**PEER REVIEW HISTORY**

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form ([http://bmjopen.bmj.com/site/about/resources/checklist.pdf](http://bmjopen.bmj.com/site/about/resources/checklist.pdf)) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

**ARTICLE DETAILS**

| TITLE (PROVISIONAL) | Bi-atrial versus left atrial ablation for patients with rheumatic mitral valve disease and non-paroxysmal atrial fibrillation (ABLATION): Rationale, design and study protocol for a multicenter randomized controlled trial |
|---------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| AUTHORS             | Yu, Chunyu; Li, Haojie; Yang, Wang; Chen, Sipeng; Zhao, Yan; Zheng, Zhe |

**VERSION 1 – REVIEW**

| REVIEWER | Michail Ovcharov  
| FSBI National Medical Research Center named after E N Meshalkin |
| REVIEW RETURNED | 05-Jul-2022 |
| GENERAL COMMENTS | First of all I want to congratulate authors with well designed and planned study protocol. In my opinion, all aim are clarified, randomisation procedure declare on high level. Study endpoints are well marked. I wish authors achieve goals and discover new knowledge for all of us. |

| REVIEWER | James Edgerton  
| Barnes-Jewish Hospital, Cardiothoracic Surgery |
| REVIEW RETURNED | 12-Jul-2022 |
| GENERAL COMMENTS | To the Authors:  
This is a well written protocol from a respected high volume medical center with which I am well familiar. The senior investigator is an experienced and articulate scientist. The protocol seeks to define the relative contribution of the right atrial lesions to the efficacy of a Maze procedure. A secondary objective is to determine if performance of the right atrial lesions is associated with an increased rate of pacemaker implantation. Others have attempted to answer these questions and fallen short for reasons of inadequate study design, inconsistent surgical technique, or inadequate powering. This well thought out protocol attempts to avoid all these pitfalls. However, I do have just a few comments in each of these areas.  

Power Analysis:  
Dr Gillinov et al. attempted to answer a similar question [N Engl J Med 2015;372(15):1399-409]. They addressed whether pulmonary vein isolation alone was equally as efficacious as the full left atrial lesion set. They arrived at an erroneous conclusion yet stated in the paper that the study was inadequately powered to address this secondary outcome. It would be a terrible shame for this eloquent
study to fall short in the end because of inadequate powering. The power analysis is well described and is sound. However, it is dependent on the assumption that there will be a 15% difference (70% for left atrial lesions vs. 85% for bi-atrial lesions) in success between the two groups. What happens if this assumption is wrong, and the difference is 10% or 8%? I offer a word of caution to be absolutely certain that the study is adequately powered.

Study Design:
Your primary endpoint is "the survival rate without any recurrence of atrial tachyarrhythmias by means of 3-day continuous Holter monitoring at 6-month and 12-month follow-up." How will you handle patients who may die of noncardiac causes prior to twelve-month endpoint but were in sinus rhythm at 6-month assessment and any ECG done during intermittent office visits? Does your statistical methodology adequately address this?

Study participants will complete a questionnaire at 6 and 12 months. The protocol does not contain a copy of that questionnaire. Important elements would include any history of neurological events, peripheral thromboembolic events, and palpitations.

Regarding permanent pacemaker implantation, it will be critical that rhythm prior to implantation be documented and submitted, and that the reason for implantation and timing of it be recorded. It will also be critical to record the percent of time a patient is paced at both six- and twelve-month follow-up.

Methodology:
Rhythm follow-up is by 72-hour Holter, and this complies with the 2017 Consensus Statement requirement for scientific studies. The definition of failure also complies with the Consensus Statement. Regarding the 72-hour Holter monitors, the protocol states that "after wearing, they are sent back to the project team for data analysis." Does this mean that a core lab will be used to analyze the monitors? Will the persons doing the analysis be blinded as to the randomization? Similarly, will any intermittent 12 lead ECGs be submitted to a core lab? Further, will the 12-month echocardiograms be interpreted by a blinded core lab?

