at being at high risk: all adults aged 75 and over, all women aged 65-74, and most men aged 65-74 by virtue of them also having HTN, HF, or vascular disease. The CHADS2 classifies a significant number of older adults as high risk, but certainly far fewer than the CHA2DS2-VASc. There must be data available demonstrating the frequencies of adults aged 65-74 and age 75 and over who would qualify for anticoagulation therapy across these 2 instruments. Providing this information would help to round out the discussion of the CHADS2 versus CHA2DS2-VASc and provide clinicians with a better context for making decisions about which instrument to use.	We agree that evaluating the tools in these specific patient populations could be useful but unfortunately this was not reported consistently across the included studies. We do include information on the ages represented in the individual studies in Appendix F.
22.	Peer Reviewer #5	General	2. Fall Risk and Other Comorbidities I could not find anywhere in the report mentioned fall risk and its relationship to decision-making for anticoagulation treatment among older adults with AF. This is an issue of huge concern to clinicians who are caring for frail elderly patients. I understand that the data on falls in this situation are scant but somewhere in the report it should be mentioned.	We agree this is of concern and that more data are needed on this issue. We have added this as a population of interest “fall risk and anticoagulation treatment among older adults” for future research.



#	Commentator & Affiliation	Section	Comment	Response
23.	Peer Reviewer #5	General	I believe there is also relatively cursory mention of multiple chronic conditions and their impact on stroke risk and side effects of anticoagulant therapy. Geriatricians and primary care clinicians are often presented with AF decision making in older adults who are more frail, and have a greater number of disparate comorbid conditions, than are seen in research subjects. Addressing this knowledge gap more overtly in the report would be helpful.	We acknowledge this is a concern and have added the following to the discussion: "In frail older patients, there may be concerns about using anticoagulation in the presence of multimorbidity due to a higher prevalence of pre-existing conditions that predispose to bleeding, concomitant interacting medications (antiplatelet therapy, nonsteroidal anti-inflammatory drugs), and additional complicating conditions such as risk of falls.". We did look at age as a subpopulation of interest however evidence was sparse.
24.	Peer Reviewer #5	General	Some areas of the report that could address these issues include: <ul style="list-style-type: none"> • Page ES-6. As mentioned above, I believe there is a gap in research on the relative risks and benefits of anticoagulation in frail older adults and those who are at increased risk of falls. 	We have added in the population subgroup of frail adults to areas of needed future research
25.	Peer Reviewer #5	General	<ul style="list-style-type: none"> • Page 5-6. To the subgroups of interest for examining these key questions, I would add older adults who are frail, those with comorbidities beyond those related to CV disease, and those at increased fall risk. 	Unfortunately this population is was not identified at the outset as a population of interest and so we have not made the requested change. We have identified this population as a target for future research



#	Commentator & Affiliation	Section	Comment	Response
26.	Peer Reviewer #5	General	<ul style="list-style-type: none"> Page 11, Table 2. Older adults with increased fall risk, multiple comorbidities, frailty, and who are on multiple mediations could be added to the Populations Inclusion criteria. 	This subgroup was not an identified subgroup of interest – although we did include age as a subgroup of interest and evaluate studies that reported findings for older adults more broadly.
27.	Peer Reviewer #5	General	<ul style="list-style-type: none"> Page 172, Table 67. Can either of the elderly subgroups listed in the table be further subdivided into those who are more frail and who have higher levels of comorbidity? 	Because these were not populations of interest predefined for this report we did not separate out studies in this way.
28.	Peer Reviewer #6	General	<p>The report summarizes evidence on the predictive utility of instrument for predicting stroke and bleeding risk in persons with AFib, and comparative effectiveness of a variety of therapies for preventing stroke.</p> <p>Given the scope of the review it is difficult to process all of the results and comparisons. Of course this is not the fault of the reviewers but it makes it hard to digest the results, especially as there are many comparisons with only a few studies.</p> <p>I do think it would have been helpful to focus more on head-to-head comparisons in the risk prediction section.</p> <p>Regardless, I think most of the results are valid though I had some concerns about some specific conclusions.</p>	Comments listed here are addressed in the comment disposition document in specific sections.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
29.	Peer Reviewer #6	General	Specifically, I think it is very difficult to justify the conclusion that the HAS-BLED tool is more accurate than other instruments, given that most studies found no differences.	We agree that several of the bleeding risk tools provide summary c-statistics that are comparable and because of this they each have moderate SOE evidence for limited risk discrimination. From the studies however that evaluated the comparative risk discrimination, the evidence favors HAS-BLED
30.	Peer Reviewer #6	General	I also have some questions about interpretation for the Xa inhibitors. At least at higher (non-renal) doses the effects look quite similar and I'm not sure it makes sense to say that one is effective and the others are not.	We do not feel that our report supports saying that one specific DOAC is more effective than another. We have tried to ensure that our findings reflect the underlying evidence and the heterogeneity in findings. We've also emphasized the need for future studies which have direct comparisons of the DOACs.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
31.	Peer Reviewer #6	General	In some places it is unclear what studies the SOE findings are based on (e.g., an RCT or an RCT + observational studies).	We have tried to clarify throughout the text and within the SOE tables what evidence supports the findings. As stated in our methods “when outcomes were assessed by RCTs and observational studies, we focused our strength of evidence rating on the findings from the RCTs and then increased or decreased the strength of evidence rating depending on whether findings from the observational studies were consistent or inconsistent with those from the RCTs. We provided greatest weight to findings from large RCTs.”
32.	Peer Reviewer #6	General	It is also unclear throughout the results what is new evidence and what was in the prior report, making it difficult to understand what is new/different.	Although this is an update of the 2013 report, this review fully incorporates the older evidence in to the updated evidence review. As such we have not separated out what was known in 2013 versus what is now known – but rather focus on the current state of the evidence given the full evidence base. We do highlight in the “Findings in Relation to What is Known” section of the discussion what our findings have added to existing knowledge in the field

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
33.	Judy Wagner (Public Reviewer #2)	General	<p>ES-3: ASA vs. VKA</p> <p>We kindly suggest that a stronger statement be made around the use of ASA for stroke prevention in atrial fibrillation (SPAF), as there is limited evidence to support its use. While we understand your statement reflects only the findings, we however believe that it misses the mission of this document and a stronger stance should be taken, in light of the overuse of ASA for SPAF.</p> <p>Reference:</p> <ul style="list-style-type: none"> Gregory Y.H. Lip, MD et al. 'Real-World' Antithrombotic Treatment in Atrial Fibrillation: The EORP-AF Pilot Survey. <i>The American Journal of Medicine</i> 2014;127: 519-529 Jonathan C. Hsu, MD, MAS et al. Aspirin Instead of Oral Anticoagulant Prescription in Atrial Fibrillation Patients at Risk for Stroke. <i>J Am Coll Cardiol</i> 2016;67:2913–23. 	We agree and we have inserted in an additional statement in the key point emphasizing with the lack of evidence supporting ASA use in AF patients for stroke prevention.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
34.	Judy Wagner (Public Reviewer #2)	General	<p>ES-5: ROCKET-AF should not be classified as a moderate SOE on the basis of on-treatment vs. ITT (repeated in the tables on page 150 and 182 and 183).</p> <p>Rationale:</p> <p>The safety on-treatment analysis set was selected for the primary superiority analysis prior to the start of the study. The on-treatment data scope was selected in advance as the most appropriate to test superiority for several reasons.</p> <p>First, the 2 study drugs have very different PK and PD properties, such that data scopes of greater duration after discontinuation of study drug would incorporate periods of time where the longer acting warfarin would still be pharmacodynamically active while the rivaroxaban would not be.</p> <p>Secondly, based on the mechanism of action of the study drugs (anticoagulation) and the persistence of the underlying predisposing factor for thrombosis (atrial fibrillation) in the study population, it would not be expected that either agent would have a durable treatment effect beyond the time of their pharmacodynamic activity. Based on the decision to use the on-treatment data scope, it was decided that the most appropriate population for the analysis was the safety population, as a requirement for this population is administration of at least one dose of study drug. The safety on-treatment analysis set was thus considered to provide the most accurate representation of the true treatment effect for superiority testing.</p>	<p>We feel that intention to treat is the best way to analyze data from the RCT and therefore have not modified our approach.</p> <p>Specifically, ITT can be viewed as an approach to estimate realistic treatment use in practice. Randomized groups are balanced with respect to measured and unmeasured patient characteristics. The ITT approach provides valid statistical tests of no treatment difference hypotheses. The “as treated” (treatment received) analysis compares the actually received treatments. However, the compared groups are not anymore fully formed by the randomized assignment but by patient characteristics or other reasons possibly related to outcome, thus likely introducing bias in the “as treated” comparisons. Although “as treated” analyses can be useful, most often they are considered secondary to the ITT approach.</p>

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



35.	Judy Wagner (Public Reviewer #2)	General	KQ 3 – Interventions for Preventing Thromboembolic Events The evidence indicates that the interplay of DOAC dosing and renal function is an important consideration. There are two key analyses: 1. In an FDA analysis, the risk of ischemic stroke with apixaban and edoxaban was increased 1.35 and 1.58, respectively. For edoxaban, this resulted in a black-box warning in the label that precluded the use of edoxaban in patients with CrCL \geq 95 ml/min. (Figure provided below) 2. An independent study by Yao X [Ref: Yao X, et. al. JACC 2017; 69:2779-90] demonstrated a statistically significant ~5x increase in stroke/systemic embolism but no statistically significant difference in bleeding with reduced dose vs standard dose apixaban in patients without a renal indication for reduced dosing; in the same analysis, there was no difference in the benefit : risk of reduced dose vs. standard dose rivaroxaban or dabigatran.	Optimal dosing for specific patient populations was considered out of scope for this review. We do however explore patients with renal impairment as a specific subgroup of interest and summarize available evidence within this population in the report.
-----	--	---------	---	---

#	Commentator & Affiliation	Section	Comment	Response
			<p>Risk of Ischemic Stroke Is Increased for Apixaban and Edoxaban Compared With Warfarin in Patients With Normal Renal Function (CrCl \geq 80 ml/min) – FDA Analysis</p> <p>The chart displays the Hazard Ratio (HR) for the first stroke and Systemic Embolic Events (SEE) for three direct oral anticoagulants (DOACs) compared to Warfarin in patients with normal renal function (CrCl \geq 80 ml/min). The y-axis represents the HR for First stroke/SEE, ranging from 1.0 to 2.0. The x-axis lists the comparisons: XARELTO (Rivaroxaban) vs Warfarin [ROCKET AF] with an HR of 1.07; ELIQUIS (Apixaban) vs Warfarin [ARISTOTLE] with an HR of 1.35; and SAVAYSA (Edoxaban) vs Warfarin [ENGAGE AF] with an HR of 1.58. The FDA logo is present in the bottom left of the chart area. Source: Lawrence J, Hung J. Statistical considerations, ENGAGE AF trial. US FDA website. http://www.fda.gov/downloads/ohrt/ovr/committees/committeesmeetingmaterials/drugs/cardiovascularandrenal/drugadvisorycommittee/ucm421612.pdf. Accessed November 20, 2015.</p>	



#	Commentator & Affiliation	Section	Comment	Response
36.	Judy Wagner (Public Reviewer #2)	General	<p>KQ 3 – Interventions for Preventing Thromboembolic Events</p> <p>Approximately 5 to 8% of patients undergoing PCI have atrial fibrillation. In those patients, the use of anticoagulants is complicated by the antiplatelets that must be used post intervention. In this subpopulation of atrial fibrillation patients, two studies with two different DOACs – dabigatran and rivaroxaban -- have been conducted that demonstrated, in secondary endpoint analyses, similar or non-inferior efficacy in thromboembolic events when dual therapy of a DOAC with a P2Y12 inhibitor is compared to triple therapy (ASA, P2Y12, and warfarin); both studies also demonstrated significantly lower major bleeding compared to triple therapy in the primary endpoint analyses. There are no other available data for other anticoagulants in this subpopulation at this time.</p>	This population was not considered a population of interest for this review and therefore not included for the report.
37.	Judy Wagner (Public Reviewer #2)	General	<p>Page 195 – this statement should be included in the ES</p> <p>“Comparative effectiveness of these direct oral anticoagulants as compared to one another, however is limited by the lack of randomized studies directly comparing their safety and effectiveness.”</p>	We have added this sentence in to the evidence summary.

