



July 21, 2016

RE: BMJ.2016.033290

Atrial fibrillation and the risks of cardiovascular disease, renal disease and death: a meta-analysis

Ayodele Odutayo, Christopher X. Wong, Sally Hopewell, Douglas Altman, Allan J. Hsiao, Connor A. Emdin

First and foremost, we would like to thank the editorial staff and reviewers for their time and efforts. All editorial changes are tracked in **yellow highlight** in the manuscript and new text is provided here as needed in "italics". As requested by the editorial staff, we have reviewed the style guidelines and adjusted our manuscript accordingly.

Finally, we would welcome editorial guidance regarding whether some of the supplementary tables or figures should be moved to the main manuscript to improve its readability.

Thank you for your consideration.

Ayodele Odutayo and Connor Emdin

Response to Reviewers

Editorial Comments

- 1. Please add a sentence or two discussion your choice of method for quality assessment. It seems a little unorthodox to us since you don't use standard instruments.**

We have now performed our quality assessment using the New Castle Ottawa scale. We have added the following information describing our approach to the methods section (page 7):

"A risk of bias assessment was performed using the Newcastle-Ottawa Scale, which assesses studies on three broad domains: the selection of participants for study groups; the comparability of study groups; and the ascertainment of the outcome. A star rating system is used and the maximum numbers of stars achievable are: Selection (4 stars), Comparability (2 stars), and Outcome (cohort studies only; 3 stars). We applied strict criteria to assess comparability (of individuals with or without AF) based on which variables were included in the multivariable models. To receive one star for comparability, studies were required to meet the aforementioned criteria for adequate adjustment. To receive two stars, studies were required to meet the criteria for being well adjusted. Finally, studies were considered at low risk of

bias if they achieved a full rating in at least two categories of selection, comparability or outcome assessment."

The risk of bias table is provided in Supplementary Table 6.

2. Does your search needed updating? [only to March 2015 currently].

We have updated the search strategy and included any studies meeting our inclusion criteria. Overall, our meta-analysis now includes 104 studies. One study reports results for peripheral arterial disease.

3. Please also add detail about exclusions. Eg. You excluded studies that did not report which co variables they adjusted for, and also excluded unadjusted studies. Could either of these exclusions introduce bias? How may were excluded for each of these reasons? [the flow chart details only one of them]

It would be inappropriate to include unadjusted studies, as the baseline characteristics of patients with atrial fibrillation (AF) will differ significantly from patients without AF. For example, patients with AF will be older and have more comorbidities than those without AF. Without requiring that studies adjust for baseline characteristics, differences in risk due to age, sex, cardiovascular risk factors and other confounders would be inappropriately attributed to the presence of AF. Accordingly, we limited our meta-analysis to studies that adjusted for these differences using multivariable regression.

We additionally excluded studies that did not report the variables included in their multivariable regression model as we could not determine whether these studies did or did not adjust for important covariates such as comorbidities. Overall, 30 studies were excluded because they were unadjusted or inadequately adjusted and 2 studies were excluded because the variables in the multivariable regression model were not included. The small number of exclusions for studies that did not report the variables in their model are unlikely to meaningfully affect the relative risk estimates obtained in our study.

We have modified our flow chart to clarify the number of studies that were unadjusted/inadequately adjusted versus the number of studies that did not report the variables in their regression model.

4. Please consider adding supplementary figure 1 [flow chart] to the main paper.

We have moved this figure into the main paper.

5. Some of the later funnel plots have very few studies in them - are they valid? Please review.

In accordance with Cochrane, we have removed funnel plots for analyses where there were fewer than 10 studies.

REVIEWER 1

- 1. It would be useful if you would reference the papers in your sub-analysis, in the text and/or in the figures (e.g. in Figure 3 and Table suppl. Figure 3 etc.).**

We have included references for studies contributing to each analysis in the main text of the manuscript.