A major failing of the previous randomized trial [N Engl J Med 2015;372(15):1399-409] was that the surgical technique was not controlled, and multiple different lesion sets were performed and grouped together. You have taken the extraordinary step of insisting that surgical technique be standardized and verifying this by requiring any enrolling surgeon to submit a video of his surgical technique. I applaud you for this. As a minor concern I note that this requirement is stated in the Introduction and the Discussion but not in the Methodology section.

The specified surgical technique requires that each bipolar lesion is performed three times. Dr Damiano had demonstrated that transmurality of a lesion is enhanced by applying energy to the bipolar clamp twice without releasing the clamp. You may wish to adopt this technique in the protocol.

I look forward with eager anticipation to the presentation of your results of this landmark study.
This is an RCT aiming to evaluate the efficacy of bi-atrial ablation for patients with RMVD and non-paroxysmal AF, which have begun recruitment 5 months ago. My major concerns are listed as below:

1. It is unusual to observe the practice adopted for DMC in this study. DMC is an independent group who will assess the safety risk from the treatment. This means they cannot be blind and treatment allocation should be shared within the DMC members in the closed meeting. Meanwhile, DMC members meet regularly (eg, half a year). What do you mean urgent review in line 304? The authors may need carefully review their DMC charter. Some references from DAMOCLES study group (Lancet. 2005 Feb 19-25;365(9460):711-22) may be helpful.

2. It is confused to me to learn the term of “survival rate” while describing the trial endpoint. What does this actually refer to? Is it incidence? It seems that “survival rate” is estimated from a KM curve. Please clarify and use associated proper word. In addition, effectiveness is repeatedly used in this trial, but this should be different from efficacy. From the trial design, it may not be suitable to use effectiveness.

3. The sample size is all calculated based on one-side 0.05, which means two sides 0.1. This is not common for a confirmative trial. Detailed explanation is needed, which is better for the readers.

4. A Gantt plot would be needed to facilitate reading.

Response to reviewer #1
Dr. Michail Ovcharov, FSBI National Medical Research Center named after E N Meshalkin
Comments to the Author:
First of all I want to congratulate authors with well designed and planned study protocol. In my opinion, all aim are clarified, randomisation procedure declare on high level. Study endpoints are well marked. I wish authors achieve goals and discover new knowledge for all of us.

Answer: We are very grateful to the reviewer for the time and effort in our manuscript. Thank you very much.

Response to reviewer #2
Dr. James Edgerton, Barnes-Jewish Hospital
Comments to the Author:
To the Authors:

This is a well written protocol from a respected high volume medical center with which I am well familiar. The senior investigator is an experienced and articulate scientist. The protocol seeks to define the relative contribution of the right atrial lesions to the efficacy of a Maze procedure. A secondary objective is to determine if performance of the right atrial lesions is associated with an increased rate of pacemaker implantation. Others have attempted to answer these questions and fallen short for reasons of inadequate study design, inconsistent surgical technique, or inadequate powering. This well thought out protocol attempts to avoid all these pitfalls. However, I do have just a few comments in each of these areas.
Power Analysis:
Dr. Gillinov et al. attempted to answer a similar question [N Engl J Med 2015;372(15):1399-409]. They addressed whether pulmonary vein isolation alone was equally efficacious as the full left atrial lesion set. They arrived at an erroneous conclusion yet stated in the paper that the study was inadequately powered to address this secondary outcome. It would be a terrible shame for this eloquent study to fall short in the end because of inadequate powering. The power analysis is well described and is sound. However, it is dependent on the assumption that there will be a 15% difference (70% for left atrial lesions vs. 85% for bi-atrial lesions) in success between the two groups. What happens if this assumption is wrong, and the difference is 10% or 8%? I offer a word of caution to be absolutely certain that the study is adequately powered.

Answer 1: Dr. Professor Edgerton, we appreciate your time and effort that dedicated to providing feedback on our manuscript. Thank you very much for your valuable suggestions, which will definitely improve our manuscript a lot!

