PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Protocol for a Single-Centre Prospective Observational Study of Postoperative Delirium Following Total Joint Arthroplasties among South East Asians
AUTHORS	abdullah, hairil; Tan, Sapphire; Lee, Si Jia; Rahmatullah Bin Abd Razak, Hamid; Seet, Rachel; Ying, Hao; Sethi, Ervin; Sim, Eileen

VERSION 1 – REVIEW

REVIEWER	R.A. Pol UMCG, The Netherlands
REVIEW RETURNED	26-Nov-2017

GENERAL COMMENTS	Overall decision: In general it can be an interesting study with a specific focus on postoperative delirium (POD) after total joint arthroplasties. I think it can be a great opportunity for a nice prospective and large registry. The study design is well explained well and clearly. However, it lacks in focus on POD. Even though the authors report a good approach for diagnosing POD and clearly underline the need for such a prospective study, they also want to look at long-term functional outcome (using several knee scores), the BANG-score (developed for sleep apnoes?) and long-term mortality. I can understand their interest in functional outcome but this should not be included in a study protocol on POD. Although the occurrence of POD certainly has effect on long-term cognitive decline and mortality, including such a long list of secondary outcome variables is very distracting from the primary outcome and purpose: identifying POD and its risk factors! I would suggest to Comments to the Authors: Major: 1) Please focus on your main outcome in this study protocol (POD!) and remove all functional parameters. Unless these outcomes are related to POD (which I don't think). 2) Introduction is too long and should be shorter with more focus on the problem. For instance, 16% POD incidence is at the low spectrum (compared to oncology and vascular surgery patients). Is there an actual knowledge gap in POD development in Asian patients? You contradict yourself at the end of the discussion section with the hypothesis that the incidence will be similar. Then why the emphasis on the Asian population? Focus on POD after joint arthroplasty. Minor: 1) Postoperative Morbidity Survey is not the best way to report surgical outcome. I would suggest to use the Comprehensive Complication Index.

	That way you do not need to dichotomize your results (and loose important data) 2) I disagree with your rational of starting POMS at day 5. In a prospective study (using CRFs) you should collect and report any deviant course from day of admission to discharge. 3) CAM is used for POD screening. But how do you make the appropriate diagnosis? I think using the DSM classification? 4) Please add a clinical epidemiologist to your group to check the statistical analysis. For instance, 'Odds ratio will be used to describe the relative risk……. These can be 2 total different thinks! Also the sample size calculation is confusing. Please emphasize on the intended study population (200 or 500?). I understand the difficulty of running a logistic regression model with a low POD incidence (risk for overfitting) but you have to clarify this more. 5) Please write the paper in present tense (instead of past tense: analyses were performed, variables were incorporated..etc) 6) Please use proper abbreviations in the reference list (ref 5 and 6).

REVIEWER	Tao Yang Life Sciences Institute, University of Michigan, Ann Arbor, MI, USA
REVIEW RETURNED	26-Nov-2017

GENERAL COMMENTS	The research about the postoperative delirium benefits all the patients of TJA. Its risk factors are unknown. The research about the risk factors should be designed more carefully and thoughtfully. This protocol describes the coming research which will be conducted in SGH to screen the postoperative delirium and characterize the risk factors related to the delirium. The screen will be done in the patients with TKA or THA operation in SGH. Preoperative, postoperative and long-term postoperative mental states will be estimated. Delirium of the patients will be observed by comparing the CAM questionnaire and some other questionnaires pre- and postoperatively. The risk factors related to the delirium might be characterized during the research. Major concern: 1, The patients between 65 and 90 years old will be invited to enroll in the study. The authors should explain why the age is relevant to this study. The younger patients may have the lower percentage of delirium, but they may give the clues of the risk factors. 2, CAM, the only method for the delirium diagnosis in the research, has the 94% - 100% sensitivity and ~90% specificity. Do the authors have the backup method to check the CAM and work together with CAM to cover higher percentage of delirium patients? 3, In the protocol, there is the method to estimate the TKA long-term outcomes, but there are no methods for THA. Will the long-term outcomes of THA patients be estimated? How will the long-term outcomes of THA be estimated? 4, The sample size is estimated roughly. However, the protocol mentioned multiple aims including the risk factors of TJA delirium, which may not be easy to figure out, especially when the difference is small, and the variety is big. The authors should think carefully about the candidates of their risk factor research and estimate the difference based on the published data or experience. Then, they can estimate their sample size more accurately. 5, Double-blind quantification is important for the research to clarify the risk factors of the postoperative delirium.