- 2. In the limitations you should briefly mention the risk of having missed studies in your search. I'm aware of one study from our group (Bang CN, AHJ 2012) with a sub-analysis on new-onset AF and the risk of all-cause mortality in AS patients. Please also include this study in your analysis if suitable.**

We have added a sentence to the limitations noting the potential for missing studies and its potential impact given the size of our meta-analysis. (page 19)

"First, despite our extensive search strategy we may have missed relevant studies for inclusion. However, the large number of included studies made our results robust to the inclusion of any single study and provided us with the power to investigate whether the associations of AF with cardiovascular disease and death differed by important patient and study characteristics."

We have also reviewed the study the reviewer proposed for inclusion. It meets our inclusion criteria and we have included it in our analysis.

REVIEWER 2

- 1. Page 5, Methods: the authors should report their search strategy, so others can replicate or build on this. The search strategy can be supplemented online.**

We have now included the search strategy as Supplemental Table 1.

- 2. Page 5, Methods: describe the study selection criteria (inclusion/exclusion) under a separate header, and explain why you chose to exclude studies with <100 participants and <6 months follow-up**

We have created a separate header for the study selection criteria and we have provided more details to explain why studies with <100 participants and <6 months follow up were excluded.

"In order to avoid overestimating the relative risks associated with AF, studies were required to include a minimum of 50 participants with AF and 50 participants without AF. Studies were also required to have at least 6 months mean/median follow up because we sought to assess the mid-term to long-term risks associated with AF."

- 3. Page 5, Methods: the authors correctly address the potential sources of bias regarding adjusting for potential confounders in multivariable analysis and the method of AF diagnosis**

at baseline. However, for complete study quality assessment the authors should consider providing details regarding the following: I) Selection of exposed (AF) and unexposed (Controls) participants; if Controls were likely to have other health issues, it has to be considered how this affects the reported relative risk and to what extent this can be projected onto the general population. II) Completeness of follow-up data; if for some reason a larger proportion of Controls were lost to follow-up than AF patients, this might bias the reported relative risk. III) Outcome ascertainment during follow-up; was it possible that AF patients were more closely monitored for outcomes than Controls, for example using routine administrative data rather than a regularly monitored prospective cohort study? This might bias the reported association for outcomes other than mortality.

We have now performed our quality assessment using the New Castle Ottawa scale which addresses the items raised by the reviewer. We have included text in the methods section describing our approach (page 7):

“A risk of bias assessment was performed using the Newcastle-Ottawa Scale, which assesses studies on three broad domains: the selection of participants for study groups; the comparability of study groups; and the ascertainment of the outcome. A star rating system is used and the maximum numbers of stars achievable are: Selection (4 stars), Comparability (2 stars), and Outcome (cohort studies only; 3 stars). We applied strict criteria to assess comparability (of individuals with or without AF) based on which variables were included in the multivariable models. To receive one star for comparability, studies were required to meet the aforementioned criteria for adequate adjustment. To receive two stars, studies were required to meet the criteria for being well adjusted. Finally, studies were considered at low risk of bias if they achieved a full rating in at least two categories of selection, comparability or outcome assessment.”

The risk of bias table is provided in Supplementary Table 6.

We are in agreement with the reviewer with respect to whether differences in outcome ascertainment may have contributed to heterogeneity in our study. We have therefore added the following sentence to our limitations (page 19):

“For instance, the criteria in well-established prospective cohort or secondary analyses of randomized trials [8] were comprehensively detailed whereas the criteria and specific diagnostic codes used in analyses of administrative data were not consistently provided [51]. Differences in the diagnostic codes used in administrative data may have contributed to the high heterogeneity in our study.”

4. Page 6, line 13: explain why studies that did not report the covariates were excluded. Are these studies presumably of less quality than studies that were minimally adjusted or unadjusted?

