

## 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>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

### ARTICLE DETAILS

<b>TITLE (PROVISIONAL)</b>	Impact of surgeon and anaesthesiologist sex on patient outcomes after cardiac surgery: a population-based study
<b>AUTHORS</b>	Sun, Louise; Etherington, Nicole; Boet, S; Chan, Vincent; Lee, Douglas; Mesana, Theirry; Bader Eddeen, Anan

### VERSION 1 – REVIEW

<b>REVIEWER</b>	Janda, Allison University of Michigan Health System, Anesthesiology
<b>REVIEW RETURNED</b>	10-Apr-2021

<b>GENERAL COMMENTS</b>	<p>Overall Summary: Well-designed study assessing the impact of surgeon-anesthesiologist dyad sex discordance in cardiac procedures. I agree that cardiac surgeries specifically require high-caliber team communication, coordination and synergy so I appreciate that this patient population was examined in this study. I recommend a few minor clarifications and considerations.</p> <ul style="list-style-type: none"><li>- Reviewer comment 1: In the abstract, the effect sizes and description of statistical significance are not included. Please add the effect sizes for the statistically significant subgroup analysis findings (i.e. OR for hospital LOS and p-value).</li><li>- Reviewer comment 2: Please delete the extra space prior to reference 11, page 4, lines 32/33.</li><li>- Reviewer comment 3: How was stroke defined in your MACE outcome in the database you utilized? Did this include hemorrhagic and ischemic stroke? Please elaborate.</li><li>- Reviewer comment 4: Starting on page 8, there are paragraph indentation inconsistencies, please correct (i.e. p. 8 line 32/33 – no indent, p. 8 line 44 – no indent, p. 8 line 53/54 – no indent, p. 9 line 7 – no indent, p. 9 line 28 – no indent, p. 9 line 44 – no indent, p. 10 line 5 – no indent, etc.).</li><li>- Reviewer comment 5: Was an a priori power analysis performed? As the findings are null for the primary analysis, the null finding could be a result of a lack of adequate sample size. There are a high number of cases in the cohort, but the prevalence of the primary outcome, 30-day all-cause mortality is low, which could contribute to a lack of power to detect a difference in the primary outcome between groups. Did you consider performing a post-hoc power analysis if an a priori power analysis was not done?</li><li>- Reviewer comment 6: On page 10, lines 53-54, please add the units for LOS.</li></ul>
-------------------------	---

	<p>- Reviewer comment 7: On page 11, lines 23-25, supplemental table 5 is referenced, but for 30-day mortality, MACE and ICU LOS for the full cohort and CABG-only cohort, should this sentence reference supplemental tables 5a, 5b, 6a and 6b?</p> <p>- Reviewer comment 8: Please provide the effect sizes and p-values for the findings you report on page 11, lines 25-28, “an all-male physician team as compared to an all-female team was associated with longer hospital LOS in CABG patients.” And please reference the correct supplemental table since one is not referenced here.</p> <p>- Reviewer comment 9: On page 11, lines 35-37 reference the incorrect supplemental table, from the version attached to the submission, it looks like it should reference supplemental tables 6a and 6b (not 5b only). Also please report the odds ratios and p-values for these results in the text.</p> <p>- Reviewer comment 10: Please discuss the lack of an a priori power analysis in the limitations section if one was not completed.</p> <p>- Reviewer comment 11: In table 2, surgeon and anaesthesiologist volume is noted “per year” but the text on page 7, describes the volume as: “total number of procedures performed since the inception of ICES databases in 1991 until the date of the index procedure.” Apologies if I’m misinterpreting table 2, but the case volumes for each surgeon or anaesthesiologist in table 2 seem high if they are actually cardiac case volume “per year” since the median for an attending anaesthesiologist would be over 2 cardiac cases per day for every day of the year. If this is general case volume instead of cardiac case volume, or not per year, please clarify this in the text or in a footnote on the tables.</p>
--	--

<b>REVIEWER</b>	Satkunasivam, Raj Sunnybrook Health Sciences Centre
<b>REVIEW RETURNED</b>	19-May-2021

<b>GENERAL COMMENTS</b>	<p>This is a retrospective cohort study investigating the association between cardiac surgical physician (surgeon and anesthesiology) team sex discordance and patient outcomes in patients who underwent cardiac surgery (defined as CABG and/or aortic, mitral, or tricuspid valvular repair). The authors sought to characterize outcomes in all-cause 30-day mortality, 30-day incidence of major adverse cardiovascular events (MACE) and lengths of stay (LOS) in intensive care.</p> <p>They found that physician sex discordance was not associated with overall patient mortality, but those who underwent CABG experienced longer LOS when treated by an all-male physician team as opposed to an all-female team.</p> <p>Abstract:</p> <ul style="list-style-type: none"> <li>- Background is unclear, better to state the existing literature (e.g. optimal performance in OR crisis situation)</li> <li>- Objective- need to include the main exposure variable that makes this study unique, “discordance”</li> </ul> <p>Methods:</p>
-------------------------	---

