Author’s response to reviews

Title: Optimal INR Level for Warfarin Therapy after Mechanical Mitral Valve Replacement

Authors:

Itthidet Kamthornthanakarn (dr.itthi@hotmail.com)
Rungroj Krittayaphong (rungroj.kri@mahidol.ac.th)

Version: 1 Date: 14 Feb 2019

Author’s response to reviews:

Response to comments from reviewers and editors
BCAR-D-18-00341
Optimal INR Level for Warfarin Therapy after Prosthetic Mitral Valve Replacement
Itthidet Kamthornthanakarn, MD; Rungroj Krittayaphong, M.D.
BMC Cardiovascular Disorders

Editor Comments:
1) Please include the heading "Declarations" at the start of this section (after the abbreviations)

Response:
We added heading “Declarations” after the List of abbreviations (page 11 line 19)

2) Under the heading “Ethics approval and consent to participate” in the Declarations, please include information on the consent to participate. If consent to participate was waived by the ethics committee, please state this here.

Response:
We added a statement ‘Consent to participate is not needed for the retrospective study design.’(page 4 line 15-18 and page 11 line 20-22)

BMC Cardiovascular Disorders operates a policy of open peer review, which means that you will be able to see the names of the reviewers who provided the reports via the online peer review system. We encourage you to also view the reports there, via the action links on the left-hand side of the page, to see the names of the reviewers.

Reviewer reports:

Miriam Silaschi, MD (Reviewer 1): Why should INR target be different for Thai population then for other ethnic groups? Please elaborate in the introduction
Response:
We added the following statements in the background section “Previous data has shown that Asian population with atrial fibrillation who received warfarin has an increased risk of intracranial hemorrhage up to 4 times compared to Caucasians. Asian population had a greater proportion of intracerebral bleeding as a stroke subtype when compared to Caucasians. There is no clear explanation for the increased risk of intracerebral bleeding in Asian population.” (page 3 line 13-17)

- Check for grammar and language

Response:
We checked and corrected the grammar and language throughout the manuscript.

- Did you use different prosthesis types? Which type of prosthesis was used? Please clarify further

Response:
We have only 1 type of prosthesis which is mechanical bileaflet valve. This has been stated in the first paragraph of the results section. (page 7 line 9) We, therefore, changed the wording from ‘prosthetic valve’ to ‘mechanical valve’. Thank you for the suggestion.

- Left atrial appendage closure during surgery?

Response:
Left atrial appendage closure was performed during surgery in 19 patients (9.5%). (page 7 line 9-10)

- Paroxysmal Afib: Did these patients have a difference in incidence in thromboembolic events

Response:
Among patients with atrial fibrillation, only 4 (2.7%) were paroxysmal atrial fibrillation in nature. The number was too small to make any comparison for the incidence of thromboembolic events. (page 7 line 11-13)

- How do you explain the relatively high event rates for ischemic stroke and bleeding (5-9%) in your study and is it comparable to other studies?

Response:
The higher rate of ischemic stroke may be related to a suboptimal TTR control in our population compared to other studies. Doctors may be fear of bleeding and keep a relatively low INR levels. It has been reported from the GARFIELD AF registry that TTR of Asian population is lower than Caucasians. But despite the lower TTR, Asian population had a higher rate of major bleeding and intracerebral hemorrhage compared to Caucasians. The higher ischemic stroke rate in our study as compared to study from China may be due to the study population. We enrolled patients with mechanical mitral valve prosthesis whereas study from China enrolled both mitral and aortic valve prosthesis. Patients with aortic valve prosthesis had a lower thrombotic complication and required a lower INR target compared to those with mitral valve prosthesis. This has been added in the discussion section. (page 9 line 25 to page 10 line 8)

- How many patients were lost to follow-up / no information available on bleeding events and how did you account for dead patients, especially those were the cause of death was unknown? Please
elaborate further

Response:
As we stated in the exclusion criteria that patients who did not have follow-up data were excluded. So we don’t have data on how many patient were lost to follow-up. Our hospital is a tertiary-care hospital and some patients were referred to our hospital for surgery and they were follow-up at the local hospitals. It is difficult to track their data on the ischemic or bleeding events.