Two studies were excluded because they did not report the variables that were included in their multivariable regression model. We excluded these studies because we could not determine whether they properly adjusted for important cofounders such as age, sex and cardiovascular risk factors. This

small number of studies was unlikely to affect relative risk estimates in our analysis. We have added the following to the methods section to clarify why we excluded these studies:

“Studies that did not report the covariates included in their regression model were also excluded because we could not determine whether they properly adjusted for important cofounders such as comorbidities, age and sex and thus whether they fulfilled the pre-specified inclusion criteria of this meta-analysis.”

5. **Page 6, line 16: I agree to distinguish between ‘minimally’ adjusted and ‘adequately’ adjusted relative risks. However, does it not depend on the outcome of interest regarding which variables should be minimally included as covariates? Age and gender are standard, but for example where CVD history is relevant for ischemic events, for stroke we know that the CHA2DS2-VASc risk factors are associated with stroke. Different risk factors will be relevant for heart failure, bleeding, and other outcomes. With your method, does it mean that the outcome-specific known risk factors are taken in account for the ‘adequately’ adjusted relative risk, rather than ‘minimally’?**

Studies were characterized as adequately adjusted if they included age and sex and baseline cardiovascular disease in their multivariable regression model. Furthermore, studies were well adjusted if they additionally adjusted for at least two of the following risk factors: hypertension, diabetes, smoking or cholesterol. We selected these four risk factors because they contribute to the development of each outcome examined in our study.

There is no accepted standard for what would be considered adequate multivariable adjustment and invariably, there will always be residual confounding present. However, inclusion of age, sex, baseline cardiovascular disease and two of the four aforementioned cardiovascular risk factors in regression models is a reasonable expectation for any well conducted study. It is also noteworthy that our criteria were more stringent than the Newcastle Ottawa Scale which only requires adjustment for two confounders in total.

We concede that we could have required adjustment for a detailed set of outcome specific risk factors (e.g. adjusting for race as a risk factor for CKD). This is a limitation of our study and we have added a sentence to acknowledge it in the limitations section of our study (page 19):

“Third, studies were classified as well adjusted if they adjusted for age, sex, baseline cardiovascular disease and at least two cardiovascular risk factors. Even though more stringent criteria may be used, such as requiring studies to adjust for medications and outcome specific risk factors (e.g. adjustment for race for the outcome of CKD), there will nonetheless be residual confounding in the estimates derived from observational studies. Accordingly, it is likely that there are other variables that contribute to the association between AF and our outcomes of interest, in addition to any possible causal disease-specific effects.”

- 6. Page 6, line 19: should “adequately adjusted” also include adjustment for certain treatments, depending on the outcome of interest? For example, in recent studies many AF patients will be on oral anticoagulation, which will lower the relative risk of a stroke.**

There is no accepted standard for what would be considered adequate multivariable adjustment and we acknowledge that a more stringent standard may have been used. We have therefore added the following sentence to our limitations (page 19):

“Third, studies were classified as well adjusted if they adjusted for age, sex, baseline cardiovascular disease and at least two cardiovascular risk factors. Even though more stringent criteria may be used, such as requiring studies to adjust for medications and outcome specific risk factors (e.g. adjustment for race for the outcome of CKD), there will nonetheless be residual confounding in the estimates derived from observational studies. Accordingly, it is likely that there are other variables that contribute to the association between AF and our outcomes of interest, in addition to any possible causal disease-specific effects.”

- 7. Page 8, line 6: clarify which statistical tests were used to test for trend.**

We used a chi-squared test for trend. We have added the following sentence to our methods section:

“We tested for trend by these characteristics across studies using random effects meta-regression.”

- 8. Page 8, line 7: clarify why AF type and % oral anticoagulation were not used for stratified analysis, as both had >9 studies.**

We have now included an analysis stratifying studies by the proportion of adults receiving oral anticoagulation (Supplemental Table 7-12). Although there were >9 total studies that reported on AF type, there were <9 studies reporting this information for any single outcome of interest. This did not meet our criteria for sensitivity analysis. We have clarified this in our methods (page 9):

“There were fewer than 9 studies that reported on AF type for any given outcome of interest (Supplementary Table 3).”