	<p>1. Why did the authors exclude re-do cardiac procedures, which represent a greater/easier definition of case complexity for subgroup analyses.</p> <p>2. Why was pulmonic valvular repair excluded?</p> <p>3. Do you have information of cardiac fellowship training? Do we have information of anesthesia sub-specialization (e.g. cardiac anesthesia)?</p> <p>4. Please expand on how you adjusted for surgeon volume. I am confused why total volume of procedures was used for adjustment (e.g. anything the surgeon uses) dating back to 1991 until index procedures. This is as opposed to focusing on the procedures of interest to this study (CABG, MV repair etc). Further, the most important aspect of volume is that which is contemporaneous to the index procedures – what the volume is for the surgeon (e.g. the year before,) doing the procedures of interest before doing the procedure of interest.</p> <p>5. I would be tempted to consider a composite outcome of any Mortality or MACE in light of a negative results for these alone. Is this a power issue?</p> <p>Interpretation:</p> <p>1. The authors report, interestingly, that an all-male physician and anesthesiologist team had a, longer hospital length of stay than a sex-discordant or all female team. They also found a longer hospital length of stay for patients when individual sex was considered for both male anesthesiologists and male surgeon. Though these differences are statistically significant, their clinical significance is not commented on. The rate ratio, for example, of male anesthesiologist v. female anesthesiologist in terms of longer hospital length of stay is 1.02. Authors do not comment on the potential clinical impact of this difference.</p> <p>2. Further to this point LOS is a process measure and not necessarily an outcome. How do we know we are not just measuring differences in practice patterns of sex-discordant pairs?</p> <p>Conclusion: Please align your conclusions only with the objective findings of your study.</p>
--	---

## VERSION 1 – AUTHOR RESPONSE

### REVIEWER 1 COMMENTS

**Comment 1:** In the abstract, the effect sizes and description of statistical significance are not included. Please add the effect sizes for the statistically significant subgroup analysis findings (i.e. OR for hospital LOS and p-value).

**Response:** Thank you for this suggestion. We have now added these values to the abstract:

“Physician sex discordance was not associated with overall patient mortality or LOS; however, patients who underwent isolated CABG experienced longer hospital LOS when treated by an all-male physician team as compared to an all-female team (adjusted odds ratio [OR]=1.07; p=0.049). When examining the impact of individual physician sex, the length of hospital stay was longer when isolated CABG procedures were attended by a male surgeon (OR=1.10; p=0.004) or anaesthesiologist

(OR=1.02; p=0.01).”

**Comment 2:** Please delete the extra space prior to reference 11, page 4, lines 32/33.

**Response:** This space has now been deleted.

**Comment 3:** How was stroke defined in your MACE outcome in the database you utilized? Did this include hemorrhagic and ischemic stroke? Please elaborate.

**Response:** We have now revised the methods section to expand on this definition:

“Secondary outcomes were hospital and intensive care unit (ICU) lengths of stay (LOS) as well as major adverse cardiovascular events (MACE). MACE was defined as a composite of stroke, repeat revascularization, hospitalization for MI and HF. Stroke included ischemic stroke and was generally defined as new focal or global neurologic deficit of cerebrovascular origin lasting 24 hours or longer that was not present before surgery”

**Comment 4:** Starting on page 8, there are paragraph indentation inconsistencies, please correct (i.e. p. 8 line 32/33 – no indent, p. 8 line 44 – no indent, p. 8 line 53/54 – no indent, p. 9 line 7 – no indent, p. 9 line 28 – no indent, p. 9 line 44 – no indent, p. 10 line 5 – no indent, etc.).

**Response:** Thank you for pointing these out. We have now revised the manuscript to ensure indentations are used consistently.

**Comment 5:** Was an a priori power analysis performed? As the findings are null for the primary analysis, the null finding could be a result of a lack of adequate sample size. There are a high number of cases in the cohort, but the prevalence of the primary outcome, 30-day all-cause mortality is low, which could contribute to a lack of power to detect a difference in the primary outcome between groups. Did you consider performing a post-hoc power analysis if an a priori power analysis was not done?