As for the 15% difference in success between the two groups, in fact, the ablation success rates for both groups referred to data from our center. According to the data from our center, the 1-year ablation success rate in the bi-atrial ablation group exceeds 90%, and the 1-year ablation success rate in the left atrial ablation group is about 60%. Therefore, if based on these two values, there will be a 30% difference, and this study requires a much smaller sample size. To avoid bias in the single-center data, we also referred to outcomes from other centers, as the references in the manuscript.

Therefore, in order to have a greater power, we finally chose the 15% difference (70% for left atrial lesions vs. 85% for bi-atrial lesions).

During data analysis, if the actual difference is less than 15%, such as 8% or lower, then this study is likely to yield negative results. In fact, if the difference is only 8% or less, we can test for a statistical difference by increasing the sample size. Even if the difference is just 1%, we can get the difference by increasing the sample size. However, in a clinical sense, such a low difference between the two is, in fact, not clinically significant. Therefore, there is also no need for us to conduct clinical studies with such a huge sample size. James L Cox demonstrated that after strictly left-sided ablation procedures, approximately 10% of patients will have postoperative atrial flutter originating from the right atrium, and if the right atrium is enlarged or stretched as a result of the left heart problem, it may become large enough to support two simultaneous macro-reentrant circuits (doi:10.1016/j.jtcvs.2010.02.027). Therefore, adding right atrial ablation might improve the efficacy by at least 10%. We believe that a 15% difference in our study is clinically meaningful and the sample size of 320 is feasible in practical terms.

Change 1: None

Study Design:
Your primary endpoint is "the survival rate without any recurrence of atrial tachyarrhythmias by means of 3-day continuous Holter monitoring at 6-month and 12-month follow-up." How will you handle patients who may die of noncardiac causes prior to twelve-month endpoint but were in sinus rhythm at 6-month assessment and any ECG done during intermittent office visits? Does your statistical methodology adequately address this?

Answer 2: Thank you very much for raising this issue! In fact, patients do die during follow-up, which equates to a lost follow-up in this study. However, with only 1 year of follow-up in this study, there will not be many missed visits or deaths. We considered a withdrawal rate of 10%, which will cover subjects who die and are lost to follow-up. In addition, we will perform sensitivity analysis during data analysis to verify the reliability of the results by filling in missing values.
Study participants will complete a questionnaire at 6 and 12 months. The protocol does not contain a copy of that questionnaire. Important elements would include any history of neurological events, peripheral thromboembolic events, and palpitations.

Answer 3: Thanks for your comment. This is a very important issue. The questionnaire includes questions on subject survival status, cardiac function classification, stroke, peripheral thromboembolic events, hospitalization for heart failure, bleeding events, medication use, and permanent pacemaker implantation. The questionnaire has been carefully described in the revised manuscript for better understanding by the readers.

Change 3: Page 13, lines 306-308

“The questionnaire includes questions on subject survival status, cardiac function classification, stroke, peripheral thromboembolic events, hospitalization for heart failure, bleeding events, medication use, and permanent pacemaker implantation.”

Regarding permanent pacemaker implantation, it will be critical that rhythm prior to implantation be documented and submitted, and that the reason for implantation and timing of it be recorded. It will also be critical to record the percent of time a patient is paced at both six- and twelve-month follow-up.

Answer 4: Thank you very much for your suggestions! Permanent pacemaker implantation was included in our questionnaire, as described above. We will record when the participant's pacemaker is implanted and the reasons. During patient enrollment, we inform patients that if they have a subsequent readmission for treatment, they need to save and submit their case information to us during follow-up. The 3-day Holter monitoring devices are mailed to participants for wearing at 6-month and 12-month follow-up. Therefore, after wearing, we will record and analyze the time taken by the pacing rhythm at 6-month and 12-month follow-up. We have described in detail the recording of pacemaker implantation in the revised manuscript.

Change 4: Page 13, lines 310-316

“If permanent pacemaker is implanted in a participant during follow-up, we will record the date and reason that the participant's pacemaker is implanted by questionnaire. During participant enrollment, we inform participants that if they have a subsequent readmission for treatment, they need to save and submit their case information to us during follow-up. In addition, we will record and analyze the time taken by the pacing rhythm by 3-day Holter monitoring at 6-month and 12-month follow-up.”