Regarding the dead patients, when the death were related to ischemic stroke we count as ischemic stroke, when it was related to severe bleeding, we counted as clinically significant bleeding. But when the death is not related to ischemic stroke or bleeding, we did not count as an outcome. We added this statement in the methods section.(page 6 line 15-17) There were 4 deaths; one from ischemic stroke, one from intracerebral hemorrhage, and 2 from septic shock. There was no unknown cause of death.

- Do you conclude that Thai patients have lower INR target than Caucasians for MVR? Is there an explanation for this observation?

Response:
Yes, we concluded that Thai (and may be Asians) required a lower INR target than Caucasians for patients with mechanical mitral valve. The explanation for the lower INR target may be related to a higher bleeding event and intracerebral hemorrhage in Asian population as compared to Caucasians. When we consider the INR level that is most suitable for the balance of risk and benefit ratio towards a lower target than the general recommendation. This has been added in the discussion section.(page 10 line 13-16)

- Were there any fatal bleeding/thrombembolic events?

Response:
Yes, there was one patient who died from intracerebral hemorrhage and another patient who died from a large cerebral infarction.(page 7 line 20-22)

- How often were INR levels controlled? Did all patients perform self-check on a daily basis?

Response:
Based on the INR recommendation at 2.5-3.5 in this group of patient, the frequency of controlled INR were 40.4±16.3%. This has been added in page 7 line 26 and page 8 line 1.
No, all patients did not have self-check for INR at home.

- What is „new“ about your study/elaborate in the conclusion as to why your study might add further knowledge to the scientific community

Response:
As we stated in the conclusion, we suggested a slightly lower INR target in Thai population which may be applied in Asian population. This suggestion should allow more room for INR adjustment which could avoid bleeding event in patients with mechanical mitral valve without losing the benefit of warfarin. We added this statement in the conclusion.(page 11 line 9-13)

Nathan Clark (Reviewer 2): This is a retrospective study evaluating adult Thai patients having
prosthetic mitral valve replacement between 2011 and 2015 to ascertain the optimal INR value for warfarin therapy. Overall, the manuscript is well written and study activities were appropriately reviewed by Institutional Review Board at the Siriraj Hospital.

Major considerations -
The term prosthetic could refer to either mechanical or tissue valve replacement. Table 1 indicates all valves are mechanical, bi-leaflet valves. I would suggest referring to the valves as "mechanical" throughout the paper for clarity and including the valve type in the first paragraph of the results.

Response:
We have changed the word ‘prosthetic’ to ‘mechanical’ as suggested. Valve type is already stated in the first paragraph of the results section.

The approach to determining the amount of time spent in each INR group is not the Rosendaal method. Rosendaal used linear interpolation where the INR travels in a straight line between points. The process of dividing the duration between INRs in half and attributing half the interval to each INR is different.

Response:
The original Rosendaal’s paper (reference number 15), they described that ‘they assumed that the INR values between 2 measurements will vary linearly from the first to second measurement. Based on this assumption, the simplest approximation is to divide the time between 2 measurements into halves, and allocate the first half to INR value of the first, and the second half to INR value of the second measurement.’ Tis is shown in the figure below. There might be some modification in the methods later on. But the method that we used is based on original description of Rosendaal.

Could you explain the rationale for excluding patients with a history of bleeding or thromboembolic events? As this is a high risk population, many patients will survive such events and having information about optimal INR goals for them would have been useful.

Response:
We excluded patients with history of thromboembolic event and bleeding due to some reasons. For patients with history of stroke, most of them had some degree of residual neurological deficit. This could lead to problem of the documentation of recurrent stroke. For patients with history of bleeding, they could have some underlying disease that prone to recurrent bleeding even without warfarin. Moreover, many patients with history of stroke of bleeding may be a referred case from other hospital and may be difficult for the investigators to collect the data on the time relation of anticoagulant use and the onset of stroke or bleeding. However, we also agreed with the comments from the reviewer that this group of patients is a vulnerable group and may be of interest. So we added a statement related to this issue under the limitation of the study. (page 10 line 23 to page 11 line 5)