	How does this protocol design the double-blind quantification method? Minus concerns: 1, The words like “characterise” should be “characterize”, “institutionalisition” should be “institutionalization”. There are lots of the similar mistakes. Please revise the whole manuscript to correct them. Please revise the English carefully for the whole manuscript. 2, In Preoperative data session, the authors mentioned CAM (page 9, 32) but didn’t mention its full name. In Postoperative data, Short-term secondary outcomes session, LOS (page 11, 43) was mentioned without the full name. Please add the full name when the acronym was mentioned the first time. 3, There is no proof to support that “poor compliance with postoperative therapy” with “the postoperative delirium” mentioned in page 7, 26-27.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1:
Major concerns:
1) Please focus on your main outcome in this study protocol (POD!) and remove all functional parameters. Unless these outcomes are related to POD (which I don’t think).

Thank you for your suggestion. We appreciate the idea, however, on reflection of our study aims, we decided that it is important to delineate and compare the outcomes between patients who suffered from Postoperative Dementia and those who did not. This postoperative delirium has been associated with increased morbidity and healthcare cost, and in our opinion, it is important to study how they impact the longer-term outcomes for the surgery itself (TJA). If POD truly effects the longer-term outcomes, targeted intervention (e.g. personalized rehabilitation/cognitive training etc.) could be investigated in the future.

We understand Dr Pol’s concerns, and have edited our manuscript to reflect these outcome measure as secondary outcomes. We have given more emphasis on our primary outcome, which is the incidence and risk factors of postoperative delirium. These can be found in (Page 7, Lines 29-36) the revised protocol.

2) Introduction is too long and should be shorter with more focus on the problem. For instance, 16% POD incidence is at the low spectrum (compared to oncology and vascular surgery patients). Is there an actual knowledge gap in POD development in Asian patients? You contradict yourself at the end of the discussion section with the hypothesis that the incidence will be similar. Then why the emphasis on the Asian population? Focus on POD after joint arthroplasty.

Thank you for your comment. We agree with your point of view and have amended our manuscript accordingly to reflect the knowledge gap and why it is important to have this study done in our region. This could be found on page 7, lines 29-32 of the revised protocol.

Minor concerns:
1) Postoperative Morbidity Survey is not the best way to report surgical outcome. I would suggest to use the Comprehensive Complication Index. That way you do not need to dichotomize your results (and loose important data)

Postoperative Morbidity Survey (POMS) is widely used in perioperative medicine literature and has been validated for use in major orthopaedic surgeries including TJAs [1,2].
POMS identifies morbidity of a type and severity that could delay discharge from hospital with a simple data collection process so that large numbers of patients can be routinely screened. We do not agree that we will lose important data as the individual domains of POMS, as well as the total morbidity burden could be presented and analysed as well.

[1] Wong DJN, Oliver CM, Moonesinghe SR. Predicting postoperative morbidity in adult elective surgical patients using the Surgical Outcome Risk Tool (SORT). Br J Anaesth. Oxford University Press; 2017;119: 95–105.
[2] Grocott MPW, Browne JP, Van der Meulen J, Matejowsky C, Mutch M, Hamilton MA, et al. The Postoperative Morbidity Survey was validated and used to describe morbidity after major surgery. J Clin Epidemiol. 2007;60: 919–928.

2) I disagree with your rational of starting POMS at day 5. In a prospective study (using CRFs) you should collect and report any deviant course from day of admission to discharge.

We agree with your comments and have reflected the changes in the revised manuscript (Page 10, Lines 26-27). We will collect POMS for POD 3, 5, 8 and 15 as per the original literature [3].

[3] Bennett-Guerrero E, Welsby I, Dunn TJ, Young LR, Wahl TA, Diers TL, et al. The use of a postoperative morbidity survey to evaluate patients with prolonged hospitalization after routine, moderate-risk, elective surgery. Anesth Analg. 1999;89: 514–519.