- 9. Page 14, Discussion: the paper could benefit from more extensive discussion of the results. Two examples: I) relative risks found in general population studies were typically larger than in specific settings. This could be related to the selection of the Controls and the proportion of AF patients treated with effective evidence-based therapies. II) the potential publication bias for reporting stroke incidence, and the lower relative risk after removing the studies that were causing most of the heterogeneity, are important findings. Many scientific papers and funding proposals are based on the same presumptions regarding the risk for stroke in AF, typically referencing the Framingham study and other well-known prospective cohort studies. However, your finding might indicate that this risk is actually lower, or has become lower in**

recent studies due to higher rates of oral anticoagulation use. Can you explore the latter and comment more specifically on this issues in the Discussion?

We agree with the reviewer that there are consistent differences between estimates in general population studies and specific population studies. However, we believe that the differences should be interpreted with caution due to the multiplicity of tests. We have added the following sentence (page 18):

“We observed that the association of AF with cardiovascular disease and death was generally consistent, irrespective of baseline history of ischemic heart disease, baseline history of stroke, mean participant age and baseline risk. However, there were two notable exceptions. First, relative risk estimates for general population studies were typically larger than estimates based on studies in specific settings. This could be related to the selection of the controls and the proportion of AF patients treated with effective evidence-based therapies.”

With respect to the change in relative risk estimates after the trim and fill procedure, this finding reflects the potential for publication bias in studies examining the association between AF and stroke. However, we respectfully disagree with the reviewer and do not think this can be interpreted as improvements in stroke care over time. For instance, there was no trend in relative risk estimates for all strokes and ischemic strokes based on the year of publication.

10. Page 17: if AF is simply a marker and not a cause of these outcomes, many risk factors are at play. However, hypertension deserves special attention as typically 70-80% of AF patients have hypertension, and it is also an important risk factor for all reported outcomes. Please add this consideration to the Discussion.

We have modified the following sentence in our discussion (page 20):

“Considering our observation that AF is also associated with an increased risk of heart failure, sudden cardiac death and chronic kidney disease (in addition to ischemic heart disease), it appears likely that AF may be acting as a marker for shared underlying risk factors for cardiovascular disease. These include, hypertension, which is diagnosed in up to 90% of patients with AF, as well as obesity, diabetes and obstructive sleep apnea.[125, 126]”

11. Page 17: if AF is a marker, but has significant associations with all of these outcomes, this might also indicate that the selected observational studies are inadequately correcting for the real causes. Please discuss as a methodological issue, or limitation.

We are in agreement with the reviewer regarding this issue and have added the following text to our manuscript (page 19):

“Third, studies were classified as well adjusted if they adjusted for age, sex, baseline cardiovascular disease and at least two cardiovascular risk factors. Even though more stringent criteria may be used, such as requiring studies to adjust for medications and outcome specific risk factors (e.g. adjustment for

race for the outcome of CKD), there will nonetheless be residual confounding in the estimates derived from observational studies. Accordingly, it is likely that there are other variables that contribute to the association between AF and our outcomes of interest, in addition to any possible causal disease-specific effects.”

- 12. Page 17: if AF has a causal relation with the outcomes, treatment of AF should lower their incidence, please discuss whether there is any evidence for this. Previous rhythm vs. rate control trials (AFFIRM, RACE, others) have addressed this issue, and with the rhythm control options they had available there was no benefit over rate control. These studies also assessed other outcomes than stroke. These results meant that either the rhythm control treatments were not very effective at treating AF, or that AF is not the cause of the non-stroke outcomes. You can include more recent evidence to discuss this.**

We have expanded our paragraph discussing treatment options for adults with AF as follows (page 21):

“Finally, our study may have implications for the prioritisation of public health resources and the development of novel interventions for adults with AF. First, the development and testing of novel oral anticoagulants has been the principal focus of clinical care in AF but recent studies have shown that these medications reduce stroke related mortality, with little benefit for reducing CHF and SCD related mortality.[123] Similarly, pharmacologic termination of AF and the restoration of sinus rhythm has shown no benefit over rate control for SCD[130], worsening heart failure[131] or for mortality[130]. It may be that rhythm control treatments may not have been effective at treating AF or that the treatments resulted in additional side effects that outweighed the benefits of restoring sinus rhythm. Alternatively, if AF is not the cause of these non-stroke outcomes, this may explain the absence of treatment benefit even when AF is terminated.”