Table 3 Differences between the four NOAC randomized controlled trials that impact the robustness of cross-study comparisons

	RE-LY ^a	ROCKET AF ^b	ARISTOTLE ^a	ENGAGE AF ^a
Dose reduction	No dose reduction-patients randomized between two doses of dabigatran (110 or 150 mg bid)	At randomization: Rivaroxaban 15 mg od for patients with Cr-Cl 30-49 mL/min	At randomization: Apixaban 2.5 mg bid for patients with ≥ 2 of the following criteria: <ul style="list-style-type: none"> • Age ≥ 80 years • Body weight ≤ 60 kg • Serum creatinine level ≥ 1.5 mg/dL 	Throughout the study period: Edoxaban 30 mg od for patients with ≥ 1 of the following criteria: <ul style="list-style-type: none"> • Cr-Cl 30-50 mL/min • Body weight ≤ 60 kg • Concomitant use of verapamil or quinidine
Patients taking reduced NOAC dose	50.0%	20.7% ¹²	4.7%	25.3%
EOS transition period ^a	No defined EOS transition period Events after the end of the study not reported Selected patients were eligible to remain on dabigatran as part of a long-term, open-label extension study (RELY-ABLE) ¹¹	30-day EOS transition period Increased NOAC event rates during this period likely caused by the slow onset of VKA therapy	30-day EOS transition period (3-day supply of apixaban or placebo) Increased NOAC event rates during this period likely caused by the slow onset of VKA therapy	30-day EOS bridging/transition period (edoxaban half-dose or placebo) No increase in NOAC event rates
Definition of NVAf	Patients with a history of heart valve disorders were excluded	Patients with AF and valvular disease (defined as mitral stenosis or prosthetic valve) were excluded	Patients with moderate or severe mitral stenosis were excluded	Patients with moderate or severe mitral stenosis, unresected atrial myxoma, or a mechanical heart valve were excluded

AF, atrial fibrillation; bid, twice daily; Cr-Cl, creatinine clearance; EOS, end of study; NOAC, non-vitamin K antagonist oral anticoagulant; od, once daily; VKA, vitamin K antagonist.
¹¹The transition period at the end of the follow-up period, from blinded study drug to standard of care.

The differences outlined in the above tables are significant. For example, the ratio between persistent/permanent and paroxysmal AF differed from trial to trial. This is relevant as those patients with persistent AF generally have higher rates of thrombo-embolic events compared to those with paroxysmal AF. Other baseline demographic differences include patients with different risk scores. For example, the mean CHADS2 score for ROCKET AF (3.5) was higher than those for ARISTOTLE (2.1), RE-LY (2.1) and ENGAGE AF (2.8). Again, a study characteristic that has clear consequences for thromboembolic and bleeding outcomes. Additionally, the percentage of patients enrolled with having a prior stroke or transient ischemic attack or the percentage of patients with heart failure differed substantially across the trials, with ROCKET AF enrolling a higher percentage of these patients when compared to the other DOAC trials. It is also important to point out the differences in definitions for stroke, major bleeding



#	Commentator & Affiliation	Section	Comment	Response
			and clinically relevant non-major bleeding (CRNM) across the trials. The differences in these definitions, in addition to the differences in patient populations, make either direct or indirect comparisons across these trials unreliable and misleading. Therefore, greater emphasis should be placed on providing caution around making these types of comparisons.	
40.	Thomas Seck (Public Reviewer #4)	General	<ul style="list-style-type: none"> Page ES-4 and 83 Within the paragraph discussing “Thrombin inhibitors versus warfarin,” we recommend that the reference to “Thrombin inhibitors” be revised to say, “Thrombin inhibitor (Dabigatran).” Inclusion of the dabigatran product name clarifies which thrombin inhibitor is being examined in this section and will bring the formatting of this section in alignment with the way other DOACs (i.e., apixaban and rivaroxaban) are referenced in the paragraphs immediately following.	We have made the suggested clarification throughout.



#	Commentator & Affiliation	Section	Comment	Response
41.	Thomas Seck (Public Reviewer #4)	General	<ul style="list-style-type: none"> Pages ES-4, 83, 84 <p>In instances throughout the report, the authors state that “Dabigatran increases MI risk although this finding did not reach statistical significance (low SOE).” In the sections highlighted within pages ES-4 and 83, an increased MI risk is called out for dabigatran 110mg despite having a similar hazard ratio and p-values for dabigatran 150 mg. BI recommends the authors use the same wording (“no significant difference in MI risk”) to describe the MI risk for the 110mg dose by revising the statement to say: “Dabigatran at a 110mg dose is similar to warfarin for the composite outcome of stroke or systemic embolism (RR 0.89 95% CI 0.73-1.09) (moderate SOE) and is associated with a reduction in major bleeding (RR 0.80 95% CI 0.70-0.93) with no significant difference in all-cause mortality (RR 0.91 95% CI 0.80-1.03), or MI risk.” , , Similarly, on page 84, the increased MI risk is again called out, this time without any reference to statistical significance. We recommend this language also be revised to more accurately describe the lack of statistical significance. These revisions provide a more concise and accurate reflection of the findings while maintaining consistency in wording across similar findings.</p>	We have clarified throughout the report whether there is statistical significance for specific findings and in this specific case have made the suggested modifications.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
42.	Thomas Seck (Public Reviewer #4)	General	<ul style="list-style-type: none"> Page ES-5 and 188 <p>In the sections entitled “Discussion/Findings in Context: What Does This Review Add to What is Already Known?” and “Findings in Relation to What is Already Known,” the authors call out findings for only edoxaban and apixaban, while omitting other DOACs (i.e., dabigatran and rivaroxaban). This selective presentation of product-specific findings creates a perceived imbalance of evidence and information. We recommend the authors revise this section to resolve this imbalance. For example, the report should, at minimum, include subheadings for all DOACs with the authors’ rationale for why the findings are not new or do add to what is already known. Alternatively, the authors may consider reorganizing the section in a way that removes the product-specific headers to alleviate the imbalance.</p>	In the evidence summary and discussion we have now added in sections related to dabigatran and rivaroxaban.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
43.	Thomas Seck (Public Reviewer #4)	General	<p>Page ES-4 and 84</p> <p>BI recommends the authors ensure that the findings under the “Key Points” sections are organized such that primary outcomes for each agent are reported consistently and in the ITT population. Any reporting on secondary endpoints should be similarly consistent across all direct oral anticoagulants (DOAC). For example, the authors highlight that “Edoxaban (either 60mg or 30mg dose) is superior in reducing hemorrhagic stroke (high SOE) and major bleeding (high SOE) though did not differ in overall stroke risk (moderate SOE), myocardial infarction (high SOE) or all-cause mortality (moderate SOE for high dose);” however, if reporting on the primary outcome for edoxaban versus warfarin, the conclusion should be that both doses of edoxaban were non-inferior to warfarin.¹ Along those lines, in regards to the findings around on percutaneous left atrial appendage (LAA) closure versus warfarin, the authors highlight the stroke and all-cause mortality outcomes, but do not report on the other primary endpoints, cardiovascular mortality and systemic embolism.² Furthermore, because all DOACs had superior reduction of hemorrhagic stroke, the authors should either state this for all agents or omit altogether.</p>	Note that the findings which are called out in the key points are those with the strongest SOE and which reflect the outcomes prioritized by the stakeholders.

¹ Giugliano RP, Ruff CT, Braunwald E, Murphy SA, Wiviott SD, Halperin JL, Waldo AL, Ezekowitz MD, Weitz JI, Špinar J, Ruzyllo W, Ruda M, Koretsune Y, Betcher J, Shi M, Grip LT, Patel SP, Patel I, Hanyok JJ, Mercuri M, Antman EM; ENGAGE AF-TIMI 48 Investigators. Edoxaban versus warfarin in patients with atrial fibrillation. *N Engl J Med*. 2013 Nov 28;369(22):2093-104.

² Reddy VY, Sievert H, Halperin J, Doshi SK, Buchbinder M, Neuzil P, Huber K, Whisenant B, Kar S, Swarup V, Gordon N, Holmes D; PROTECT AF Steering Committee and Investigators. Percutaneous left atrial appendage closure vs warfarin for atrial fibrillation: a randomized clinical trial. *JAMA*. 2014 Nov 19;312(19):1988-98. Erratum in: *JAMA*. 2015 Mar 10;313(10):1061.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



44.	Thomas Seck (Public Reviewer #4)	General	<p>Page ES-4, 84, 189</p> <p>BI recommends that the authors present findings from observational studies in a consistent manner across all DOACs. On pages ES-4 and 84, the authors claim to highlight inconsistencies between findings in the observational studies and RCTs for dabigatran when, in fact, the findings from the observational studies support those of the RCTs. For example, the authors state that “observational studies demonstrated a benefit [in all-cause mortality] for patients on dabigatran, while RCT studies suggested no difference;” however, the RCT did report a reduction in all-cause mortality with dabigatran vs warfarin that was just under the threshold for statistical significance ($p=0.051$).^{3,4} On the other hand, the authors also reported that while “observational studies did not show a difference [in MI risk], RCT studies suggested an increase with dabigatran;” however, in this instance, they failed to note that this increase did not reach statistical significance.⁵ In both cases, the findings from the observational studies in fact supported findings from the RCTs. As such, these findings should not be interpreted as an inconsistency between the studies, rather they should be presented with the appropriate context and same objectivity as the findings for other products. Moreover, in the context of comparing findings from RCTs and observational studies, the authors did not report that both found statistically significant reduction in cardiovascular death. We therefore recommend that the authors revise this section to more accurately reflect the findings from the observational studies and RCTs for dabigatran versus warfarin.</p>	<p>This inconsistency has been fixed and we now consistently note whether statistical significance was reached.</p>
-----	--	---------	---	---



#	Commentator & Affiliation	Section	Comment	Response
45.	Thomas Seck (Public Reviewer #4)	General	<ul style="list-style-type: none"> Page ES-6 <p>The report notes that inconsistent effect estimates from observational studies “likely resulted from confounding, selection bias, different endpoint definitions and rigor and completeness of follow-up.” While these are valid considerations, inconsistencies could also have resulted from variations in decision-making in practice between trials and real world scenarios. We recommend the authors acknowledge this as an additional consideration.</p>	We thank the reviewer for their comment and have made the suggested addition.
46.	Thomas Seck (Public Reviewer #4)	General	<ul style="list-style-type: none"> Page ES-5 <p>The report states that “both lower (30 mg) and higher (60 mg) once-daily doses of edoxaban is similar to warfarin in preventing stroke or systemic embolism...;” however, for optimal clarity, AHRQ should note that the 60 mg once-daily dose of edoxaban is approved by the FDA to treat only NVAf patients with creatinine clearance (CrCL) >50 to ≤ 95 mL/min, while 30 mg once-daily dose of edoxaban is approved to treat NVAf in patients with renal dysfunction (CrCL 15 to 50 mL/min).</p>	We have made the suggested modification.

³ 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 Sep 17;361(12):1139-51. Epub 2009 Aug 30.

⁴ 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 Nov 4;363(19):1875-6.

⁵ Ibid.