3) CAM is used for POD screening. But how do you make the appropriate diagnosis? I think using the DSM classification?

Thank you for your comment. We agree that DSM-V provides the framework for formally diagnosing delirium, hence all CAM positive patients will be referred to our psychiatry team for a formal assessment and management. This has been reflected in (Page 10, Lines 18-19) of the revised protocol. For the purpose of our study, CAM is an acceptable primary outcome measure. This is possible due to the preoperative baseline cognitive assessment, as well as the known physiological stressor (surgery) which the patient underwent. This is reflected in a systematic review by Wei et al 2008 [4] which demonstrated CAM as one of the most common outcome measure in delirium studies. We will however follow up on the patients referred to the psychiatrist and report on the prevalence.

[4] Wei LA, Fearing MA, Sternberg EJ, Inouye SK. The Confusion Assessment Method: a systematic review of current usage. J Am Geriatr Soc. 2008;56: 823–830.

4) Please add a clinical epidemiologist to your group to check the statistical analysis. For instance, ‘Odds ratio will be used to describe the relative risk……. These can be 2 total different thinks! Also the sample size calculation is confusing. Please emphasize on the intended study population (200 or 500?). I understand the difficulty of running a logistic regression model with a low POD incidence (risk for overfitting) but you have to clarify this more.

We agree with your comments. We have modified our use of odds ratio so describe the association of the risk factors with postoperative delirium, instead of the relative risk. This can be seen on page 13, lines 3-4 of the revised protocol. We have also clarified the intended study population size of 500 to allow for logistic regression modelling. This can be seen be found on page 12, lines 22-25 the revised protocol.

5) Please write the paper in present tense (instead of past tense: analyses were performed, variables were incorporated..etc)
Thank you for your comment. We agree with your point and have amended our manuscript accordingly. This can be found in (Page 13, Lines 17) the revised protocol.

6) Please use proper abbreviations in the reference list (ref 5 and 6).

Thank you for your comment. We agree with your point and have amended our manuscript accordingly. This can be found in (Page 15, Lines 9-14) the revised protocol.

Reviewer 2:
Major concerns:
1) The patients between 65 and 90 years old will be invited to enroll in the study. The authors should explain why the age is relevant to this study. The younger patients may have the lower percentage of delirium, but they may give the clues of the risk factors.

Thank you for your comment. We have chosen a minimum age of 65 years based on the original study which developed and validated the CAM as a tool that enables non-psychiatric clinicians to detect delirium quickly in high-risk settings [5]. We have clarified this in the revision, which can be found on page 8, lines 11-13 of the revised protocol.

[5] Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegal AP, Horwitz RI. Clarifying confusion: the confusion assessment method. A new method for detection of delirium. Ann Intern Med. 1990;113(12):941-948.

2) CAM, the only method for the delirium diagnosis in the research, has the 94% - 100% sensitivity and ~90% specificity. Do the authors have the backup method to check the CAM and work together with CAM to cover higher percentage of delirium patients?

Thank you for your comment and we agree with your point of view. Patients which are CAM positive will be referred to our psychiatry team for a formal assessment based on DSM-V and management. This has been reported in (Page 10, Line 18-19) the revised protocol. We will follow up on these patients referred to the psychiatrist and report the prevalence.

3) In the protocol, there is the method to estimate the TKA long-term outcomes, but there are no methods for THA. Will the long-term outcomes of THA patients be estimated? How will the long-term outcomes of THA be estimated?

We agree with you comment and we will be analysing long-term outcomes specifically for THA in addition to those for TKA. We will be using the Harris Hip Score and the Parker Mobility Score to evaluate the functional outcomes of patients following THA. This can be found in (Page 11, Lines 12-21) the revised protocol. These will be analysed using the Wilcoxon matched pairs sign rank test, similar to the outcomes following TKA.

4) The sample size is estimated roughly. However, the protocol mentioned multiple aims including the risk factors of TJA delirium, which may not be easy to figure out, especially when the difference is small, and the variety is big. The authors should think carefully about the candidates of their risk factor research and estimate the difference based on the published data or experience. Then, they can estimate their sample size more accurately.