- 13. Page 17: another important AF—specific scenario for illustration: the new oral anticoagulants are at least as effective at preventing stroke and cause fewer bleedings than warfarin. However, warfarin might prevent some more myocardial infarctions, one of your non-stroke outcomes.**

Consistent with comment 12 by the reviewer, we have expanded our paragraph discussing treatment options for adults with AF as follows (page 21):

“Finally, our study may have implications for the prioritisation of public health resources and the development of novel interventions for adults with AF. First, the development and testing of novel oral anticoagulants has been the principal focus of clinical care in AF but recent studies have shown that these medications reduce stroke related mortality, with little benefit for reducing CHF and SCD related mortality.[123] Similarly, pharmacologic termination of AF and the restoration of sinus rhythm has shown no benefit over rate control for SCD[130], worsening heart failure[131] or for mortality[130]. It may be that rhythm control treatments may not have been effective at treating AF or that the treatments resulted in additional side effects that outweighed the benefits of restoring sinus rhythm. Alternatively, if AF is not the cause of these non-stroke outcomes, this may explain the absence of treatment benefit even when AF is terminated.”

- 14. Page 17: a next research step could be to use your absolute risk increases for AF patients to determine what types of interventions or strategies need to be tested next to further reduce their burden of disease, please comment.**

As our study is a meta-analysis of observational studies, a causal relationship between AF and our outcomes of interest cannot be concluded. Accordingly, it would be misleading to apply our estimates to stratify treatment strategies.

Reviewer 3

- 1. The majority of the results are contained in the very extensive Supplementary Tables (9 Tables) and Figures (23 Figures) and I feel that this impedes the readability and understanding of the results. I'm not sure how this can be rectified though or how many Supplementary Tables and Figures the journal allows.**

We have asked for editorial guidance regarding which additional tables should be added to the manuscript to improve its readability. Based on editorial comments, we have moved the flow diagram into the main manuscript and eliminated funnel plots with less than ten studies.

- 2. How were the outcomes of congestive heart failure, peripheral arterial disease and CKD defined in the studies?**

We have compiled all outcome definitions in Supplementary Table 5. We have also provided a discussion of how the outcome definitions may have affected relative risk estimates in the discussion section of our manuscript:

"For instance, the criteria in well-established prospective cohort or secondary analyses of randomized trials [8] were comprehensively detailed whereas the criteria and specific diagnostic codes used in analyses of administrative data were not consistently provided [51]. Differences in the diagnostic codes used in administrative data may have contributed to the high heterogeneity in our study."

- 3. Please provide a clear rationale for why American Heart Association estimates of the incidence of cardiovascular mortality, IHD, HF, SCD were used, and Centers for Disease Control and Prevention estimates were used to calculate the absolute risk increase for each vascular outcome?**

The United States was the individual country that accounted for the largest share of studies. As absolute risk increases depend on the population specific incidence of the outcome of interest, we elected to use American estimates. We have clarified this in the manuscript.

We have also added a sentence to our methods information regarding how the absolute risk increases were derived and these estimates can be readily calculated by readers for other countries:

“Specifically, $ARI = (RR - 1) \times (ACR)$ where ARI is the absolute risk increase, RR is the relative risk and ACR is the assumed control risk.”

4. MOOSE checklist should be included.

We have included the MOOSE checklist as a supplementary appendix.