47.	Thomas Seck (Public Reviewer #4)	General	<p>BI recognizes that OACs are the focus of this systematic review, however, we feel the review has not been sufficiently updated to reflect the existence of newly available reversal agents. , In particular, PRAXBIND® (idarucizumab), a specific reversal agent to PRADAXA® (dabigatran) indicated to reverse the anticoagulant effects of PRADAXA® in case of emergency surgery, urgent procedures or in case of life-threatening or uncontrolled bleeding, was never explicitly mentioned in the report. PRAXBIND® was granted Breakthrough Therapy Designation in June 2014 and Priority Review in April 2015. The Biologics License Application received accelerated approval, and FDA approval was granted on October 16, 2015. It is currently the only FDA approved therapy to specifically reverse the anticoagulant effect of dabigatran. Despite PRAXBIND®'s recent approval, it is only indirectly acknowledged on page 194, when the authors note that “despite all the potential advantages of the direct oral anticoagulants demonstrated in the clinical trials when compared with warfarin, except for dabigatran, these drugs still do not have an approved immediate antidote.”</p> <p>As highlighted in Palacio et al. (2015), when selecting anticoagulation treatments, awareness of the availability of an antidote can help alleviate patient concerns around elevated bleeding risk and even influence decision-making. Therefore, highlighting pertinent information about the role, availability, indication, and effectiveness of FDA-approved reversal agents within this systematic review can ensure the reader has a comprehensive understanding of anticoagulation management in patients with non-valvular atrial fibrillation.</p>	<p>Thank you for the information</p> <p>Reversal agents were not included as interventions of interest and therefore were considered out of scope for the current review. No changes made.</p>
-----	-------------------------------------	---------	--	--

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
48.	Thomas Seck (Public Reviewer #4)	General	<p>Opportunities to strengthen the report through acknowledgement and/or more robust discussion of reversal agents include:</p> <ul style="list-style-type: none"> • Page 9 <p>Similarly, availability of a reversal agent may be incorporated under Key Question 2 (“Clinical tools and individual risk factors for assessment/evaluation of intracranial hemorrhage bleeding risk”). We understand that because this review was intended to update an existing review, AHRQ did not provide an opportunity to comment on the KQs, however, we recommend that future updates more extensively consider evidence around the role and impact of reversal agents.</p>	Although reversal agents are clinically important they are not within the scope of this systematic review and therefore no changes have been made.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
49.	Thomas Seck (Public Reviewer #4)	General	<p>BI has notes that the following relevant articles were not included in this review. Given their relevance to findings throughout the report, we recommend the authors add findings from these studies to the report wherever relevant:</p> <ul style="list-style-type: none"> • Connolly SJ, Wallentin L, Yusuf S. Additional events in the RE-LY trial. N Engl J Med. 2014 Oct 9;371(15):1464-5. Epub 2014 Sep 24. • Flaker G, Ezekowitz M, Yusuf S, Wallentin L, Noack H, Brueckmann M, Reilly P, Hohnloser SH, Connolly S. Efficacy and safety of dabigatran compared to warfarin in patients with paroxysmal, persistent, and permanent atrial fibrillation: results from the RE-LY (Randomized Evaluation of Long-Term Anticoagulation Therapy) study. J Am Coll Cardiol. 2012 Feb 28;59(9):854-5. • Gibson CM, Mehran R, Bode C, Halperin J, Verheugt F, Wildgoose P, van Eickels M, Lip GY, Cohen M, Husted S, Peterson E, Fox K. An open-label, randomized, controlled, multicenter study exploring two treatment strategies of rivaroxaban and a dose-adjusted oral vitamin K antagonist treatment strategy in subjects with atrial fibrillation who undergo percutaneous coronary intervention (PIONEER AF-PCI). Am Heart J. 2015 Apr;169(4):472-8.e5. Epub 2014 Dec 20. • Hernandez I, Zhang Y, Brooks MM, Chin PK, Saba S. Anticoagulation Use and Clinical Outcomes After Major Bleeding on Dabigatran or Warfarin in Atrial Fibrillation. Stroke. 2017 Jan;48(1):159-166. Epub 2016 Dec 1. 	<p>Thank you for these additional citations.</p> <p>The citations from Connolly and Flaker were both research letter correspondence and therefore were not included.</p> <p>The citation by Gibson and colleagues (was excluded at the full text level since it is a description of the study design for the PIONEER AF-PCI trial but does not include findings.</p> <p>The study by Hernandez and colleagues was excluded at the full text because it was unclear whether the study population was restricted to those with nonvalvular AF.</p>

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
50.	Thomas Seck (Public Reviewer #4)	General	<p>The articles below were published after the initial June 2017 cut-off date, however, we urge the authors to revise these parameters and include the following studies to ensure comprehensiveness and completeness of findings:</p> <ul style="list-style-type: none"> • Adeboyeje G, Sylwestrzak G, Barron JJ, White J, Rosenberg A, Abarca J, Crawford G, Redberg R. Major Bleeding Risk During Anticoagulation with Warfarin, Dabigatran, Apixaban, or Rivaroxaban in Patients with Nonvalvular Atrial Fibrillation. <i>J Manag Care Spec Pharm.</i> 2017 Sep;23(9):968-978. • Cannon CP, Bhatt DL, Oldgren J, Lip GYH, Ellis SG, Kimura T, Maeng M, Merkely B, Zeymer U, Gropper S, Nordaby M, Kleine E, Harper R, Manassie J, Januzzi JL, Ten Berg JM, Steg PG, Hohnloser SH; RE-DUAL PCI Steering Committee and Investigators. Dual Antithrombotic Therapy with Dabigatran after PCI in Atrial Fibrillation. <i>N Engl J Med.</i> 2017 Oct 19;377(16):1513-1524. Epub 2017 Aug 27 • Løfgren B, Pareek M, Larsen JM. Uninterrupted Dabigatran versus Warfarin for Ablation in Atrial Fibrillation. <i>N Engl J Med.</i> 2017 Aug 3;377(5):494-5. 	<p>Thank you for these citations. The study by Adeboyeje was part of our search update and is now included in the systematic review. The study by Cannon focused on patients with PCI and therefore was excluded as not a population of interest. Similarly the study by Lofgren focused on patients undergoing ablation and therefore was also excluded for not a population of interest.</p>
51.	Peer Reviewer #1	Introduction	<p>clear concise review-helpful to outline results of last review as you did. objectives for this review clearly stated.</p>	<p>Thank you. No response needed</p>



#	Commentator & Affiliation	Section	Comment	Response
52.	Peer Reviewer #2	Introduction	In the discussion of “Risk Stratification” on page 2, I recommend a more detailed discussion of how the trade-offs should actually be assessed and decisions made. The summary on lines 34-36 does not seem quite right, because it will usually not be possible to have both maximum benefit and the lowest risk of complications. I believe that the optimization to be attempted, and the role of risk stratification in it, is more complex than this, with the goal being to optimize the expected net benefit of treatment (or lack thereof). A good choice could be achieved, for example, by very high expected benefit despite a substantial increase in risk of complications. The discussion on page 185 of the tool developed by Fraenkel (reference 363) suggests to me that the framework developed there may be a good way to explain the importance and roles of the different quantities assessed in the three Key Questions. At a minimum, I would hope that the Introduction would somewhere make clear that the key quantities needed for rational treatment choice are the differences in absolute risks with one treatment strategy versus with another.	We have added some additional detail about the complexity of balancing potential harms and benefits when risk stratifying patients with atrial fibrillation.
53.	Peer Reviewer #2	Introduction	For the contextual question, what is meant by a “best evidence” approach?	We now describe how a “best evidence” approach prioritizes evidence based on study design, reporting, and relevance
54.	Peer Reviewer #3	Introduction	no comments	Thank you. No response needed
55.	Peer Reviewer #4	Introduction	Reads well, sets up paper nicely	Thank you. No response needed

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
56.	Peer Reviewer #5	Introduction	No comments	Thank you. No response needed
57.	Peer Reviewer #6	Introduction	The section on use of anticoagulation therapy focuses on monitoring of PT/INR which is mainly relevant for warfarin, but doesn't address lab monitoring (assays and need for monitoring or not) for other meds.	Given the size of the report we have focused the introduction on the most important clinical evidence and uncertainties and as such not all components of stroke prevention strategies are included in detail.
58.	Peer Reviewer #6	Introduction	There is no discussion of non-drug options for stroke prevention, other than in the table.	We have added additional discussion about the non-drug options immediately preceding the table.
59.	Peer Reviewer #6	Introduction	P 26—the paragraph on how the report will be useful to various stakeholders reads a little self-congratulatory and I would consider deleting it.	We have removed this paragraph.
60.	Peer Reviewer #1	Methods	very clear	Thank you. No response needed

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
61.	Peer Reviewer #2	Methods	<p>I have concerns about the criterion stated at page 16 lines 17-21: “we assumed that a HR between 0.8 and 1.2 with a narrow confidence interval that also crossed 1.0 suggested no clinically significant difference between treatment strategies; in such cases, we describe the treatment strategies being compared as having ‘comparable efficacy.’” This criterion seems inappropriate for whether an estimated HR is clinically important. The width of the CI and whether it crosses 1.0 are relevant for strength of evidence, rather than for assessing the clinical implications of the point estimate. Whether the CI crosses 1.0 is equivalent to assessing statistical significance at $p < 0.05$, but the recent statement on p-values from the American Statistical Association states as its Principle #5: “A p-value, or statistical significance, does not measure the size of an effect or the importance of a result.” For example, I would see no reason to interpret HR=0.90 with CI 0.80 to 1.01 as evidence for comparable efficacy, while HR=0.92 with CI 0.85 to 0.99 is evidence for superiority. If HR=0.92 is far enough from 1.0 to be clinically important, then so is HR=0.90. An additional problem is that this criterion is often not interpreted as described here: the wording “comparable efficacy” has been forgone in favor of “no difference” in many places throughout the report. In addition, HRs outside the stated range have been interpreted as “no difference” (e.g., P 43 Li 49-51, 7% vs 4%; P 43 Li 46-49, entry in Table 16 with HR 0.78). My impression is that $p > 0.05$ has been interpreted as evidence of no difference, but this is not valid statistical reasoning.</p>	<p>We have clarified throughout that rather than saying “no difference” we instead say “no evidence of a difference”. The reviewer is correct that we do not focus on the minimally important difference from a clinical perspective.</p>

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
62.	Peer Reviewer #3	Methods	Yes for all questions except I did not see criteria pertaining to study type and quality. See comment on this in my general comments.	Thank you – we have responded to this concern in the general comments section
63.	Peer Reviewer #4	Methods	Are the definitions or diagnostic criteria for the outcome measures appropriate? A table with a comparison of the outcome definitions in each of the clinical trials and corresponding observational studies would be useful. The studied are compared to each other, but in many cases there are differences in the outcome definitions (+/- traumatic head bleeds, +/- inclusion of TIA in the stroke outcome, definitions of major bleeding).	We agree that the outcomes assessed are often heterogeneous. We have attempted to clarify throughout the specific outcomes assessed in the individual studies.
64.	Peer Reviewer #5	Methods	No comments	Thank you. No response needed
65.	Peer Reviewer #6	Methods	The analytic framework is structured and uses shapes/symbols differently than what I'm used to seeing (e.g. USPSTF analytic frameworks, which often incorporate diagnostic accuracy and risk assessment elements). I believe it is based on the AF in the prior report and I don't feel that there is a strong need to re-vamp it, but might be good to be clear that the prior AF was used if that was the case. It is not clear to me what is the difference between therapeutic efficacy and patient outcome efficacy for KQ 1 and KQ 2 and how these are different than the outcomes in the box in the far right of the AF.	The reviewer is correct that the analytic framework is based on the prior report and updated to reflect the additional outcomes of interest.
66.	Peer Reviewer #6	Methods	The ES has no Methods and goes right to Results—as the ES is often a standalone document I think it would be a good idea to have a brief Methods section.	The structure of the evidence summary follows the template for PCORI systematic reviews.



#	Commentator & Affiliation	Section	Comment	Response
67.	Peer Reviewer #6	Methods	Is the “Organization of this Report” section necessary—won’t there be a table of contents?	We thank the reviewer for this comment although have maintained the “organization of this report” section based on comments from AHRQ/PCORI
68.	Peer Reviewer #6	Methods	There’s a bit of repetition and lengthy explanation in the Search Strategy section for the search start dates—I think it can be stated succinctly that studies in the prior report were included, what the search dates for the prior report were, and that you conducted new searches starting from 2011 to capture new studies.	We have streamlined the explanation as suggested.
69.	Peer Reviewer #6	Methods	A sample size threshold of >1000 for drug studies seems pretty high, and I’m not sure that requiring a study to be big is the most important criterion. E.g. it might be better to have used a smaller sample size threshold but restrict to studies that performed statistical adjustment or had adequate adjudication of VTE and bleeding events. Anyway I don’t think you necessarily need to change the inclusion criteria at this point, but this should be discussed in the Limitations section of the Discussion.	We explored several mechanisms for maintaining the scope of the review within assigned resources – and as noted considered excluding studies of lower methodological quality. The final inclusion/exclusion criteria tried to balance the limitations of any exclusions with available resources. We have added some additional clarification to the limitations section.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
70.	Peer Reviewer #6	Methods	Is management of afib really that different in places like Japan, Taiwan, Korea and Israel compared with the U.S.? I would think places like that would be more similar than some low-income countries in South and Central America, for example.	We explored several mechanisms for maintaining the scope of the review within assigned resources – and as noted considered excluding studies of lower methodological quality. The final inclusion/exclusion criteria tried to balance the limitations of any exclusions with available resources.
71.	Peer Reviewer #6	Methods	Many of the studies evaluated prognosis/predictive utility, not just diagnostic accuracy—there are some tools available to assess prognostic studies (e.g., Jill Hayden’s tool published in Annals, recent GRADE paper on applying GRADE to prognosis studies)—was there a reason these were not used? The prognostic studies of course have the time element but may also have statistical issues (e.g., when reporting risk estimates, need for adjustment). There is also the issue of evaluating risk instruments in the derivation cohort vs. an independent validation cohort, which is mentioned later in the Methods but relevant for assessing risk of bias/quality.	Although we did not evaluate prognostic/predictive utility, in addition to diagnostic accuracy, we also attempted to look for evidence exploring diagnostic thinking efficacy, therapeutic efficacy, and patient outcome efficacy. These outcomes were indicated to be highest interest by PCORI, AHRQ, and the stakeholders involved in the topic refinement process. Unfortunately the evidence regarding outcomes other than diagnostic accuracy was negligible.