**Response:** We did not conduct a power analysis a priori but have now included one post hoc.

Results: “We conducted a post-hoc power analysis to determine whether the lack of observed between group mortality difference was due to the small number of outcome events. Using logistic regression with a sample size of 79,862 patients (24.9% treated by sex discordant surgeon-anesthesiologist pairs) and an observed OR of 0.93, we were able to achieve 19% power at a 0.05 significance level. At the request of the reviewers, we repeated our analysis for the composite endpoint of death and MACE. The findings of this post hoc analysis also did not reach statistical significance (adjusted OR, 0.96 [0.88-1.05], p=0.37; Supplemental Table 7).”

**Comment 6:** On page 10, lines 53-54, please add the units for LOS.

**Response:** The units have now been provided:

“Median ICU and hospital LOS were 2 days (IQR, 2-3) and 7 days (6-9), respectively, both in patients who were treated by sex discordant and concordant physicians (Table 3).”

**Comment 7:** On page 11, lines 23-25, supplemental table 5 is referenced, but for 30-day mortality, MACE and ICU LOS for the full cohort and CABG-only cohort, should this sentence reference supplemental tables 5a, 5b, 6a and 6b?

**Response:** Thank you for pointing this out. We have corrected the referencing of these tables:

“We did not observe an independent association between teams comprised of male surgeon - male anaesthesiologist, male surgeon - female anaesthesiologist, female surgeon - male anaesthesiologist, and female surgeon - female anaesthesiologist, and 30-day mortality, MACE, or ICU LOS (Supplemental Tables 5a and 5b). However, an all-male physician team as compared to an all-female team was associated with longer hospital LOS in CABG patients (adjusted OR=1.07 [1.00-1.15]; p=0.049) (Supplemental Table 6a).”

“Male as compared to female surgeon (adjusted OR=1.10 [1.03-1.18]; p=0.004), and male vs. female anaesthesiologist (adjusted OR=1.02 [1.00-1.04]; p=0.01), was associated with longer hospital LOS in the overall and CABG patient groups (Supplemental Table 6b).”

**Comment 8:** Please provide the effect sizes and p-values for the findings you report on page 11, lines 25-28, “an all-male physician team as compared to an all-female team was associated with longer hospital LOS in CABG patients.” And please reference the correct supplemental table since one is not referenced here.

**Response:** Thank you for this suggestion. We have now provided the effect sizes and p-values, as well as corrected the table numbers:

“However, an all-male physician team as compared to an all-female team was associated with longer hospital LOS in CABG patients (adjusted OR=1.07 [1.00-1.15]; p=0.049) (Supplemental Table 6a).”

“Male as compared to female surgeon (adjusted OR=1.10 [1.03-1.18]; p=0.004), and male vs. female anaesthesiologist (adjusted OR=1.02 [1.00-1.04]; p=0.01), was associated with longer hospital LOS in the overall and CABG patient groups (Supplemental Table 6b).”

**Comment 9:** On page 11, lines 35-37 reference the incorrect supplemental table, from the version attached to the submission, it looks like it should reference supplemental tables 6a and 6b (not 5b only). Also please report the odds ratios and p-values for these results in the text.

**Response:** Thank you for this insight. These have now been corrected as outlined in our responses to your comments 7 and 8.

**Comment 10:** Please discuss the lack of an a priori power analysis in the limitations section if one was not completed.

**Response:** This has now been added to the limitations section.

**Comment 11:** In table 2, surgeon and anaesthesiologist volume is noted “per year” but the text on page 7, describes the volume as: “total number of procedures performed since the inception of ICES databases in 1991 until the date of the index procedure.” Apologies if I’m misinterpreting table 2, but the case volumes for each surgeon or anaesthesiologist in table 2 seem high if they are actually cardiac case volume “per year” since the median for an attending anaesthesiologist would be over 2 cardiac cases per day for every day of the year. If this is general case volume instead of cardiac case volume, or not per year, please clarify this in the text or in a footnote on the tables.

**Response:** Thank you for pointing this out. The statement “total number of procedures performed since inception of ICES databases in 1991 until the date of the index procedure” correctly describes surgeon and anesthesiologist volume. We have now corrected this inconsistency by removing “per year” from the tables and have added the footnotes as the reviewer has advised.

## **REVIEWER 2 COMMENTS**

**Comment 1:** Abstract:

- Background is unclear, better to state the existing literature (e.g. optimal performance in OR crisis situation)
- Objective- need to include the main exposure variable that makes this study unique, “discordance”

**Response:** Thank you for this suggestion. We have revised the abstract as follows:

**“Background:** Effective teamwork between anaesthesiologists and surgeons is essential for optimizing patient safety in the cardiac operating room (OR). While many factors may influence the relationship between these two physicians, the role of sex and gender have yet to be investigated.

**Objectives:** We sought to determine the association between cardiac physician team sex discordance and patient outcomes.”