5. Since the Table of characteristics of the included studies is not included in the main manuscript it would be helpful to add a short paragraph summarising these in the Results section.

We have added a few sentences summarizing the table of characteristics of included studies to the first paragraph of our result section:

“Accordingly, 104 studies involving 9,686,513 patients were included in this meta-analysis. Of these individuals, 587,867 (6.1%) had AF. The method of AF ascertainment was not specified in five studies. Although outcome definitions were broadly consistent, the criteria applied in large prospective studies and secondary analyses of randomized trials were often more detailed than the criteria applied in studies that were strictly reliant on administrative datasets. Finally, sixty-seven studies were at low risk of bias based on the Newcastle Ottawa Scale and sixty-nine studies provided estimates from well adjusted regression models.”

For each outcome, we have also reported the prevalence of AF as well as the median follow up.

6. Please explain what Figure 1 represents in the text in the Results section.

In order to preserve the readability of the results section, we have instead adjusted the caption for the figure (now figure 2):

“Summary relative risks for each outcome examined in our study are provided.”

7. Page 10, lines 18-19 “...relative risk of major...” does not make sense. Please clarify.

We have clarified this sentence and it now reads as follows (page 12):

“In subgroup analyses, the relative risk of cardiovascular mortality was lower with increasing age ($p=0.052$, Supplementary Figure 3) and the relative risk of major cardiovascular events declined when studies were stratified based on the absolute event rate in their control group ($p=0.027$, Supplementary Figure 4).”

8. In Figure 3, when stratified by age, the age groups are overlapping, 64 and 71 appear in 2 groups. It is not clear to me what the other figures in square brackets on the left hand side for CHD and stroke represent.

We have clarified the caption for the figures to demonstrate that they are non overlapping:

“A square bracket indicates that the range is inclusive of the number within the bracket and a parenthesis indicates that the range is exclusive of the number within the bracket.”

- 9. Confidence intervals should be written in the format (1.36 to 2.60) within parentheses, using the word "to" rather than as (1.36, 2.60).**

We have made this change.

- 10. The proportion of patients on anticoagulation will likely affect all the outcomes of interest (particularly all-cause mortality, stroke (ischemic and hemorrhagic). The implications of this on the findings should be considered further and discussed.**

We have added an analysis stratifying studies by the proportion of adults receiving anticoagulation. There were sufficient studies for the following outcomes: all-cause mortality, all stroke and ischemic stroke. We have also discussed the implications of our findings in the discussion section of our manuscript:

“Second, the relative risk of all-cause mortality was lower in studies with a higher proportion of participants receiving anticoagulation, but we did not observe a similar pattern for stroke. The absence of any association between the relative risk of stroke and the use of anticoagulation should be interpreted with caution because of the small number of studies included in this analysis. Furthermore, if warfarin was used, it is difficult to determine the effectiveness of anticoagulation without INR measurements. Differences in the proportion of adults with a low INR would affect our stratified analyses for stroke.”

- 11. A reference(s) to a clinical guideline(s) would be more appropriate than reference 5 in the Introduction. It would be better if you did not reference just one of the NOAC clinical trials (currently reference 4) but included a reference(s) which summarised the findings of VKA trials (e.g., Hart et al, 2007) and NOAC trials (e.g., Ruff et al, 2014); there are many others to choose from.**

We have adjusted the references accordingly.

- 12. The references in the text should be given as numbers in square brackets or superscript throughout not a mixture of the two.**

We have corrected the references in the manuscript.

- 13. The references from 19 onwards are missing from the main reference list.**

We have corrected the references in the manuscript.