#	Commentator & Affiliation	Section	Comment	Response
72.	Peer Reviewer #6	Methods	The report focuses on discrimination as the measure of predictive utility but there are many other ways to evaluate predictive utility—e.g. sensitivity/specificity/LR's at various cutoffs, calibration, net reclassification index, adjusted risk estimates, etc. All of these can provide some useful information. In fact I would suggest that for risk prediction of CVA that calibration is more important than discrimination—the goal is not necessarily to discriminate every person that has a CVA from those that don't. But if an instrument is well-calibrated and can clearly identify people who are at higher risk for CVA with higher scores (which CHADS2 seems to do pretty well) that is extremely useful clinically and would be a strong reason to use it in clinical practice. There should be some explanation of why the report focuses solely on discrimination and this should be noted as a Limitation.	The specific measures of the predictive accuracy of the various tools were heterogeneous and sparse. The report focused on those outcomes most consistently reported within the included studies.
73.	Peer Reviewer #6	Methods	The Methods describes how the c-statistic was interpreted but not interpretation of magnitude of effects for efficacy/harms outcomes i.e. what would be considered a clinically relevant HR/RR for major CVA or bleed?	We have clarified within the methods that we do not make recommendations on whether specific differences are clinically relevant but instead highlight whether there is evidence of a statistical difference for outcomes of interest.



#	Commentator & Affiliation	Section	Comment	Response
74.	Peer Reviewer #6	Methods	p. 43—I am not sure that having a large number of studies means there is no need to worry about publication/reporting bias. There are plenty of examples of drugs for which there was a lot of published studies (e.g., antidepressants in kids) for which there was evidence of publication/reporting bias. Also there isn't really a good way identify unpublished observational studies. Anyway I would suggest just saying that you found no indication of a large number of unpublished studies and just leave it at that.	We have revised the discussion related to publication bias and whether the large number of studies reduces the potential for such bias.
75.	Peer Reviewer #6	Methods	I don't think that being a large single study makes a body of evidence "consistent". It may make it more (or much more) precise, but that is not the same thing. You cannot rate consistency if there is only one study, and I would argue that in general the body of evidence should be downgraded for consistency when there is only a single study regardless of the sample size. Throughout the Results there are a number of SOE ratings of moderate and especially high for single studies that I would suggest reviewing.	We have corrected any SOE ratings where outcomes with only one study were rated as consistent to not be not applicable.
76.	Peer Reviewer #1	Results	outstanding and comprehensive. I don't see many readers ingesting this however. I guess it depends on your audience.	Thank you. No response needed



#	Commentator & Affiliation	Section	Comment	Response
77.	Peer Reviewer #2	Results	Many results on event rates, notably in Tables 6-8, could be converted to a common scale, such as events per 100 person-years, and then plotted in forest plots, separately for each score (or score range/category). The plots could also use different colors or symbols for different event definitions. I believe that such plots would more clearly show the risk scores' utility, or lack thereof. For example, if the absolute rates of events at a given score vary widely from study to study, including studies that nominally have the same event definition, then using the risk score in a clinical decision becomes problematic, because estimating the absolute risk with or without any given treatment is too uncertain. If differing population characteristics are what cause the rates to vary widely, then the risk scores need to be revised to incorporate those characteristics. On the other hand, a quantitative synthesis of estimated of event rates could be quite useful if such a synthesis would be valid.	We have modified these tables in KQ1 and KQ2 to use consistent outcome measures when possible and then structured the tables to organize the rates by scores. Although the differing population characteristics are such that combining these event rates is not valid we feel that these revised tables now allow the reader to better compare findings from the numerous studies.
78.	Peer Reviewer #2	Results	The word "continuous" is used in several places for scores that only take on a range of integer values (P 26 li 10, Table 21, P 60 Li 31). This could be misunderstood as meaning that a linear regression was used (a common meaning of treating a predictor as "continuous"). I believe that the meaning is instead that the risk score was used in full detail, rather than some distinct scores being lumped together.	We have clarified that continuous scores mean that there are used in full detail rather than categorical when distinct scores are lumped together.
79.	Peer Reviewer #2	Results	At page 46 lines 34-36, the numbers appear to be inconsistent. The HRs are very similar, but $p=0.03$ for interaction.	We have clarified the listed p value



#	Commentator & Affiliation	Section	Comment	Response
80.	Peer Reviewer #2	Results	At page 47 lines 10-13, the two groups/conditions being compared do not seem clear for all of the HRs.	We have clarified the groups being compared.
81.	Peer Reviewer #2	Results	At page 48 line 28, it does not seem clear what two factors are interacting.	We have clarified what is being compared.
82.	Peer Reviewer #2	Results	At page 49 line 52, I recommend avoiding the wording “not being statistically different”. This can give the impression that statistical analysis has shown that the apparent difference is not real, which is not a correct interpretation of $p>0.05$.	We have revised the text to clarify.
83.	Peer Reviewer #2	Results	Why does the format change for Figure 5, vs Figs 3,4,6,7?	Given that there were no new studies evaluating the Framingham risk scores, Figure 5 was repeated from the original systematic review report unchanged.
84.	Peer Reviewer #2	Results	P 60 Li 37-45. I believe that similar discussion/orientation would also be useful for Key Question 1. Also, at lines 42-43, what is the “data” that is not directly compared? If rates cannot be compared and estimated, then what is the clinical utility? This would seem to assume a priori that the scores must be clinically useless, because they cannot be used to estimate actual risk.	We have clarified in KQ1 the limitations of our quantitative analyses.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
85.	Peer Reviewer #2	Results	P 62 Li 17-19. P 63 Li 28-30. P 65 Li 33-35. P 68 Li 52-54. I believe that this phenomenon should be clearly illustrated and interpreted as a severe limitation of the risk scores. For a risk score to be helpful for clinical decisions, it must give information about the magnitude of risk. It is not enough to know only that higher scores have higher risk. As noted above, if absolute rates differ substantially between different populations, then the populations' characteristics that cause those differences should be incorporated in to the risk scores. The information about event rates in Tables 24-27 could be shown clearly by use of forest plots as noted for Key Question 1.	We agree that there are limitations both in the risk scores and in the evidence supporting these scores. Similar to KQ1, we now try to structure the tables in KQ2 to allow better comparison of the scores across studies.
86.	Peer Reviewer #2	Results	P71 Li 42-48. P 72 Li 19-22. The information given does not seem relevant to the topic being addressed, because it is not evaluating CKD's influence on bleeding risk. The influence of CKD on the effect of warfarin is a different issue, and would probably belong under Key Question 3 rather than here. There is a similar problem at page 73 line 5.	We have removed this study from the individual risk factors table
87.	Peer Reviewer #2	Results	P 72 Li 40. Two numbers are apparently given for one HR.	This typo has been fixed.
88.	Peer Reviewer #2	Results	P 80 Li 42. If the study only included patients presenting with stroke, how were incidence rates determined? Was this using ischemic stroke presentation as a "control" group for those presenting with ICH? That seems odd.	This error has been corrected.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
89.	Peer Reviewer #2	Results	P 86 Li 24-25. Why is no quantitative estimate of the synthesized HR given in this sentence? Most or all of the mentioned results can be converted to HRs and synthesized. I have the same question for P 86 Li 43-44.	Given the heterogeneity in the populations we did not combine this evidence quantitatively.
90.	Peer Reviewer #2	Results	Table 34 has “Magnitude of effect” in a column heading, but does not include this information in the table entries.	We have modified the column text to clarify that it includes the SOE and effect.
91.	Peer Reviewer #2	Results	In the forest plots, Figures 8-35, “favors treatment” and “favors control” do not seem like clear labels, since both are active treatments.	We have clarified in the headers of all the figures what treatment and control correspond with in terms of interventions
92.	Peer Reviewer #2	Results	P 98-99. Why is there a change in format for presenting synthesized evidence between subsections 1-8 and subsection 9?	The evidence in subsection #9 onwards represent the majority of new studies and data and therefore those subsections are formatted for presentation to aid the reader in reviewing the entirety of the large evidence base.
93.	Peer Reviewer #2	Results	P 99 Li 44. Characterizing an HR of 0.80 as not “significant” does not seem reasonable. The intended meaning appears to be not “statistically significant”, which is a different concept.	We have globally clarified text to read “no statistically significant...”
94.	Peer Reviewer #2	Results	P 167, Li 45 or 46. The interpretation “Low strength of evidence for no benefit in all-cause mortality” seems to be inconsistent with the HR’s and p-values given in the previous sentence.	We have clarified that this SOE rating was based on the entirety of the included studies findings.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
95.	Peer Reviewer #2	Results	P 167 Li 51. What is a “cumulative primary safety rate”? If this is really looking at safety, then it seems very low (90% unsafe).	We now clarify that the cumulative primary safety rate included events related to excessive bleeding (eg, intracranial or gastrointestinal bleeding) or procedure-related complications (eg, serious pericardial effusion, device embolization, procedure-related stroke).
96.	Peer Reviewer #2	Results	P 168 Li 40. Missing which day of followup: “At - day followup”.	We have corrected this typo
97.	Peer Reviewer #2	Results	Similar to Table 34, Table 66 has “Magnitude of effect” in a column heading, but does not include this information in the table entries. Some entries give two separate rates, but I would expect that separate event rates could be compared to obtain a HR. Other entries only indicate the direction of effect, not the magnitude.	We have clarified the column headings to reflect the content of the tables.
98.	Peer Reviewer #2	Results	P 171 Li 13-16. Large interaction p-values do not provide evidence for similarity.	The interaction information has been removed.
99.	Peer Reviewer #2	Results	Table 70. A few entries (near the top and at the end) do not give quantitative effects and CIs.	We have provided quantitative effects where the data supports such synthesis.
100.	Peer Reviewer #3	Results	The amount of data presented is massive and hard to follow. In some places there was inconsistency in the way the results were described. There were times the practices recommended below were adhered to and other times they were not. I would recommend adhering to them consistently.	We have attempted to clarify the text and tables as well as the consistency of the presentation.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
101.	Peer Reviewer #3	Results	All types of observational studies were often combined into a single category and described as “observational”. This is perhaps acceptable when describing a large number of studies, such as 10 observational studies, but it is better to quantify the number of each type, for instance, 5 cohort and 5 case-control. (See page 83 first key point as an example where this could have been done)	Given the number of studies we have not included this breakdown in the main report although it is included in Appendix F. Of note there was only one case-control study and one cross-sectional study. All other studies were cohorts.
102.	Peer Reviewer #3	Results	When describing individual studies and their results the study type, including specific observational type, and quality should be mentioned. (Such as was done on page 93, lines 32-33)	We have included information on study quality and design when appropriate but given the large number of studies don't include this information every time a study is discussed. We do include this information in the Appendix and in the summary key points
103.	Peer Reviewer #3	Results	In addition to hazard ratios and RR's, some description of absolute differences should be included (number needed to treat to achieve a benefit or number needed to harm to achieve an adverse outcome). This information would be useful in the tables as well as the text but at least in one of the two. Table 51 is a good example as is table 66. This would be particularly helpful in the summary tables.	We have added in the NNT for the stroke outcome in Table 54 for ARISTOTLE where a reduction was demonstrated.
104.	Peer Reviewer #3	Results	When there are inconsistent findings in the studies, the highest quality studies should be given some preference in the interpretation. (For example ES-6 lines 13-18, and page 84 lines 3-7)	When there are inconsistent findings in the studies (and between study designs) we have attempted to be clear about how we have used this information in modifying our associated SOE rating.



#	Commentator & Affiliation	Section	Comment	Response
105.	Peer Reviewer #4	Results	Is the amount of detail presented in the results section appropriate? YES, its a ton of information, but the key point bullets make it manageable and allow the reader to find the information they are looking for	Thank you. No response needed
106.	Peer Reviewer #4	Results	Could include this paper in the section "Predicting Thromboembolic Risk" Comparison of the CHA2DS2-VASc, CHADS2, HAS-BLED, ORBIT, and ATRIA Risk Scores in Predicting Non-Vitamin K Antagonist Oral Anticoagulants-Associated Bleeding in Patients With Atrial Fibrillation. Yao X, Gersh BJ, Sangaralingham LR, Kent DM, Shah ND, Abraham NS, Noseworthy PA. Am J Cardiol. 2017 Nov 1;120(9):1549-1556. doi:10.1016/j.amjcard.2017.07.051. Epub 2017 Jul 31.PMID: 28844514	Thank you for the citation, the suggested citation is included in the revised report.
107.	Peer Reviewer #5	Results	See general comments (#s 24-36)	No response needed. Comments addressed previously
108.	Peer Reviewer #6	Results	Abstract: The results describe results as providing "limited prediction" which sounds a little awkward and nonspecific to me. Consider something like "...CHADS2 scores provides moderate (or fair, or whatever descriptor is appropriate) discrimination for identification of future stroke events..." or something like that.	We have attempted to balance readability with ensuring that we are accurately reflecting what the results are demonstrating and therefore have retained our use of the wording "limited prediction ability"
109.	Peer Reviewer #6	Results	ES: The comparative accuracy bullet seems to focus on the amount of evidence for the different instruments rather than how the instruments perform relative to one another, which I think is of more interest clinically.	We have modified the key messages bullet to make clear that the specific tools have the best evidence to support risk prediction.