**Comment 2:**

Methods:

1. Why did the authors exclude re-do cardiac procedures, which represent a greater/easier definition of case complexity for subgroup analyses.
2. Why was pulmonic valvular repair excluded?
3. Do you have information of cardiac fellowship training? Do we have information of anesthesia sub-specialization (e.g. cardiac anesthesia)?
4. Please expand on how you adjusted for surgeon volume. I am confused why total volume of

procedures was used for adjustment (e.g. anything the surgeon uses) dating back to 1991 until index procedures. This is as opposed to focusing on the procedures of interest to this study (CABG, MV repair etc). Further, the most important aspect of volume is that which is contemporaneous to the index procedures – what the volume is for the surgeon (e.g. the year before,) doing the procedures of interest before doing the procedure of interest.

5. I would be tempted to consider a composite outcome of any Mortality or MACE in light of a negative results for these alone. Is this a power issue?

**Response:** Thank you for these detailed comments.

1. Redo cardiac surgery was not excluded from the analyses. These were not automatically included in the “complex” surgery category as the actual complexity of redo cases depended on cardiac anatomy and the type of planned procedure. Redo procedures are also few in number, especially with the advent of transcatheter valvular procedures.

2. We studied the impact of physician sex discordance in patients undergoing routine cardiac surgery. Pulmonic valve surgery is most often performed in patients with congenital heart disease, a specialized patient group whose baseline profiles are not comparable with the rest of the cardiac surgery population. In addition, these cases are often performed by few dedicated congenial surgeons and anesthesiologists, thus making physician sex composition is somewhat fixed.

3. We unfortunately do not have information on fellowship/subspecialization training.

4. We appreciate the reviewer’s concern. However, the procedures of interest in this study are routine cardiac cases. Using total procedure volume since 1991 (i.e., lifetime procedure volume) provides an indication of the overall technical and clinical experience of the surgeon, which is an important consideration especially in combination with number of years in practice (i.e., a new graduate may have a large procedure volume in their first year of practice but is relatively inexperienced overall, whereas a seasoned surgeon scientist could have low annual volume that is well compensated by many years of operative experience).

5. We agree with the reviewer and have conducted a post hoc analysis as the reviewer suggested, using the composite outcome of death and MACE. This is now described in the results section:

“We conducted a post-hoc power analysis to determine whether the lack of observed between group mortality difference was due to the small number of outcome events. Using logistic regression with a sample size of 79,862 patients (24.9% treated by sex discordant surgeon-anesthesiologist pairs) and an observed OR of 0.93, we were able to achieve 19% power at a 0.05 significance level. At the request of the reviewers, we repeated our analysis for the composite endpoint of death and MACE. The findings of this post hoc analysis also did not reach statistical significance (adjusted OR, 0.96 [0.88-1.05], p=0.37; Supplemental Table 7).”

**Comment 3:** Interpretation:

1. The authors report, interestingly, that an all-male physician and anesthesiologist team had a longer hospital length of stay than a sex-discordant or all female team. They also found a longer hospital length of stay for patients when individual sex was considered for both male anesthesiologists and male surgeon. Though these differences are statistically significant, their clinical significance is not commented on. The rate ratio, for example, of male anesthesiologist v. female anesthesiologist in terms of longer hospital length of stay is 1.02. Authors do not comment on

the potential clinical impact of this difference.

2. Further to this point LOS is a process measure and not necessarily an outcome. How do we know we are not just measuring differences in practice patterns of sex-discordant pairs?

**Response:**

1. The reviewer raises a great point. We have now made revisions in the Discussion:

Key Findings: “[...]2) Patients who underwent isolated CABG experienced longer hospital LOS when treated by an all-male physician team as compared to an all-female team; 3) When examining the impact of individual physician sex, the length of hospital stay was clinically and statistically significantly longer when procedures were attended by a male surgeon.”

Interpretation: “We observed clinically and statistically significant longer lengths of hospital stay in those treated by all-male surgeon-anaesthesiologist teams as compared to all-female teams, as well as individually by male surgeons.”

**Comment 4:** Conclusion: Please align your conclusions only with the objective findings of your study.

**Response:** We have revised the conclusion in both the abstract and manuscript as follows:

“Patient mortality and length of stay after cardiac surgery may vary by sex discordance of the attending surgeon-anaesthesiologist team. Further research is needed to examine the underlying mechanisms of these observed relationships.”

**VERSION 2 – REVIEW**

<b>REVIEWER</b>	Satkunasivam, Raj Sunnybrook Health Sciences Centre
<b>REVIEW RETURNED</b>	03-Jul-2021

<b>GENERAL COMMENTS</b>	The authors have sufficiently addressed comments/concerns.
-------------------------	--