Reviewer 4

- 1. There is no reference in the entire manuscript to risk prediction of stroke/thromboembolism which is the focus of management in AF using scores such as CHADS2 or CHA2DS2-VASc. The derivation and validation studies for these scores (which are widely used) suggest different absolute stroke rates based on baseline characteristics, which seems to be in conflict with the conclusion of this study. What may have caused this difference?**

The CHADS2 and CHA2DS2-Vasc scores were designed for risk prediction among adults with AF and demonstrate that baseline comorbidities are associated with higher risk of stroke among adults with AF. However, we meta-analysed studies that compared adults with AF to those without AF to examine a separate question – whether AF has a similar proportional effect across baseline comorbidities. Our observation of similar proportional effects across baseline characteristics is not in conflict with the observation of increased risk with baseline comorbidities.

As an example of how this can be true, see the table below. In this table, AF is associated with two times the risk of stroke regardless of whether someone is >65 or <65 (our finding) and individuals over 65 are at two times the risk of stroke than individuals under 65 (CHADS finding).

Hypothetical risk of stroke by the presence of AF and age.

	AF Present	AF Not present
<65	2X	1X
>65	4X	2X

- 2. What is the implication for future risk prediction in AF?**

We have added the following paragraph to our discussion:

“Despite uncertainty about whether AF causes the non-stroke outcomes in our study, there is merit to developing clinical risk prediction models for outcomes such as CHF; particularly given the relative and absolute risk estimates obtained in our study. To date, three models have been developed with c-statistics ranging from 0.7 to 0.84 but none have been externally validated. Future models may also benefit from including non-invasive measures of cardiac function and assessments of novel biomarkers.”

- 3. The background section in the abstract states "a wide range of cardiovascular outcomes". Renal disease is not classically included in CVD. Why is it included here?**

Although chronic kidney disease (CKD) is not classically considered a cardiovascular disease, it is an important source of morbidity. Furthermore, the association between AF and CKD may be overlooked

by clinicians. Given that the aim of our study was to summarize the association of AF with outcomes beyond stroke, we consider the inclusion of CKD as appropriate.

To more accurately reflect the objectives of our study, we have adjusted the background section of the abstract to now read as follows:

“The relationship between atrial fibrillation (AF) and the development of cardiovascular and renal disease is unclear. We aimed to quantify the associations between AF and cardiovascular disease, renal disease and death.”

- 4. Why are Congestive Heart Failure and Chronic Kidney Disease together in the results section? Please separate these two unrelated categories.**

We have made this change.

Reviewer 5

- 1. The discussion of the implications for clinicians is excellent in that it addresses identifying the possibility of AF as a prognostic marker for other conditions, and the importance of primary prevention and risk management. I could ask for strengthening these statements by including clear recommendations for how to communicate this risk information to patients and how to follow up - regular follow-up visits and testing protocols come to mind - but I really think clinicians are going to understand the implications for their patients regardless.**

We are in agreement with the reviewer and have added the following sentence to our manuscript about how to communicate cardiovascular risk (page 21):

“Nonetheless, reducing the burden of non-stroke events in adults with AF would benefit from a focus on primary prevention and cardiovascular risk factor management. Evidence based strategies in this regard include discussing the concept of predicted cardiovascular risk with patients and calculating their cardiovascular age – which is the increase in life expectancy associated with reduction of cardiovascular risk factors. Regular updates should also be provided to patients after lifestyle changes and/or pharmacotherapy have begun as a way to encourage further progress.”

Reviewer 6

- 1. The prevalence of atrial fibrillation amongst the participants was 14.8% overall and varied substantially between the analyses for different endpoints (5.4% to 15.5%). It would be useful to illustrate the considerable variation in the prevalence of atrial fibrillation between individual studies (currently only available as numbers in the on-line supplement).**

We have added this information for each outcome.

2. **The heterogeneity might partly reflect variation in endpoint definitions between the studies and over time (e.g. impact of troponin on diagnosis of myocardial infarction) but the authors provide little information about this aspect of their systematic review. Did the authors document variation in endpoint definitions and can they comment on the extent to which this may have contributed to heterogeneity between studies?**

We have compiled all outcome definitions in Supplementary Table 5. We have also discussed how outcome definitions may have affected relative risk estimates (page 19):

“For instance, the criteria in well-established prospective cohort or secondary analyses of randomized trials [8] were comprehensively detailed whereas the criteria and specific diagnostic codes used in analyses of administrative data were not consistently provided [51]. Differences in the diagnostic codes used in administrative data may have contributed to the high heterogeneity in our study.”