Methodology:
Rhythm follow-up is by 72-hour Holter, and this complies with the 2017 Consensus Statement requirement for scientific studies. The definition of failure also complies with the Consensus Statement. Regarding the 72-hour Holter monitors, the protocol states that "after wearing, they are sent back to the project team for data analysis." Does this mean that a core lab will be used to analyze the monitors? Will the persons doing the analysis be blinded as to the randomization? Similarly, will any intermittent 12 lead ECGs be submitted to a core lab? Further, will the 12-month echocardiograms be interpreted by a blinded core lab?
Answer 5: Thank you for your comment. The 72-hour Holter monitors will be analyzed by a core lab blinded to the group allocation, and 12-lead ECG and echocardiograms are also analyzed by the core lab. We have made it clear in the revised manuscript.

Change 5: Page 13, lines 319-320

“The 3-day Holter monitoring, 12-lead electrocardiograms and echocardiograms will be analyzed by a core lab blinded to the group allocation.”

A major failing of the previous randomized trial [N Engl J Med 2015;372(15):1399-409] was that the surgical technique was not controlled, and multiple different lesion sets were performed and grouped together. You have taken the extraordinary step of insisting that surgical technique be standardized and verifying this by requiring any enrolling surgeon to submit a video of his surgical technique. I applaud you for this. As a minor concern I note that this requirement is stated in the Introduction and the Discussion but not in the Methodology section.

Answer 6: Thank you for your comment. We have added the requirement in the Method section.

Change 6: Page 10, lines 245-249

“All surgeons in this study are required to watch the video of standard Cox-Maze IV procedure and their surgical ablation procedures will be recorded before the trial, and incorrect or irregular manipulation will be reported back to surgeons, which is initiated to eliminate the impact of different tools and lesions on the results.”

The specified surgical technique requires that each bipolar lesion is performed three times. Dr Damiano had demonstrated that transmurality of a lesion is enhanced by applying energy to the bipolar clamp twice without releasing the clamp. You may wish to adopt this technique in the protocol.

Answer 7: Thank you for your suggestion, and we are very happy to adopt your technique. We have adopted this technique in the revised manuscript.

Change 7: Page 10, lines 238-239

“Each site is effectively ablated at least 3 times with radiofrequency clamp without releasing the radiofrequency clamp.”

Response to reviewer #3
Dr. Tao Chen, Liverpool School of Tropical Medicine
Comments to the Author:
This is an RCT aiming to evaluate the efficacy of bi-atrial ablation for patients with RMVD and non-paroxysmal AF, which have begun recruitment 5 months ago. My major concerns are listed as below.

1. It is unusual to observe the practice adopted for DMC in this study. DMC is an independent group who will assess the safety risk from the treatment. This means they cannot be blind and treatment allocation should be shared within the DMC members in the closed meeting. Meanwhile, DMC members meet regularly (eg, half a year). What do you mean urgent review in line 304? The authors may need carefully review their DMC charter. Some references from DAMOCLES study group (Lancet. 2005 Feb 19-25;365(9460):711-22) may be helpful.

Answer 1: We appreciate your time and effort that dedicated to providing feedback on our manuscript. Thank you very much for your questions.
After we submitted this manuscript, and before we enrolled participants, our team discussed the whole study protocol again in detail, including the necessity of data monitoring committee (DMC) and clinical event committee (CEC) again.

We know that the responsibilities of DMC include the following aspects: safety monitoring, effectiveness monitoring, trial operation quality monitoring, trial design adjustment suggestions, etc. First, in our study, we don't have a pre-set interim analysis. Second, the experimental group is bi-atrial ablation and the control group is left atrial ablation. Both of these modalities are routine clinical treatments, and bi-atrial ablation does not increase additional risks for patients. Third, the study will not last long from recruitment to the end of follow-up, which is expected to be 2 to 3 years. Therefore, considering these aforementioned reasons, we believe that setting up DMC in this study is unnecessary.