#	Commentator & Affiliation	Section	Comment	Response
110.	Peer Reviewer #6	Results	ES: On tools for predicting bleeding risk, the bullet on CKD seems to just describe the association between CKD and bleeding, not the performance of instruments in persons with CKD on anticoag.	We have modified the key point to clarify this point.
111.	Peer Reviewer #6	Results	ES: For dabigatran vs. warfarin it is unclear if the first two subbullets are based on the RCT. The third subbullet also refers to “RCTs” (plural) when the main bullet says there was one RCT. Also is the SOE based purely on the single RCT? How was inconsistency between the RCT and the observational studies accounted for?	We provide additional details in the text about the SOE ratings and how these are based on RCT and observational studies (and potential inconsistencies). We do not provide such details in the evidence summary.
112.	Peer Reviewer #6	Results	ES: For apixaban vs. warfarin and rivaroxaban vs. warfarin it is unclear if the HR's is based on some kind of summary estimate or if it is from the RCT.	We provide additional details in the text about the SOE ratings and how these are based on RCT and observational studies (and potential inconsistencies). We do not provide such details in the evidence summary.
113.	Peer Reviewer #6	Results	KQ 1—Should have some discussion of methodological issues in the studies rated fair or poor quality,	We have no inserted text in the introduction to each KQ about the major limitations related to quality/risk of bias.
114.	Peer Reviewer #6	Results	The results and tables for individual risk prediction instruments focus on stroke rates at different scores but don't seem to report the c-statistic which is described as the main outcome of interest for predictive utility.	We present the c-statistics in Table 21 and use those findings to determine our SOE ratings. We do however present the event rate findings within Tables 6 through 16 based on feedback from our clinical stakeholders.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
115.	Peer Reviewer #6	Results	The section on imaging findings, INR etc report other measures of predictive utility (adjusted risk estimates, etc)—there seems to be inconsistency in how predictive utility is assessed from instrument to instrument.	We agree that the evidence related to imaging findings is sparse and heterogeneous. As such, we have rated the evidence as insufficient.
116.	Peer Reviewer #6	Results	Presence of CKD in itself isn't really a risk prediction instrument—though adding CKD to an existing instrument makes sense to assess as a tool. Again this section is scattered with regard to what predictive utility outcomes are reported.	We have clarified the limitations of the evidence exploring CKD and bleeding risk in both the key point and within the text.
117.	Peer Reviewer #6	Results	P 69—Would be better to report the range in differences between risk instruments for c-statistics across studies rather than describing the range of c-statistics across all of the studies and instruments. i.e. it is more informative to note that the difference in c-statistic ranged from 0.00 to 0.10 than to say that the c-statistics ranged from 0.56-0.78 or whatever the numbers are. Could focus on comparisons against CHADS2 as that is likely to be the most common comparator.	We feel that the findings currently represent the best synopsis of the comparative data. We have focused both on the c-statistics as well as whether their 95% confidence intervals overlap. As discussed there is sparse data evaluating direct comparisons.
118.	Peer Reviewer #6	Results	The SOE table focuses on pooled c-statistic for each instrument rather than pooled differences in c-statistic between instruments, which I think is a more relevant comparison. In effect you are comparing the predictive utility of instruments based on indirect comparisons (even though you say in the Discussion that indirect comparisons are not likely to be helpful).	We feel that the findings currently represent the best synopsis of the comparative data. We have focused both on the c-statistics as well as whether their 95% confidence intervals overlap. As discussed there is sparse data evaluating direct comparisons.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
119.	Peer Reviewer #6	Results	In this section and other sections of the Results, it would be helpful to summarize what the prior report found and then describing how new evidence impacted findings.	Although this is an update of the 2013 report, this review fully incorporates the older evidence in to the updated evidence review. As such we have not separated out what was known in 2013 versus what is now known – but rather focus on the current state of the evidence given the full evidence base. We do however have a section in the discussion which highlights what this updated report adds to what is already known.
120.	Peer Reviewer #6	Results	KQ 2. The results aren't really described in a parallel fashion to KQ 1, e.g. the summary bullets don't describe the c-statistics for the various instruments.	Because the evidence for KQ2 did not support quantitative synthesis of the c-statistics, we did not include these findings in the key points.
121.	Peer Reviewer #6	Results	The table in the bleeding risk section report the c-statistic (unlike the tables for KQ 1).	Given the greater number of studies in KQ1, we felt that separating out the event rate findings and the c-statistics was easier for the reader. This was not the case in KQ2.
122.	Peer Reviewer #6	Results	Most studies seemed to show no clear difference between HAS-BLED and the other bleeding prediction instruments, so it is quite difficult to justify the conclusion that HAS-BLED is superior.	Based on the studies where direct comparisons of the bleeding risk scores were evaluated, HAS-BLED does appear to have the strongest evidence supporting its comparative predictive abilities.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
123.	Peer Reviewer #6	Results	Again when reporting results would focus on the head-to-head comparisons and describing the difference in AUROC's . In most cases the CI's are wide and overlapping with no clear difference. ?Why was there no attempt to pool	We do not directly compare data from different studies, as we felt that it would not be appropriate given inter-study differences in patient populations, followup times, and definitions of bleeding outcomes.
124.	Peer Reviewer #6	Results	In these Tables would be helpful to have some description of what intervention the patient is on.	We have separated the tables out as to whether the patients are on therapy or not, the specific therapies are available in the study characteristics tables in the Appendix
125.	Peer Reviewer #6	Results	KQ 3. Unfortunately given the large number of comparisons and outcomes, and results based on a single study, this section is very hard to digest.	We now include a summary table (Table 35) of the different comparisons and hope that this table aids the reader in digesting the large amount of evidence
126.	Peer Reviewer #6	Results	It is unclear to me if there was only 1 RCT of ASA vs. warfarin total, or just one new RCT (likewise for other comparisons—it sounds like the Results focus on new studies but it is not clear).	Throughout the results report the findings from 2000 to 2018 inclusive and we do not indicate studies as being from the original report versus the update. The findings as reported are for the complete current evidence base.
127.	Peer Reviewer #6	Results	It would be helpful to have some description of the patient populations in the studies in terms of CVA risk (e.g., average CHADS scores, bleeding risk scores) as well as some other characteristics (mean age, etc.)	This information is provided in the study characteristics appendix.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
128.	Peer Reviewer #6	Results	Would be helpful to report absolute rates of major CVA and major bleeds to help interpret the relative risk estimates.	Given the amount of evidence we have focused our report on the relative risk estimates.
129.	Peer Reviewer #6	Results	There is no discussion of heterogeneity (or subgroup/stratified analyses if heterogeneity is present) with respect to the pooled analyses.	We have now inserted information on the heterogeneity for all of the pooled analyses
130.	Peer Reviewer #6	Results	The analyses that stratify results (qualitative) according to propensity matched vs. non-propensity matched are interesting but don't seem to show much difference. I'm not aware of studies showing that propensity matching gives different results than other methods of matching or adjustment—what was the reason for doing this analysis? (There are many methodological features that studies could have been stratified by).	We agree that there are many methodological features that we could have stratified the analyses by, though concerns about quality, we did not include observational studies in quantitative synthesis that did not use propensity matching for controls or similar methods. We do include their findings in the associated tables or text to allow the reader to compare their findings with the full set of included studies.
131.	Peer Reviewer #6	Results	The results on p 143 are very confusing. It first says there were 6 RCT's but it looks like data are only reported for 2. The forest plot shows no difference in risk of CVA, based on three trials, but this seems mostly related to differential dosing in one of the trials (low-dose was associated with increased risk), and the SOE ratings are high showing a benefit for one but no benefit for the other two, when the estimates essentially look the same.	Although we present graphically the meta analysis of the three Xa inhibitors, given the heterogeneity in the trials we evaluate the SOE for the individual Xa inhibitors compared to warfarin.
132.	Peer Reviewer #6	Results	Table 51 doesn't provide the comparison for the first two trials.	These details have now been added in to the column headings

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
133.	Peer Reviewer #6	Results	Table 52-54 (and others)—would make sense to organize by medication.	The tables are organized such that findings with the same study but for different drugs are grouped together. We felt that allowing the reader to see how the individual drugs performed in the same study was important. We have then grouped the studies by drug when we combine them for individual meta analyses.
134.	Peer Reviewer #6	Results	Table 67—the description of results is pretty lengthy and could be shortened.	We have tried to reduce the results when possible though acknowledge that there are a lot of included studies, interventions, and outcomes and therefore the results section is quite lengthy.
135.	Judy Wagner (Public Reviewer #2)	Results	Page:117 Para: 5 Recommendation: While the authors mention that there are differences among the pivotal RCT DOAC trials, they fail to mention that the 3 trials differ substantially relative to the patients' baseline stroke (i.e. CHADS2 score) and underlying comorbidities. The differences in patient populations and underlying comorbidities, make either direct or indirect comparisons across these trials unreliable and misleading.	We have now included text about this difference in patient population in the introduction text to the section.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
136.	Judy Wagner (Public Reviewer #2)	Results	<ul style="list-style-type: none"> The 3 pivotal Phase 3 NOAC trials differ substantively relative to the patients' baseline stroke risk (ie, CHADS2 score) and underlying comorbidities. Variations in baseline risk influence the rates of efficacy and bleeding, confounding comparisons among the trials. Of note, the mean CHADS2 score in ROCKET AF was 3.48, reflecting a high stroke risk, whereas it was 2.1 in ARISTOTLE and 2.8 in ENGAGE. In ROCKET AF, 87% of patients had CHADS2 score of ≥ 3, compared to 30% in ARISTOTLE and 53% in ENGAGE AF. Thus, ROCKET AF reflects a much higher risk population than ARISTOTLE and ENGAGE, which would be expected to have higher rates of both bleeding and strokes. 	We now include this additional detail in the introduction text to this section.
137.	Judy Wagner (Public Reviewer #2)	Results	<ul style="list-style-type: none"> The rivaroxaban ROCKET AF trial had more patients with comorbidities, thus reflecting a more complex population. The ROCKET AF trial included a substantially higher number of patients with prior stroke/TIA (55%) compared and ARISTOTLE (20%) and ENGAGE (29%). Prior stroke has been linked to increased risk of bleeding, including ICH, again limiting the usefulness of indirect comparison on bleeding results. Moreover, the ROCKET AF trial included a higher proportion of patients with comorbidities such as diabetes (ROCKET 40%; ARISTOTLE 25%, ENGAGE 36%) and CHF (ROCKET 63%; ARISTOTLE 35%, ENGAGE 58%). 	We have added this information to the introduction text of this section and highlight that given the differences between the trial populations our SOE ratings focus on evidence for individual drugs rather than for Xa inhibitors more broadly.