The definition of major bleeding does not follow that provided by the International Society of Thrombosis and Haemostasis. Specifically, classifying patients with GI bleeding or hemoptysis that required hospital admission for observation only without the need for blood transfusion or with a corresponding drop in hemoglobin of at least 2 g/dl could have led to overestimation in major bleeding rates compared to ISTH.
Response:
We set our criteria for major bleeding from partly from the International Society of Thrombosis and Haemostasis (ISTH) and we think that bleeding required medical attention is more than simple minor bleeding and should be included in the criteria of major bleeding. The issue on the heterogeneity of the bleeding definition used in major cardiovascular trials has been discussed and Bleeding Academic Research Consortium (BARC) proposed a definition of bleeding and considered bleeding required medical attention as type 2 and separate from minor bleeding (type 1). The bleeding required medical attention has been included together with major bleeding as criteria for safety endpoint (for example in PIONEER AF study). We believed that bleeding required medical attention should be included. Since reviewer suggested that this may overestimate the rate of major bleeding, we therefore changed the wording from ‘major bleeding’ to ‘clinically significant bleeding’. We added this explanation in methods section. (page 6 line 11-14)

Research Square (Reviewer 3): "STATISTICAL REVIEWER ASSESSMENT:

Is the study design appropriate for the research question (considering whether the analyzed population accurately reflects the design and whether you see any problems with control/comparison groups, e.g., likely confounders)?
No - there are minor issues

Are methodologies adequate and well implemented (considering whether assumptions are addressed and whether analyses are robust)?
No - there are minor issues

Are the analyses adequately communicated (considering whether reporting details are adequate and whether figures and tables are well labeled and described)?
No - there are minor issues

Does the interpretation accurately reflect the analyses without overstatement (considering whether limitations/bias are acknowledged and whether accurate descriptors, e.g., 'significant', are used)?
No - there are minor issues

Could an appropriately REVISED version of this work represent a statistically sound contribution?
Probably - with minor revisions

STATISTICAL REVIEWER COMMENTS:

The findings of the study may help cardiovascular surgeons decide optimal INR range for warfarin dose optimization in patients undergoing prosthetic mitral valve replacement in the Thai population. Since warfarin is a narrow therapeutic index medicine and needs long-term therapy in patients undergoing prosthetic mitral valve replacement, findings of the present study would ultimately be beneficial for the target population.

The authors have studied the INR levels of 200 patients for 3.53±1.27 years with 707.81 patient-years. Application of patient-year concept is a good statistical and epidemiological interpretation in terms of descriptive statistic. This helped to depict the INR levels of the study population over a fairly long follow up period, i.e., 3.5 years (although it is not sufficiently long follow up period for patients taking warfarin). They have shown their results in frequency and percentage form (i.e., applying descriptive statistics) in an illustrative manner and also shown p values in figures 1A, 1B, 1C, 1D, 2 and 3.
The authors have used a chi-square test to compare the incidence density of thromboembolic and bleeding complications, which seems an underutilization of better parametric alternative because chi square test better gives the association between two or among more than two variables; comparison of proportions and goodness of fit. Since the authors have studied INR values of more than 3.5 years' follow-up data, repeated measure ANOVA (if data were normally distributed) or Friedman's test (if data were not normally distributed) would be better inferential statistic of choice. Authors are suggested to check normality of their dataset with Kolmogorov Smirnov test and follow the above-mentioned suggestion.

REQUESTED REVISIONS:

ADDITIONAL REQUESTS/SUGGESTIONS:

General comments:
* Use proper subject-verb agreement with the term 'data' as it is plural term.

Response:
We made correction as suggested throughout the manuscript.

* Use 'who' with persons instead of 'that'.

Response:
We made correction as suggested throughout the manuscript.

* Follow citation within text properly such as in 'Marieke Torn, et al., 2009'.

Response:
We made correction as suggested throughout the manuscript.

Specific comments:

Abstract:
Results: Rewrite the sentence "Two hundred patients were included and followed over a period of 707.81 patient-years." The period was 3.53±1.27 years whereas 707.81 was patient-years.

Response:
We made changes as suggested in the results section. (page 7 line 6)

Conclusion: It is not necessary to redefine 'optimal INR level' as it is already defined in Methods. This also applies to the Conclusion section in the main text.

Response:
We made changes as suggested in the conclusion of the abstract and conclusion at the end of the manuscript.

Main text:
Methods:
* Mention study design properly; just mentioning data extracted retrospectively is not sufficient. However, it seems that the authors followed the time-series study as they took the follow-up data of
200 patients till 707 patient-years (i.e., more than 3.5 years' follow-up data). This has a significant impact on the choice of statistical test.

Response:
We revised the wording in the study design of the methods section.(page 3 line 26 to page 4 line 2)

* Rewrite Methods section under various sub-headings such as study design, study area and site, study population and sample size calculation, ethics approval, study procedure, sampling technique, inclusion and exclusion criteria, data collection tool and data collection, reliability and validity, and statistical analysis although you have mentioned majority of sections (except data collection tool, reliability and validity and other) in the unstructured format.