The responsibility of CEC is judging the endpoint events under uniform review criteria, and CEC is routinely used to adjudicate questionable endpoint events. In our study, the primary endpoint is the probability of freedom from atrial tachyarrhythmias at 12 months after operation documented by 3-day Holter monitoring, which is analyzed by a core lab blinded to the randomization. Therefore, the primary endpoint in our study is objective and clear, and the other secondary endpoints are likewise. Therefore, we believe that setting up CEC in this study is also unnecessary.

In summary, our study no longer establishes DMC and CEC and the description of DMC and CEC has been removed from our manuscript.

Change 1: Page 15, lines 349-361 (marked copy)
Data monitoring and clinical event committee
Data quality and safety are monitored by an independent DMC. The DMC has no competing interests in the ABLATION trial and monitors the study implementation and adverse event occurrence which is blinded to the group allocation. All serious adverse events are reported to the DMC for urgent review. DMC monitors the quality of study implementation by reviewing the study data, including monitoring protocol compliance, recruitment status, shedding rate of subjects, and the integrity of study data, and etc. If serious quality problems are found during study execution, DMC shall advise sponsors to improve the quality of study.

An independent clinical events committee (CEC) will adjudicate all clinical outcomes in accordance with the study’s prespecified adverse event definitions and in accordance with the CEC charter, which comprises experienced experts in the field blinded to the randomization schemes.

2. It is confused to me to learn the term of “survival rate” while describing the trial endpoint. What does this actually refer to? Is it incidence? It seems that “survival rate” is estimated from a KM curve. Please clarify and use associated proper word. In addition, effectiveness is repeatedly used in this trial, but this should be different from efficacy. From the trial design, it may not be suitable to use effectiveness.

Answer 2: Thanks for your suggestion. We are sorry that our endpoint is misleading. The primary endpoint is the probability of freedom from atrial tachyarrhythmias at 12 months. Therefore, we have changed the primary endpoint to “the probability of freedom from atrial tachyarrhythmias at 12 months”, and changed “survival rate” to “the probability of freedom from” in other endpoints. We have changed all the “effectiveness” to “efficacy” in the revised manuscript.

Change 2: Page 10, lines 252-254
“The primary endpoint is the probability of freedom from atrial tachyarrhythmias off antiarrhythmic drugs at 12 months after operation documented by 3-day Holter monitoring.”

Page 11, lines 259-261
“The key secondary endpoint is the probability of freedom from permanent pacemaker implantation at 12 months after operation, that is, the percentage of participants who do not have a new implanted permanent pacemaker.”

Page 11, lines 263-267
The secondary endpoints are the probability of freedom from atrial tachyarrhythmias with antiarrhythmic drugs, AF burden, incidence of adverse events (including cardiac death, stroke, hospitalization for heart failure, hospitalization for embolism events or major bleeding events), and cardiac function documented by echocardiography at 12 months after operation.

Revised Table 2. Endpoints in this trial

| Primary endpoint                                    |
|----------------------------------------------------|
| The probability of freedom from atrial tachyarrhythmias without AADs at 12 months after operation |

| Key secondary endpoint                             |
|----------------------------------------------------|
| The probability of freedom from permanent pacemaker implantation at 12 months after operation |

| Secondary endpoints                                  |
|----------------------------------------------------|
| The probability of freedom from atrial tachyarrhythmias with AADs at 12 months after operation |
| Burden of AF (Evaluating with 3-day Holter monitoring at 12 months after operation) |
| Incidence of adverse events (including cardiac death, stroke, hospitalization for heart failure, hospitalization for embolism events or bleeding events) |
| Cardiac function documented by echocardiography at 12 months after operation |

AF: Atrial fibrillation; AADs: Antiarrhythmic drugs

Page 3, lines 55-57
“The study aims to compare the efficacy and safety of bi-atrial ablation with those of left atrial ablation among patients with RMVD and persistent or longstanding persistent AF.”