Thank you for your comment. We agree that the incidence for each outcome is low, and a large sample size is needed to run multiple logistic models to detect potential factors associated with postoperative delirium. We thus aim to recruit 500 patients in this study.
Even though we may only be able to test a few variables, the primary aim of this study it still to estimate the incidence of postoperative delirium following TJA, and not to come up with a risk prediction model. Furthermore, many existing studies recruit around 200 patients [6]. Thus, we believe that 500 patients should be sufficient for our purposes, and this has been reported in (Page 12, Lines 22-25) the revised protocol.

[6] Scott JE, Mathias JL, Kneebone AC. Incidence of delirium following total joint replacement in older adults: A meta-analysis. Gen Hosp Psychiatry. 2015;37(3):223-229. doi:10.1016/j.genhosppsych.2015.02.004.

5) Double-blind quantification is important for the research to clarify the risk factors of the postoperative delirium. How does this protocol design the double-blind quantification method?

We agree with your comments and our protocol does align to the double-blind quantification method. The team conducting the postoperative assessments is blinded from the preoperative baseline assessment. The preoperative assessment team is also not the clinical team taking care of the patient, hence all aspects of the care will be unchanged.

Minor concerns:
1) The words like “characterise” should be “characterize”, “institutionalisition” should be “institutionalization”. There are lots of the similar mistakes. Please revise the whole manuscript to correct them. Please revise the English carefully for the whole manuscript.

Thank you for your comment. However, we believe that the British spelling of these words are spelled with an “s” instead of a “z”, unlike the American spelling. As we are submitting to a British journal, we will use the British spelling of these words instead.

2) In Preoperative data session, the authors mentioned CAM (page 9, 32) but didn’t mention its full name. In Postoperative data, Short-term secondary outcomes session, LOS (page 11, 43) was mentioned without the full name. Please add the full name when the acronym was mentioned the first time.

Thanks for your comments, and we have amended our manuscript accordingly. This can be seen in Page 9, Line 1 and Page 10, Line 24 of the revised protocol.

3) There is no proof to support that “poor compliance with postoperative therapy” with “the postoperative delirium” mentioned in page 7, 26-27.

Thank you for your comment. We have removed this statement and it is no longer present in the revised protocol.

VERSION 2 – REVIEW

REVIEWER	R.A. Pol University Medical Center Groningen, University of Groningen, The Netherlands
REVIEW RETURNED	07-Jan-2018

GENERAL COMMENTS	Overall decision: I would like to thank and compliment the authors for their comprehensive and detailed revision. The introduction has much more focus on the primary outcome measure POD. The statistical analysis has improved with the clarification on the appropriate power but this still holds a major methodological risk. The reference list needs further adjusting according to the guidelines. With minor revision I think this study protocol will lead to a nice prospective and large registry and provide novel and important insights in POD characteristics and new ways to prevent it. Comments to the Authors: Major: 1) Please focus on your main outcome in this study protocol (POD!) and remove all functional parameters. Unless these outcomes are related to POD (which I don't think). Authors reply: Thank you for your suggestion. We appreciate the idea, however, on reflection of our study aims, we decided that it is important to delineate and compare the outcomes between patients who suffered from Postoperative Dementia and those who did not. This postoperative delirium has been associated with increased morbidity and healthcare cost, and in our opinion, it is important to study how they impact the longer-term outcomes for the surgery itself (TJA). If POD truly effects the longer-term outcomes, targeted intervention (e.g. personalized rehabilitation/cognitive training etc.) could be investigated in the future. We understand Dr Pol's concerns, and have edited our manuscript to reflect these outcome measure as secondary outcomes. We have given more emphasis on our primary outcome, which is the incidence and risk factors of postoperative delirium. These can be found in (Page 7, Lines 29-36) the revised protocol. Reviewer reply: Sufficiently addressed and adjusted and thereby much improved 2) Introduction is too long and should be shorter with more focus on the problem. For instance, 16% POD incidence is at the low spectrum (compared to oncology and vascular surgery patients). Is there an actual knowledge gap in POD development in Asian patients? You contradict yourself at the end of the discussion section with the hypothesis that the incidence will be similar. Then why the emphasis on the Asian population? Focus on POD after joint arthroplasty. Author reply: Thank you for your comment. We agree with your point of view and have amended our manuscript accordingly to reflect the knowledge gap and why it is important to have this study done in our region. This could be found on page 7, lines 29-32 of the revised protocol. Reviewer reply: Sufficiently adjusted and thereby much improved Minor: 1) Postoperative Morbidity Survey is not the best way to report surgical outcome. I would suggest to use the Comprehensive Complication Index. That way you do not need to dichotomize your results (and loose important data) Author reply: Postoperative Morbidity Survey (POMS) is widely used in perioperative medicine literature and has been validated for use in