Response:
We rewrite the methods using various subheadings as suggested.(page 3 line 26 to page 7 line 3)

* The authors have mentioned "Using a power of 90%, a type I error of 0.05, and a type II error of 0.10, a minimum sample size of 200 patients was calculated." Still, sample size calculation is not clear from the statement.

Response:
We calculated sample size based on the results of a previous publication by Marieke Torn, et al., 2009. They reported the incidence of thromboembolic and bleeding events in the <2, 2.0-2.4, 2.5-2.9, 3.0-3.4, 3.5-4.5, and >4.5 INR groups to be 0.319, 0.067, 0.02, 0.025, 0.033, and 0.247, respectively (modified from Marieke Torn, et al., 2009)[12]. We calculate sample size of each of the 6 INR groups by using NQuery 6.0 program (Statistical Solutions Ltd, Cork, Ireland) by the formula based on proportion and unequal N. The total sample size of all groups were 200 patients.(page 4 line 6-14)

* The authors have mentioned "The incidence density of thromboembolic and bleeding complications was compared using chi-square test." Chi square test better gives the association between two or among more than two variables; comparison of proportions and goodness of fit. Since the authors have studied INR values of more than 3.5 years' follow-up data (although the frequency of follow ups are not clear from the text and that needs further elaboration), repeated measure ANOVA (if data were normally distributed) or Friedman's test (if data were not normally distributed) would be better inferential statistic of choice. Chi square test seems an underutilization of better parametric alternative. Authors are suggested to check normality of their dataset with Kolmogorov Smirnov test and follow the above-mentioned suggestion.

Response:
The dataset were normally distributed based on Kolmogorov Smirnov test. Linear Mixed models (fixed effect) was use interpret the repeated measurement INR levels related outcome measures at time. (page 6 line 19 to page 7 line 3) Since the number of INR test were different among each patient, we do not use repeated measure ANOVA (repeated measure ANOVA may be suitable when the number of repeated data are almost similar). We added statements in the statistical analysis in the methods section and added the result in the results section and in Table 3.(page 8 line 3-7)

Results:
* The authors have not shown the dose of warfarin being studied in the study sample.
Response: The average dose of warfarin in the study sample was 25.6±9.5 mg per week. (page 7 line 25-26)

* They have also not shown the frequency of INR level monitoring and frequency of warfarin optimization based on that.

Response: Average frequency of INR monitoring was once every 3.1±1.0 month. (page 8 line 1-2) One-hundred ninety-one patients (91.5%) had warfarin dose optimization. The frequency of based on the INR test was 3.31±1.8 times during the study period. (page 8 line 2-3)

Discussion:

* Rewrite the sentence "The study population of these studies were slightly difference than ours since they enrolled both mitral and aortic valve replacement especially ARENA study which enrolled aortic valve replacement in the majority" clearly.

Response: We rewrote the sentence as follows: `The study population of our study were different from the aforementioned studies. They enrolled both mitral and aortic valve replacement but we enrolled only patients with mechanical mitral valve. Majority of patients enrolled in AREVA study were those with aortic valve replacement.` (page 9 line 15-18)

* The sentence "As a result, our study may have lacked sufficient power to identify all significant associations" mentioned as the limitation seems controversial with the previous statement "Using a power of 90%, a type I error of 0.05, and a type II error of 0.10, a minimum sample size of 200 patients was calculated" mentioned in Methods section. Justify how can the power of 90% be insufficient?

Response: We just felt that 200 patients are relatively small number of patients. However, this number should have enough statistical power to identify the objective of the study. We therefore deleted this statement.

* The authors have mentioned "Third, the patients enrolled in this study were from a single center." Are single-centric studies real limitations? Or are all multi-centric studies are free from such limitations?

Response: We deleted this statement. We think that the limitation of generalizability already covered what we intend to say.

Conclusion: Rewrite Conclusion section more specifically.

Response: We rewrote the conclusion statement. (page 11 line 9-13)

Figures 1A, 1B, 1C, 1D, 2 and 3: Show all p values in between each interval of INR.

Response: We redo all figures and showed p-values in between each interval of INR. We put NA (not applicable) when there is no event occurred in the two INR groups that were compared.