3. The sample size is all calculated based on one-side 0.05, which means two sides 0.1. this is not common for a confirmative trial. Detailed explanation is needed, which is better for the readers.

Answer 3: Thanks for your useful suggestion. In designing the study protocol, we discussed in detail the choice of parameters “one-side 0.05 and 90% power” or “two-side 0.05 and 80% power”. We chose the former for several reasons.

First, this study is not a clinical drug trial and does not require registration with the FDA, so “false positive” can be controlled less strictly because “false negative” is equally important. From the overall study design, a hierarchical testing procedure is applied to the primary and key secondary endpoints, the hypothesis test of the key secondary endpoint is required to be carried out on the basis of the primary endpoint reached positive. Therefore, it is very important to obtain the positive result of the primary endpoint with greater power (90%), and then to carry out the sequential test of the key secondary endpoint. Once the hypothesis test of primary endpoint fails, there is no need for hypothesis test of key secondary endpoint. Therefore, in order to take into account this goal, we tend to choose a greater power.

Second, both the hypothesis tests of the primary endpoint and the key secondary endpoint are one-sided tests, and to ensure the uniformity of significance levels, we chose a significance level of one-side 0.05.
Third, “one-side 0.05 and 90% power” requires a larger sample size than “two-side 0.05 and 80% power”, and from a practical point of view, “one-side 0.05 and 90% power” is a more conservative parameter.

For the readers to better understand the design of our study, we explain this in detail in the Discussion section in the revised manuscript.

Change 3: Pages 16-17, lines 408-416

“This is an investigator-initiated study, and false positive can be controlled less strictly because the issue of false negative is equally important. From the overall study design, a hierarchical testing procedure is applied to the primary and key secondary endpoints, the hypothesis test of the key secondary endpoint can be carried out only if the primary endpoint reached positive. Therefore, it is very important to obtain the positive result of the primary endpoint with greater power, and then to carry out the sequential test of the key secondary endpoint. Once the hypothesis test of primary endpoint fails, there is no need for hypothesis test of key secondary endpoint. In order to take into account this goal, we tend to choose a greater power (90%). Therefore, we chose a significance level of one-side 0.05 and 90% power.”

4. A Gantt plot would be needed to facilitate reading.

Answer 4: Thanks for your comment. We have drawn a Gantt plot in the revised manuscript.

Change 4: Page 8, lines 189-191

“The ABLATION trial began recruitment in May 2022 and is expected to complete recruitment by the end of April 2024 and follow-up will be completed by the end of April 2025 (Figure 1).”

Figure 1.

VERSION 2 – REVIEW

| REVIEWER                      | James Edgerton  |
|-------------------------------|-----------------|
| Barnes-Jewish Hospital, Cardiothoracic Surgery |
| REVIEW RETURNED              | 13-Nov-2022     |

| GENERAL COMMENTS             | Thank you for considering all suggestions carefully and responding appropriately. Congratulations on a great study. I look forward to reading your results. |

| REVIEWER                      | Tao Chen        |
|-------------------------------|-----------------|
| Liverpool School of Tropical Medicine, Clinical Sciences |
| REVIEW RETURNED              | 14-Nov-2022     |

| GENERAL COMMENTS             | Thanks for sending me the response. I have no particular comments. However, a detailed Gannt chart including the items listed would be better. |

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2

Dr. James Edgerton, Barnes-Jewish Hospital

Comments to the Author:
Thank you for considering all suggestions carefully and responding appropriately. Congratulations on a great study. I look forward to reading your results.

Answer: Thank you again for your valuable time to provide us with very useful comments on the manuscript. Thanks very much!

Reviewer: 3
Dr. Tao Chen, Liverpool School of Tropical Medicine
Comments to the Author:
Thanks for sending me the response. I have no particular comments. However, a detailed Gantt chart including the items listed would be better.

Answer: Thank you again for your valuable time to provide us with very useful comments. We have added “Formulation of clinical research protocol, Team formation and investigator meetings, Central ethics approval, Formulation of project management plan, Pre-initiation training” to the Gantt chart to make it more detailed.

Change: Revised Figure 1.