	major orthopaedic surgeries including TJAs [1,2]. POMS identifies morbidity of a type and severity that could delay discharge from hospital with a simple data collection process so that large numbers of patients can be routinely screened. We do not agree that we will lose important data as the individual domains of POMS, as well as the total morbidity burden could be presented and analysed as well. [1] Wong DJN, Oliver CM, Moonesinghe SR. Predicting postoperative morbidity in adult elective surgical patients using the Surgical Outcome Risk Tool (SORT). Br J Anaesth. Oxford University Press; 2017;119: 95–105. [2] Grocott MPW, Browne JP, Van der Meulen J, Matejowsky C, Mutch M, Hamilton MA, et al. The Postoperative Morbidity Survey was validated and used to describe morbidity after major surgery. J Clin Epidemiol. 2007;60: 919–928. Reviewer reply: I would still advice to use the Comprehensive Complication Index (CCI) with the same key argument. Dichotomizing results leads to loss of data. The CCI takes the quantity of appearance of each complication into account, using a specific calculation that yields a score from 0 to 100, thereby giving a very detailed assessment for every patient. 2) I disagree with your rational of starting POMS at day 5. In a prospective study (using CRFs) you should collect and report any deviant course from day of admission to discharge. Author reply: We agree with your comments and have reflected the changes in the revised manuscript (Page 10, Lines 26-27). We will collect POMS for POD 3, 5, 8 and 15 as per the original literature [3]. [3] Bennett-Guerrero E, Welsby I, Dunn TJ, Young LR, Wahl TA, Diers TL, et al. The use of a postoperative morbidity survey to evaluate patients with prolonged hospitalization after routine, moderate-risk, elective surgery. Anesth Analg. 1999;89: 514–519. Reviewer reply: Somewhat adjusted to starting at day 3. Please keep in mind the biological substrate of POD which can be triggered by something as simple as a blood transfusion at day 1. Or a short episode of fever. 3) CAM is used for POD screening. But how do you make the appropriate diagnosis? I think using the DSM classification? Author reply: Thank you for your comment. We agree that DSM-V provides the framework for formally diagnosing delirium, hence all CAM positive patients will be referred to our psychiatry team for a formal assessment and management. This has been reflected in (Page 10, Lines 18-19) of the revised protocol. For the purpose of our study, CAM is an acceptable primary outcome measure. This is possible due to the preoperative baseline cognitive assessment, as well as the known physiological stressor (surgery) which the patient underwent. This is reflected in a systematic review by Wei et al 2008 [4] which demonstrated CAM as one of the most common outcome measure in delirium studies. We will however follow up on the patients referred to the psychiatrist and report on the prevalence. [4] Wei LA, Fearing MA, Sternberg EJ, Inouye SK. The Confusion Assessment Method: a systematic review of current usage. J Am Geriatr Soc. 2008;56: 823–830. Reviewer reply: Sufficiently addressed and adjusted and thereby

	improved 4) Please add a clinical epidemiologist to your group to check the statistical analysis. For instance, 'Odds ratio will be used to describe the relative risk……. These can be 2 total different thinks! Also the sample size calculation is confusing. Please emphasize on the intended study population (200 or 500?). I understand the difficulty of running a logistic regression model with a low POD incidence (risk for overfitting) but you have to clarify this more. Author reply: We agree with your comments. We have modified our use of odds ratio so describe the association of the risk factors with postoperative delirium, instead of the relative risk. This can be seen on page 13, lines 3-4 of the revised protocol. We have also clarified the intended study population size of 500 to allow for logistic regression modelling. This can be seen be found on page 12, lines 22-25 the revised protocol. Reviewer reply: Sufficiently addressed and adjusted. I would still advice you to add an epidemiologist to the group to anticipate the methodological problems when overfitting occurs. 5) Please write the paper in present tense (instead of past tense: analyses were performed, variables were incorporated..etc) Author reply: Thank you for your comment. We agree with your point and have amended our manuscript accordingly. This can be found in (Page 13, Lines 17) the revised protocol. Reviewer reply: Sufficiently and adequately adjusted 6) Please use proper abbreviations in the reference list (ref 5 and 6). Author reply: Thank you for your comment. We agree with your point and have amended our manuscript accordingly. This can be found in (Page 15, Lines 9-14) the revised protocol. Reviewer reply: The ref list is not in accordance with the BMJ guidelines and the reference examples provided on the website. For instance, some (not all) references are displayed with the pubmed URL. Please be consistent and stick to the guidelines.