3. **The authors used American estimates for the incidence of cardiovascular mortality and other cardiovascular events to translate relative risks into estimates of absolute risk. Can the authors comment on the extent to which this might limit generalisability of the absolute risk estimates to other populations?**

The United States was the individual country that accounted for the largest share of studies. As absolute risk increases depend on the population specific incidence of the outcome of interest, we decided to use American estimates. We have clarified this in the manuscript.

We have also added a sentence to our methods information regarding how the absolute risk increases were derived and these estimates can be readily calculated by readers for other countries:

“Specifically, $ARI = (RR - 1) \times (ACR)$ where ARI is the absolute risk increase, RR is the relative risk and ACR is the assumed control risk (ACR).”

4. **Some of the analyses suggest that atrial fibrillation is associated with very large increases in the risk of adverse cardiovascular outcomes, but these risks likely reflect the particular populations under study. For example, only 6 of the 100 studies reported on the association between atrial fibrillation and heart failure (RR 4.99). Can the authors comment on the extent to which their findings are generalizable to a wider unselected population of patients with atrial fibrillation, particularly as sensitivity analyses were not done for some outcomes?**

In the case of heart failure, 4 of the 6 studies were conducted in general population cohorts and accounted for 75% of weight for the summary relative risk. Only the studies by Smit et al. and by Ruel et al. were conducted in specific populations. We agree that this clarification is important and we have added the following sentence to our manuscript (page 20):

“Fourth, due to our strict selection criteria, we identified less than nine studies for some outcomes and were unable to conduct sensitivity analyses. This is particularly important for CHF because the relative an absolute risk estimates for incident CHF was the highest among the outcomes we examined. It is therefore noteworthy that 4 of the 6 studies included in the meta-analysis for incident CHF were conducted in general population cohorts and accounted for 75% of the weight for the summary relative risk.”

Furthermore, our study also noted that estimates from general population studies were frequently larger than estimates from specific populations, which may reflect differences in the selection of controls or the use of evidence based therapies. It is therefore unlikely that studies conducted in specific populations has inflated the relative risk estimates in our study.

- 5. The authors conclude that additional interventions to reduce cardiovascular risk (beyond the use of anticoagulation to reduce the risk of stroke) are warranted in patients with atrial fibrillation, but also correctly state that they have only demonstrated a statistical association between atrial fibrillation and adverse cardiovascular outcomes. Can the authors comment on the extent to which the increased risks of adverse outcomes in patients with atrial fibrillation are likely to be modifiable?**

We have now expanded our discussion related to the limited benefit derived from anticoagulation and restoration of sinus rhythm for non-stroke outcomes. Accordingly, we believe that reducing the burden of non-stroke events in adults with AF would benefit from a focus on primary prevention. We have have elaborated on this concept in our manuscript as follows (page 21):

“Nonetheless, reducing the burden of non-stroke events in adults with AF would benefit from a focus on primary prevention and cardiovascular risk factor management. Evidence based strategies in this regard include discussing the concept of predicted cardiovascular risk with patients and calculating their cardiovascular age. Regular updates should also be provided to patients after lifestyle changes and/or pharmacotherapy have begun as a way to encourage further progress.”

- 6. The reference notation in the paper appears inconsistent. For example, in line 7 on page 4, the authors refer to reference 5 (super script) and reference 4 (in brackets). The text suggests that reference 5 relates to a guideline, but reference 5 in the main reference list is a research paper. Please can the authors review and correct the referencing in the paper.**

We have corrected the references in the manuscript.

- 7. The abbreviation CHF (page 6, line 7) is not defined until line 11.**

We have defined CHF after its first use in the manuscript.