#	Commentator & Affiliation	Section	Comment	Response
138.	Judy Wagner (Public Reviewer #2)	Results	<p>Page: 119 Table 51 Existing Text: Stroke or Systemic Embolism for ROCKET AF – only results from the ITT analysis and per-protocol analysis are provided. Recommendation: Include data from ROCKET AF on rivaroxaban vs. warfarin for stroke or systemic embolism for the safety, as-treated population. Rates were: 1.7% per year for rivaroxaban vs. 2.22% per year for warfarin; HR 0.79;95% CI,0.65, 0.95; p=0.02. (safety on treatment population). Reference: Patel MR, Mahaffey KW, Garg J, et al. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. N Engl J Med. 2011;365(10):883-91.</p>	We have included the as-treated findings in the main results text but focus on the ITT findings in the SOE summary.
139.	Judy Wagner (Public Reviewer #2)	Results	<p>Page: 123 Para:1 Existing Text: Stroke or Systemic Embolism for ROCKET AF – only results from the ITT analysis and per-protocol analysis are provided. Recommendation: Include data from ROCKET AF on rivaroxaban vs. warfarin for stroke or systemic embolism for the safety, as treated population. Rates were: 1.7% per year for rivaroxaban vs. 2.22% per year for warfarin; HR 0.79;95% CI,0.65, 0.95; p=0.02. (safety on treatment population). Reference: Patel MR, Mahaffey KW, Garg J, et al. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. N Engl J Med. 2011;;365(10):883-91.</p>	We have included the as-treated findings in the main results text but focus on the ITT findings in the SOE summary.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
140.	Judy Wagner (Public Reviewer #2)	Results	<p>Page 132 Para:1 Recommendation: For context, consider adding statement highlighting the differences in the definitions of major and nonmajor clinically relevant major bleeding among the various DOAC RCTs. The differences in these definitions, in addition to the differences in patient populations, make either direct or indirect comparisons across these trials unreliable and misleading (definitions provided below). Reference: Camm et al. Challenges in comparing the non-vitamin K antagonist oral anticoagulants for atrial fibrillation-related stroke prevention. <i>Europace</i>;2018:20, 1-11.</p>	We have added in a textual description of the definitions of major bleeding used in the various trials.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
141.	Judy Wagner (Public Reviewer #2)	Results	<p>Page: 139 Para: 1 Recommendation: Suggest adding context on the nature of GI bleeds from Sherwood, et al publication.</p> <p>For the most severe bleeds, such as events that were fatal and those that required ≥ 4 U of whole blood or PRBC transfusion, there was no significant difference between rivaroxaban and warfarin. Fatal GI bleeds were as follows: rivaroxaban 0.01 events per 100 patient years vs. warfarin: 0.04 events per 100 patient years; $p=0.15$. Transfusion ≥ 4 U were as follows: rivaroxaban 0.47 events per 100 patient years vs. warfarin 0.41 events per 100 patient years; $p=0.39$</p> <p>Reference: Sherwood MW, Nessel CC, Hellkamp AS, et al. Gastrointestinal Bleeding in Patients With Atrial Fibrillation Treated With Rivaroxaban or Warfarin: ROCKET AF Trial. J Am Coll Cardiol. 2015;;66(21):2271-81.</p>	We feel that this level of detail is beyond what is needed in the report and have not made any changes to the text.
142.	Judy Wagner (Public Reviewer #2)	Results	<p>Page: 146 / Para 2/ Page 9 Please Note: Typo for p value “$p=0.22$” should be “$p=0.022$”</p>	This error has been corrected.
143.	Judy Wagner (Public Reviewer #2)	Results	<p>Page: 154/ Para: 7 Thrombo-embolic stroke section Please note that “nearly significant” is not an accurate way to describe findings that do not reach statistical significance. Report should be consistent on how nonsignificant findings are reported.</p>	We have corrected throughout the report to describe studies as demonstrating a trend towards an outcome.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
144.	Judy Wagner (Public Reviewer #2)	Results	<p>Page:156 / Para 7 Medication Adherence Section Recommendation: Consider including the following additional adherence studies: Reference:</p> <ul style="list-style-type: none"> • Crivera C, Nelson WW, Bookhart B, et al. Pharmacy quality alliance measure: adherence to non-warfarin oral anticoagulant medications. <i>Curr Med Res Opin.</i> 2015;31(10):1889-95. • McHorney CM, Crivera C, Laliberté F, et al. Adherence to Non-VKA Oral Anticoagulant Medications Based on the Pharmacy Quality Alliance Measure. <i>Curr Med Res Opin.</i> 2015;31(12):2167-73. • Nelson WW, Song X, Thomson E, et al. Medication Persistence and Discontinuation of Rivaroxaban and Dabigatran Therapy among Patients with Non-valvular Atrial Fibrillation. <i>Curr Med Res Opin.</i> 2015;31(10):1831-40. 	<p>The study by Crivera and colleagues was not included (excluded at full text) as it was unclear that the patients were restricted to nonvalvular AF. The first citation from McHorney and colleagues was an abstract and therefore excluded for not being a full article. The study by Nelson is now included in the report</p>
145.	Judy Wagner (Public Reviewer #2)	Results	<p>Page: 156 Para: 3 Mortality section Please note that “nearly significant” is not an accurate way to describe findings that do not reach statistical significance. This report should be consistent on how nonsignificant findings are reported.</p>	<p>We have corrected throughout the report to describe studies as demonstrating a trend towards an outcome.</p>



#	Commentator & Affiliation	Section	Comment	Response
146.	Judy Wagner (Public Reviewer #2)	Results	<p>Page: 159/ Para 3 Medication Persistence Section</p> <p>Recommendation: We suggest that this section be renamed to “Medication Adherence” to be consistent with the previous sections and consider including the following additional adherence studies:</p> <p>Reference:</p> <ul style="list-style-type: none"> • Crivera C, Nelson WW, Bookhart B, et al. Pharmacy quality alliance measure: adherence to non-warfarin oral anticoagulant medications. <i>Curr Med Res Opin.</i> 2015 Oct;31(10):1889-95. • McHorney CM, Crivera C, Laliberté F, et al. Adherence to Non-VKA Oral Anticoagulant Medications Based on the Pharmacy Quality Alliance Measure. <i>Curr Med Res Opin.</i> 2015;31(12):2167-73. • McHorney CM, Peterson E, Laliberte F, et al. Comparison of Adherence to Rivaroxaban Versus Apixaban Among Patients With Atrial Fibrillation. <i>Clin Ther.</i>2016;38(11):2477-2488. • Coleman C, Yuan Z, Schein J, Crivera C, Ashton V, Laliberte F, Lefebvre P, Peterson E. Importance of Balancing Follow-up Time and Impact of Oral Anticoagulant Users Selection when Evaluating Medication Adherence in Atrial Fibrillation Patients Treated with Rivaroxaban and Apixaban. <i>Curr Med Res Opin.</i> 2017;33(6):1033-1043. 	<p>We have renamed the section as suggested.</p> <p>The study by Crivera and colleagues was not included (excluded at full text) as it was unclear that the patients were restricted to nonvalvular AF. The first citation from McHorney and colleagues was an abstract and therefore excluded for not being a full article. The second study by McHorney and the study by Coleman are now both included in our report.</p>



#	Commentator & Affiliation	Section	Comment	Response
147.	Thomas Seck (Public Reviewer #4)	Results	<ul style="list-style-type: none"> Pg 123 <p>The authors appear to examine both intent to treat (ITT) and modified intent to treat (mITT) populations when reporting findings on outcomes related to stroke or systemic embolism. Because mITT has been found to be inconsistently applied in RCTs due to inconsistent and subjective application, BI recommends that results should be consistently reported across all trials using the ITT, while not interweaving mITT. Use of mITT may lead to confusion, inaccurate results and bias.</p>	We have kept the modified ITT within Table 54 but have changed the order of the findings so that they are after the main ITT findings. We have removed the mITT findings from the results text. Note also that the ITT findings are the focus of the findings within the SOE evidence tables.
148.	Thomas Seck (Public Reviewer #4)	Results	<ul style="list-style-type: none"> Page 98 <p>In reference to this particular paragraph, we recommend that the statement “although 110mg is approved for other uses and therefore can be used off-label” be removed and instead, replaced with the following statement: “Dabigatran 110mg is not approved for stroke prevention in atrial fibrillation in the US.”</p>	We have made the suggested edit.
149.	Thomas Seck (Public Reviewer #4)	Results	<ul style="list-style-type: none"> Pg 83 <p>The word “similar” is misspelled in the paragraph starting with ‘Dabigatran at a 110mg...’</p>	The typo has been corrected.



#	Commentator & Affiliation	Section	Comment	Response
150.	Thomas Seck (Public Reviewer #4)	Results	<p>In addition to ensuring evidence is comprehensively reflected, we urge the authors to ensure all studies, both those already referenced and those to be added, are cited appropriately and accurately within in the review. In particular, we highlight the below instances where findings should reflect the cited literature or be updated to include missing evidence.</p> <ul style="list-style-type: none"> • Page 47 BI notes that findings from Bohm et al. (2015) and Hijazi et al. (2016) are not included when discussing “Renal Impairment and Stroke Risk Studies.” We recommend the authors incorporate findings from these studies, which found that patients with NVAf being treated with warfarin exhibited a more pronounced decline in renal function compared to those being treated with dabigatran. , 	<p>The study by Bohm and colleagues was reviewed and although it compares renal function decline between treatment groups it does not compare any outcomes of interest between the two treatment groups and therefore was excluded at the full text level.</p> <p>The study by Hijazi was included in our report and is discussed in the subgroup table within KQ3.</p>



#	Commentator & Affiliation	Section	Comment	Response
151.	Thomas Seck (Public Reviewer #4)	Results	<p>Page 154 BI notes that when reporting findings comparing Factor Xa Inhibitors (Apixaban, Rivaroxaban, or Edoxaban) versus Dabigatran, the authors cite 12 observational studies evaluating Xa inhibitors compared with dabigatran; however, we note that the following relevant observational studies, while cited elsewhere in the review, were not included in this section:</p> <p>Gorst-Rasmussen A, Lip GY, Bjerregaard Larsen T. Rivaroxaban versus warfarin and dabigatran in atrial fibrillation: comparative effectiveness and safety in Danish routine care. <i>Pharmacoepidemiol Drug Saf.</i> 2016 Nov;25(11):1236-1244. Epub 2016 May 27.</p> <p>Lip GY, Pan X, Kamble S, Kawabata H, Mardekian J, Masseria C, Bruno A, Phatak H. Major bleeding risk among non-valvular atrial fibrillation patients initiated on apixaban, dabigatran, rivaroxaban or warfarin: a "real-world" observational study in the United States. <i>Int J Clin Pract.</i> 2016 Sep;70(9):752-63. Epub 2016 Aug 23.</p> <p>To ensure evidence is comprehensively reflected, we recommend the authors incorporate findings from these studies in this section.</p>	We have now included these two studies in the section targeting Xa inhibitors compared with dabigatran.
152.	Thomas Seck (Public Reviewer #4)	Results	<ul style="list-style-type: none"> Page 156 Under GI bleeding, please incorporate findings from Norby et al. (2017) which indicate that a higher rate of GI bleeding was observed with rivaroxaban compared to dabigatran. 	This study is now included in our pooled analysis of GI bleeding



#	Commentator & Affiliation	Section	Comment	Response
153.	Thomas Seck (Public Reviewer #4)	Results	<ul style="list-style-type: none"> Page 156 <p>When discussing mortality outcomes under the Other Clinical Outcomes section, the authors only cite “One prospective cohort study using Medicare claims data for adults ≥65 years of age and using dabigatran vs. rivaroxaban for nonvalvular AF;” however BI notes that data from Gorst-Rasmussen et al. (2016) is not discussed. We recommend that data from this study, which found a higher rate of all-cause mortality for rivaroxaban versus dabigatran, be incorporated in to the findings for this section.</p>	This study is incorporated in to the table and pooled analysis of all-cause mortality.
154.	Thomas Seck (Public Reviewer #4)	Results	<ul style="list-style-type: none"> Page 170 <p>Under “Outcomes for Specific Subgroups of Interest,” we recommend the inclusion of an additional subgroup to include patients with Nonvalvular Atrial Fibrillation (NVAF) undergoing percutaneous coronary intervention (PCI). , We also recommend the inclusion of a subgroup for patients undergoing ablation.</p>	Patients undergoing ablation was considered out of scope for this report.
155.	Peer Reviewer #1	Discussion/ Conclusion	<p>page 197.</p> <p>when discussing what we do with bleeding risk scores I appreciate that the authors clearly state we generally do not use bleeding risk score to decide not to prescribe anticoagulation. would recommend they mention that bleeding risk scores may help guide intensity of follow up/monitoring and decision making around combination therapy (ie safety of antiplt therapy in addition to anticoagulation).</p>	We agree and we have added the following to the report. “While bleeding risk scores are generally not used to decide whether or not to use an oral anticoagulant in a given patient, high scores may help guide intensity of patient follow-up and monitoring.”