REVIEWER	Tao Yang Institute of Life Sciences, University of Michigan, Ann Arbor, MI, USA
REVIEW RETURNED	15-Jan-2018

GENERAL COMMENTS	In the revised manuscript, all of my concerns are corrected or explained. It is great to see the improvement that the delirium patients will be referred to the psychiatrist to make the diagnosis. The protocol matches the goals of this research. I support the publication of this protocol on BMJ open.

VERSION 2 – AUTHOR RESPONSE

Reviewer 1:

Minor comments:

1) Reviewer reply: I would still advice to use the Comprehensive Complication Index (CCI) with the same key argument. Dichotomizing results leads to loss of data. The CCI takes the quantity of

appearance of each complication into account, using a specific calculation that yields a score from 0 to 100, thereby giving a very detailed assessment for every patient.

Thank you for suggestion, and we have decided to incorporate it into the protocol. The CCI will be used in assessing postoperative morbidities, and will be conducted on the day-of-discharge and POD 30. This can be found on Page10, Lines 29-36 of the revised protocol.
4) Reviewer reply: Sufficiently addressed and adjusted. I would still advice you to add an epidemiologist to the group to anticipate the methodological problems when overfitting occurs.

Thank you for your advice. We have a biostatistician in our research group, which we incorrectly assigned to another role. We have amended the protocol accordingly, which is reflected in (Page 1, Lines 34-37) the revised protocol.

6) Reviewer reply: The ref list is not in accordance with the BMJ guidelines and the reference examples provided on the website. For instance, some (not all) references are displayed with the pubmed URL. Please be consistent and stick to the guidelines.

We agree your comment. We have looked into the reference list again and reviewed all the references. They are now all according to the BMJ guidelines. This is reflected in Pages 15-17 of the revised protocol.

Reviewer 2:
Thank you for reviewing our protocol and supporting our publication.

VERSION 3 – REVIEW

REVIEWER	R.A. Pol University Medical Center Groningen The Netherlands
REVIEW RETURNED	29-Jan-2018

GENERAL COMMENTS	1) Reviewer reply: I would still advice to use the Comprehensive Complication Index (CCI) with the same key argument. Dichotomizing results leads to loss of data. The CCI takes the quantity of appearance of each complication into account, using a specific calculation that yields a score from 0 to 100, thereby giving a very detailed assessment for every patient. Thank you for suggestion, and we have decided to incorporate it into the protocol. The CCI will be used in assessing postoperative morbidities, and will be conducted on the day-of-discharge and POD 30. This can be found on Page10, Lines 29-36 of the revised protocol. Reviewer reply: Thank you and I have no further comments or suggestions. 4) Reviewer reply: Sufficiently addressed and adjusted. I would still advice you to add an epidemiologist to the group to anticipate the methodological problems when overfitting occurs. Thank you for your advice. We have a biostatistician in our research group, which we incorrectly assigned to another role. We have

	amended the protocol accordingly, which is reflected in (Page 1, Lines 34-37) the revised protocol. Reviewer reply: Agree, no further comment. 6) Reviewer reply: The ref list is not in accordance with the BMJ guidelines and the reference examples provided on the website. For instance, some (not all) references are displayed with the pubmed URL. Please be consistent and stick to the guidelines. We agree your comment. We have looked into the reference list again and reviewed all the references. They are now all according to the BMJ guidelines. This is reflected in Pages 15-17 of the revised protocol. Reviewer reply: thank you, no further comments. Thank you for your patience during the review process and good luck with the study!