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
156.	Peer Reviewer #1	Discussion/ Conclusion	page 190 emphasizing the lack of effectiveness of aspirin is greatly appreciated. I think the fallacy of benefit of aspirin is responsible for much of the underuse of anticoagulation-providers think they are doing something with ASA. highlighting the AVVEROES trial here was wise.	Thank you. No response needed
157.	Peer Reviewer #1	Discussion/ Conclusion	p 195 "Additional work will be required to develop risk tools for patients to discriminate those individuals with AF where the bleeding risk may outweigh the stroke prevention benefit."-I am not sure about this statement. I don't foresee bleeding risk scores resulting in prevention of use of anticoagulation. I see them guiding therapy and monitoring and allowing illumination of balance of risk benefit to include the patient in informed decision making.	We have modified this statement as follows: "Additional work will be required to develop risk tools for patients to discriminate those individuals with AF where the bleeding risk may be high enough to warrant more intensive follow-up and monitoring"-
158.	Peer Reviewer #2	Discussion/ Conclusion	P 189 Li 4-11. A more quantitative assessment of consistency seems warranted here. How different were point estimates, and were confidence intervals incompatible? The phrases "did not demonstrate" and "did not show" may just mean that p-values were >0.05, which is not very informative. See https://urldefense.proofpoint.com/v2/url?u=http-3A__www.stat.columbia.edu_-7Egelman_research_published_signif4.pdf&d=DwlFaQ&c=imBPVzF25OnBgGmVOlcsiEgHoG1i6YHLR0Sj_gZ4adc&r=9HI35bRoP7Lbw6vdF6tBw_ijOOJWHA7dyYcqHbH849A&m=qPmmKnezC5gLht44cxEd8lwM82FWw7L4lxpuao08jrs&s=F15z9Okf1lhUaT4GkgAtBnS0acO4PyKKeLt2iQQTj2o&e=	We have clarified that the studies did not demonstrate evidence of a difference.
159.	Peer Reviewer #2	Discussion/ Conclusion	Table 71. I would clarify what the entries are in the table's cells.	We have clarified the contents of this table.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
160.	Peer Reviewer #5	Discussion/ Conclusion	See general comments (#s 24-36)	No response needed – addressed in general comments section
161.	Peer Reviewer #6	Discussion/ Conclusion	ES: The “What Does This Review Add...” section doesn’t really explain how the info on builds on what is already known—it reads like it is just presenting more results. Several bullets talk about inconsistency or consistency between RCTs and observational studies but again it is not clear what is new and how it adds to our understanding of the relevant comparisons. In fact some of the bullets on inconsistencies and need for more validation of tools sound more like limitations and research gaps.	We have expanded this section to include information on the various treatments. We agree however that this section overlaps with the limitations and future research section and so have clarified this in the introduction to the text
162.	Peer Reviewer #6	Discussion/ Conclusion	ES: In the Limitations/Future Research section it is not entirely clear why indirect comparisons are less relevant in these field than in others—what makes the heterogeneity in this case more of an issue? I think the bigger issue is that there are relatively few RCT’s to contribute to a network analysis.	We now clarify in to the introduction text within the results section the significant differences between these trial populations – highlighting the issues with indirect comparisons given these studies.
163.	Peer Reviewer #6	Discussion/ Conclusion	ES: The bullet on needing strategies to improve use and adherence of using oral anticoag seems to come out of the blue—I don’t recall seeing anything earlier about issues with adherence and use of oral anticoag.	We have removed this bullet.
164.	Peer Reviewer #6	Discussion/ Conclusion	KQ 1: Having more evidence on CHADS2 is not the same as evidence showing that it performs better than the other instruments—again would focus on the head-to-head comparisons in terms of differences in discrimination.	We now clarify the comparability of the CHADS2, CHA2DS2-VASc, and ABC stroke risk scores



#	Commentator & Affiliation	Section	Comment	Response
165.	Peer Reviewer #6	Discussion/ Conclusion	KQ 2: As noted above, I think it is very difficult to justify the conclusion that HAS-BLED is superior to the other instruments, since most studies showed no difference in discrimination.	Based on the studies where direct comparisons of the bleeding risk scores were evaluated, HAS-BLED does appear to have the strongest evidence supporting its comparative predictive abilities.
166.	Peer Reviewer #6	Discussion/ Conclusion	KQ 3: See previous comments about the Xa inhibitors and interpretation of results on p 143. At least at higher (non-renal) doses the effectiveness looks quite similar.	We do not feel that our report supports saying that one specific DOAC is preferred versus another. We have tried to ensure that our findings reflect the underlying evidence and the heterogeneity in findings. We've also emphasized the need for future studies which have direct comparisons of the DOACs.
167.	Michael Rich (Public Reviewer #1)"	Discussion/ Conclusion	I do not think that the Key Messages and Conclusions are complete and balanced. Specifically -- 1. It is stated correctly that 'Dabigatran (150mg dose) is superior to warfarin in reducing stroke with no difference in harms', but there should be a 2nd statement that 'Dabigatran (110mg dose) is equivalent to warfarin in reducing stroke with less major bleeding', an issue of substantial importance in the care of older adults.	We have added a second bullet about the 110 mg dose of dabigatran
168.	Michael Rich (Public Reviewer #1)"	Discussion/ Conclusion	2. For rivaroxaban, a more accurate statement that is also more consistent with the style of the dabigatran statement might be 'Rivaroxaban reduced the risk of hemorrhagic stroke compared to warfarin with no difference in other outcomes or harms.'	We have maintained our current key point related to rivaroxaban



#	Commentator & Affiliation	Section	Comment	Response
169.	Michael Rich (Public Reviewer #1)"	Discussion/ Conclusion	3. For edoxaban, a more accurate and consistent statement might be 'Edoxaban reduced the risk of hemorrhagic stroke and major bleeding compared to warfarin with no difference in other outcomes or harms.'	We have maintained our current key point related to edoxaban which highlights the differences in findings for the different doses of edoxaban.
170.	Michael Rich (Public Reviewer #1)"	Discussion/ Conclusion	4. I think it would be worth noting that all of the newer agents reduced the risk of hemorrhagic stroke compared to warfarin.	We have revised the concluding statements to more fully describe the effectiveness of the different DOACs on outcomes of interest.
171.	Michael Rich (Public Reviewer #1)"	Discussion/ Conclusion	These points should also be integrated into the Abstract and the manuscript itself.	Changes noted above have been made throughout the report.
172.	Judy Wagner (Public Reviewer #2)	Discussion/ Conclusion	Page: 195 Para: 3 Line: 2 Existing text: Apixaban shows safety and efficacy in patients who are not candidates for warfarin. Recommendation: Based on the evidence, all DOACS have shown safety and efficacy in patients who are not candidates for warfarin.	Although potentially true, included studies did not explicitly explore the effectiveness of other DOACs in this patient population. Apixaban is the only agent that has been studied in a large RCT in these specific patients. We have retained the text as is.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
173.	Ryne Carney (Public Reviewer #3)	Discussion/ Conclusion	A key challenge in this space is improving the current risk stratification tools that could inform clinical decision making. Both clinicians and AFib patients need stratification tools that weigh the balance between bleeding episodes and stroke risk. We agree with the report's assessment that a single tool balancing both thromboembolic risk and bleeding risk for decision making on antithrombotic therapy would be a huge step forward for this disease area. In 2015, the Task Force submitted comments to the National Heart, Lung, and Blood Institute (NHLBI) for its Strategic Vision, which was released in August 2016. Due to our advocacy, NHLBI has specifically identified the need more expanded research on bleeding risk in older adults with AFib. Such research is necessary to develop more accurate risk stratification tools, and we encourage both AHRQ and PCORI to partner with NHLBI and proactively follow this effort, as we will do within the Task Force. The lack of research in this area is a critical barrier to progress in reducing stroke incidence.	We agree that there is much needed research for risk stratification tools that balance outcomes of interest. This is highlighted in our research gaps section of the discussion.
174.	Ryne Carney (Public Reviewer #3)	Discussion/ Conclusion	We encourage studies of head-to-head comparisons of the safety and efficacy of dabigatran, rivaroxaban, apixaban, and edoxaban. As cited in the draft report, cross-trial comparisons of these DOACs may not be currently possible because of inconsistencies in dosing strategies, definitions, and the stroke and bleeding risk of recruited populations. When more evidence becomes available on these head-to-head studies such as NCT02666157 and NCT03129490, we would encourage the EPC to conduct another review of the relevant literature when these trials are complete.	As we highlight in our report, we agree with the reviewer that head-to-head comparisons of the DOACs are needed and that it will be important to incorporate their findings in the future reviews.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
175.	Ryne Carney (Public Reviewer #3)	Discussion/ Conclusion	<p>We would encourage research into strategies for improving patient adherence to the use of oral anticoagulants. The Task Force has been active on the issue related to “nonmedical switching” of oral anticoagulants. Non-medical switching refers to a formulary decision-making process designed by Part D drug plans to limit prescription coverage to less expensive medications (also called formulary-driven switching). The change in medication is determined by the plan formulary without any consideration of the medical repercussions or a physician’s knowledge and reasoning behind the selection of the original prescription medication. While addressing non-medical switching is likely outside the scope of this CER, we encourage the report to highlight the dearth for research into the medical consequences of switching between oral anticoagulants, including novel oral anticoagulants (NOAC) to NOAC and NOAC to warfarin. The Task Force sent a letter of request to CMS to explore the scale and impact of this issue, and Principal Deputy Administrator of CMS and Director of the Center for Medicare, Demetrios Kouzoukas, responded in an October 2017 letter that simply reviewed the current Part D process for notice, transition fill, and the process for enrollees or prescribing providers to appeal to their Part D plan sponsor. It did not in any way address the specific concerns we raised. Therefore, a mention in this current evidence review would be particularly meaningful to encourage future focus on this issue.</p>	<p>We agree that adherence is an important topic although given the consistencies in the evidence regarding adherence it was not highlighted in the current report as an area of future research to prioritize.</p>

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
176.	Ryne Carney (Public Reviewer #3)	Discussion/ Conclusion	There is a deficit of studies assessing the safety and effectiveness of different combinations of antithrombotic medications at varying dosages and duration.	We agree that in addition to direct comparisons of the various DOACs, that studies of their specific dosages and durations are needed.
177.	Ryne Carney (Public Reviewer #3)	Discussion/ Conclusion	More research is required to develop effective tests and measures for blood-thinning ranges of AFib patients and patient subgroups who take oral anticoagulants. When NOACs were initially being studied, there was an emphasis on comparing these drugs to warfarin, without the need for blood monitoring. Future research on reductions in stroke and bleeding from blood monitoring on NOACs would be useful. A joint Hadassah-Hebrew University Medical Center and Mayo Clinic assessing NOAC drug level monitoring in clinical patient management was published in March 2018. The study concluded that “future studies are warranted to establish associations between drug levels and outcomes, and better delineate the role of DOAC monitoring.”	We agree that this is an important area of needed future research.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
178.	Thomas Seck (Public Reviewer #4)	Discussion/ Conclusion	<ul style="list-style-type: none"> Page 195 The authors call out that “Apixaban shows safety and efficacy in patients who are not candidates for warfarin” in the last paragraph of the conclusion section. We are concerned that the placement and wording of this sentence could be interpreted to suggest that apixaban is the solely recommended DOAC and that warfarin is the primary OAC of choice. Further, apixaban is the only DOAC to be mentioned in the conclusion section, which places an arbitrary emphasis on this recommendation. We recommend the authors remove this sentence or further expand this section to include other DOACs so that the conclusions are balanced and representative of the review findings. 	We have removed this sentence from the concluding paragraph.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
179.	Thomas Seck (Public Reviewer #4)	Discussion/ Conclusion	<p>Opportunities to strengthen the report through acknowledgement and/or more robust discussion of reversal agents include:</p> <ul style="list-style-type: none"> Page 195 The limitations section begins to discuss the potential for reversal or antidote to impact patient outcomes; however, we note this section remains largely unchanged from the 2013 review. We support the authors' continued recommendation that further study on the development of antidotes for severe bleeding events is needed; however, the authors should update this review to acknowledge that idarucizumab is currently approved as a specific reversal agent to PRADAXA® (dabigatran). At minimum, we recommend AHRQ provide references to product information and published studies on PRAXBIND® to note its availability and acknowledge the foundation for further evidence generation in this area. 	We believe that the text currently included in the report appropriately identifies this as an area of needed research while acknowledging that this is outside of the scope of the current report.
180.	Judy Wagner (Public Reviewer #2)	Figures/ Tables	Please Note: Multiple 95% CI's in Figures 18, 19, 24, 26, 27, 28, 30, 31 - Forest Plot do not match the CI's in Tables 52, 53, 54, 55, 57, 58	The confidence intervals reported in the tables are reported directly from the included studies. In the forest plots these confidence intervals are calculated directly and given rounding (normally down to two digits) in papers may differ slightly.



#	Commentator & Affiliation	Section	Comment	Response
181.	Judy Wagner (Public Reviewer #2)	Figures/ Tables	Pages: 125-126 Figures: 18 & 19 Recommendation: Consider also including the rivaroxaban 20mg HR from Gorst-Rasmussen in the Forest Plot. Reference: Gorst-Rasmussen A, Lip GY, Bjerregaard Larsen T. Rivaroxaban versus warfarin and dabigatran in atrial fibrillation: comparative effectiveness and safety in Danish routine care. <i>Pharmacoepidemiol Drug Saf.</i> 2016 Nov;25(11):1236-44	The forest plot included data from the 15mg dose for this specific study.
182.	Judy Wagner (Public Reviewer #2)	Figures/ Tables	Page: 129 Figure: 22 Recommendation: The rivaroxaban HR from Coleman, 2016 (REVISIT-US) is not included in the forest plot, please add. 0.71 (0.47 to 1.07) Reference: Coleman CI, Antz M, Bowrin K, et al. Real-world evidence of stroke prevention in patients with nonvalvular atrial fibrillation in the United States: the REVISIT-US study. <i>Curr Med Res Opin.</i> 2016 Dec;32(12):2047-53	The study is now included in the forest plot.
183.	Judy Wagner (Public Reviewer #2)	Figures/ Tables	Page: 142 Table: 59 Recommendation: Add missing a decimal point in 95% CI for study reference 291 Apixaban 5mg bid 0.47 (0.29 to 0.76)	This typo has been fixed



#	Commentator & Affiliation	Section	Comment	Response
184.	Judy Wagner (Public Reviewer #2)	Figures/ Tables	<p>Page: 143 Figure: 33 Recommendation: Consider including the rivaroxaban 20 mg HR for Gorst-Rasmussen as well in the Forest plot. Rivaroxaban 20 mg vs. Warfarin 0.93 (0.75 to 1.16) Reference: Gorst-Rasmussen A, Lip GY, Bjerregaard Larsen T. Rivaroxaban versus warfarin and dabigatran in atrial fibrillation: comparative effectiveness and safety in Danish routine care. <i>Pharmacoepidemiol Drug Saf.</i> 2016 Nov;25(11):1236-44.</p>	The forest plot included data from the 15mg dose for this specific study.
185.	Judy Wagner (Public Reviewer #2)	Figures/ Tables	<p>Page: 149 Table: 62 Exiting text: "N" value for ROCKET AF RCT rivaroxaban vs. warfarin is listed as 15,544 in several rows of the table.</p> <p>Recommendation: The reference is cited as #24, Patel, et al (ROCKET-AF study), which only includes 14, 264 randomized patients. Please verify correct N value.</p>	This error has been fixed



#	Commentator & Affiliation	Section	Comment	Response
186.	Judy Wagner (Public Reviewer #2)	Figures/ Tables	<p>Page: 149 Table: 62 Existing text: For rivaroxaban vs. warfarin, the endpoint of stroke/systemic embolism is listed as “inconsistent ”</p> <p>Recommendation: Suggest reconsidering this classification and revising to “consistent”.</p> <p>As noted in the table, in the ROCKET AF RCT, there was “No difference in stroke/systemic embolism between rivaroxaban vs. warfarin (HR 0.88; 95% CI 0.74 to 1.03) “. As noted on page 124 the document ” We also synthesized the findings for individual drugs. Figure 19 demonstrates that the observational studies combining evidence from the individual drugs found no difference for rivaroxaban versus warfarin (HR 0.82, 95% 0.67 to 1.00).” This demonstrates consistency of the results for stroke/systemic embolism across these studies.</p>	In the revised report which incorporates additional observational studies, rivaroxaban demonstrates a reduction in stroke or systemic embolism as compared to warfarin. We have therefore retained the inconsistent notation.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
187.	Judy Wagner (Public Reviewer #2)	Figures/ Tables	<p>Page: 150 Table: 62 Existing text: For rivaroxaban vs. warfarin for the endpoint of “hemorrhagic stroke” SOE is listed as “low ”</p> <p>Recommendation: Suggest reconsidering this classification and revising to “high”.</p> <p>The authors cite 2 studies, the ROCKET AF RCT and “a smaller study showed a trend toward no difference (HR 0.73; 95% CI 0.16 to 3.25)”. Please clarify which smaller study is this referring to? It is not included in the references. Of note, the recommendation for apixaban is listed as “High” on the basis of the Aristotle RCT. The same standard should be applied for classification of the rivaroxaban data.</p>	<p>The SOE rating was low given the imprecision of the RCT findings, the inconsistencies in findings from individual observational studies, and since the RCT findings were from on-treatment analyses rather than intention-to-treat. We have retained the SOE rating of low for hemorrhagic stroke.</p>
188.	Judy Wagner (Public Reviewer #2)	Figures/ Tables	<p>Page: 151 Table: 62 Existing Text: SOE for intracranial bleeding is listed as” Moderate ”.</p> <p>Recommendation: Suggest reconsidering this classification and revising to “High”. The basis of the recommendation is the ROCKET AF RCT trial. The safety population is used, which is the same population used in other DOAC trials to assess safety endpoints. Of note, the recommendation for apixaban is listed as “High” on the basis of the Aristotle RCT. The same standard should be applied for classification of the rivaroxaban data.</p>	<p>We agree that there was an inconsistency in this rating and we have been changed the rating to be high SOE.</p>



#	Commentator & Affiliation	Section	Comment	Response
189.	Judy Wagner (Public Reviewer #2)	Figures/ Tables	<p>Page: 172 Table: 67 Recommendation: For AF and PAD (peripheral artery disease) subgroup, consider adding the following reference:</p> <p>Reference:</p> <ul style="list-style-type: none"> Schuyler-Jones, et al.. Efficacy and safety of rivaroxaban compared with warfarin in patients with peripheral artery disease and non-valvular atrial fibrillation: insights from ROCKET AF European Heart Journal 2014;35, 242–249 	The suggested citation was excluded because patients with peripheral artery disease and non-valvular atrial fibrillation was not defined as a subgroup of interest.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
190.	Judy Wagner (Public Reviewer #2)	Figures/ Tables	<p>Page: 172 Table: 67 Recommendation: Consider adding the following publication assessing obese patients from ROCKET AF (rivaroxaban vs. warfarin) as a subgroup of interest</p> <p>Reference:</p> <ul style="list-style-type: none"> • Balla, et al. Relation of Risk of Stroke in Patients With Atrial Fibrillation to Body Mass Index (from Patients Treated With Rivaroxaban and Warfarin in the Rivaroxaban Once Daily Oral Direct Factor Xa Inhibition Compared with Vitamin K Antagonism for Prevention of Stroke and Embolism Trial in Atrial Fibrillation Trial). <i>Am J Cardiol</i> 2017;119(12):1989-1996. <p>Recommendation: Consider adding an additional subgroup of interest; AF and CAD (coronary artery disease) subgroup with the following reference:</p> <p>Reference:</p> <ul style="list-style-type: none"> • Kochar, et al. Efficacy and safety of rivaroxaban compared with warfarin in patients with carotid artery disease and nonvalvular atrial fibrillation: Insights from the ROCKET AF trial. <i>Clinical Cardiology</i> 2018;41(1):39-45. 	<p>The suggested study was reviewed but excluded at full text since BMI was not a risk factor or subgroup of interest.</p> <p>The study by Kochar is included in the revised report.</p>

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
191.	Thomas Seck (Public Reviewer #4)	Figures/ Tables	<ul style="list-style-type: none"> Page ES-4 Please note the primary outcome and major bleeding figures for dabigatran 150mg are incorrect; they should be updated to : <ul style="list-style-type: none"> Stroke or systemic embolism (RR 0.65 95% CI 0.52-0.81) Major bleeding (RR 0.94 95% CI 0.82-1.08) 	These findings have been checked and come directly from the primary study by Connolly.
192.	Thomas Seck (Public Reviewer #4)	Figures/ Tables	Opportunities to strengthen the report through acknowledgement and/or more robust discussion of reversal agents include: <ul style="list-style-type: none"> Page ES-8, Table ES-1 The table, which provides a comprehensive list of pharmacologic interventions, should mention the existence of a reversal agent for dabigatran	The report did not focus on reversal agents and we do not feel that this information is needed in the summary tables.
193.	Thomas Seck (Public Reviewer #4)	Figures/ Tables	<ul style="list-style-type: none"> Page 171, Table 67 Under the subgroup on Patients with paroxysmal versus sustained AF, the authors should include findings from Flaker et al. (2012), which notes that “Dabigatran etexilate 150 mg bid was more effective than warfarin for stroke and systemic emboli in paroxysmal AF (hazard ratio [HR]: 0.61, 95% confidence interval [CI]: 0.42 to 0.90), in persistent AF (HR: 0.64, 95% CI: 0.43 to 0.93), and in permanent AF (HR: 0.70, 95% CI: 0.48 to 1.01).”	The suggested citation is a research letter rather than a full publication and as such was excluded at the full text level.



#	Commentator & Affiliation	Section	Comment	Response
194.	Thomas Seck (Public Reviewer #4)	Figures/ Tables	<ul style="list-style-type: none"> Page 173; Table 6 <p>Study results from Brambatti et al. (2015) are not accurately reflected in the table summarizing findings for the “patients with AF and diabetes” subgroup. Specifically, the authors found that the absolute reduction in stroke or systemic embolism with dabigatran compared to warfarin was greater among patients with diabetes than those without diabetes (dabigatran 110 mg: 0.59% per year vs. 0.05% per year; dabigatran 150 mg: 0.89% per year vs. 0.51% per year). This absolute reduction is not clearly reflected in the findings.</p>	<p>This summary states that in the third study dabigatran reduced stroke in diabetic patients which was not seen in non-diabetic patients.</p>
195.	Peer Reviewer #1	Clarity/ Usability	<p>nothing new with risk scores...still with chads2 and chadsvasc. has bled still in the lead.so not a lot added there</p> <p>I think the comparative info between anticoagulants and antiptls is helpful and highlighting results from RCTs vs observations trials is important.</p> <p>this report helps to inform new areas of study.</p> <p>I would like to see some inclusion of the LACK of benefit of warfarin when TTR is low. I think this lost on many providers and this information should drive policy and practice around anticoagulation choice. low TTR (ie < 70%) could trigger transition to DOAC.. or you could suggest further inquiry into this area.</p>	<p>We agree with this great point. We have added the following to the discussion: “It is important to note that for warfarin to be effective, time in the therapeutic range has to be high; patients in whom this is hard to achieve should be considered for other types of oral anticoagulants”.</p>
196.	Peer Reviewer #2	Clarity/ Usability	<p>The report seems to me to be well-structured and clear. It may be useful to number sections, subsections, and sub-subsections, such as “1.”, “1.1”, “1.1.1”.</p>	<p>Thank you for the suggestion</p> <p>The format used reflect requirements from both AHRQ and PCORI and has not been modified.</p>



#	Commentator & Affiliation	Section	Comment	Response
197.	Peer Reviewer #2	Clarity/ Usability	As noted above, I believe that usability would be enhanced by more complete presentation of and emphasis on the important quantitative estimates. I do not believe that c-statistics are usable, and the qualitative labels given to the different ranges does not change that.	We have restructured the tables in KQ1 and KQ2 to emphasize the event rates by score and to aid the reader in their comparison.
198.	Peer Reviewer #2	Clarity/ Usability	The introductory discussion for Key Question 2 also seems applicable to Key Question 1.	We now include similar introductory text in the KQ1 section.
199.	Peer Reviewer #3	Clarity/ Usability	Very relevant review on an important clinical question.	Thank you. No response needed
200.	Peer Reviewer #4	Clarity/ Usability	Is the report well structured and organized? YES, in parts it appears somewhat repetitive, which is frustrating for a large document, but overall it is well organized	Thank you. No response needed
201.	Peer Reviewer #4	Clarity/ Usability	Are the conclusions relevant to policy or practice decisions? The conclusions are relevant to practice and research. There are no concrete implications for policy. In this review, am not sure a discussion of policy implications is needed, however.	Thank you. No response needed
202.	Peer Reviewer #5	Clarity/ Usability	See general comments (#s 24-36)	No response needed – discussed previously in the general comments section
203.	Peer Reviewer #6	Clarity/ Usability	As noted in the General Comments, giving the very broad scope of this report making it digestible is a challenge.	We agree that the broad scope and extensive evidence base makes this report a challenge. We have attempted to clarify and simplify throughout the final report where possible.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019



#	Commentator & Affiliation	Section	Comment	Response
204.	Peer Reviewer #6	Clarity/ Usability	If possible would focus more on head-to-head comparisons particularly in the section on risk prediction instruments.	We have tried to emphasize the head-to-head comparisons throughout the report – and the need for additional studies across key questions with such comparisons.

Source: <https://effectivehealthcare.ahrq.gov/topics/stroke-afib-update/research-2018>

Published Online: October 30, 2019