Author's response to reviews

**Title:** Reporting quality of statistical methods of surgical observational studies: protocol for systematic review

**Authors:**

Robert Wu (rowu@ottawahospital.on.ca)
Peter Glen (pglen@ottawahospital.on.ca)
Tim Ramsay (tramsay@ohri.ca)
Guillaume Martel (gumartel@ottawahospital.on.ca)

**Version:** 2  
**Date:** 8 May 2014

*Author's response to reviews:* see over
Ottawa, May 8th, 2014

Prof. Mark Rodgers  
Associate Editor, Systematic Reviews  
CRD, University of York, UK

Re: MS 1516731886119555 “Statistical quality of surgical observational studies in medical and surgical journals: protocol for systematic review”

Dear Prof. Rodgers,

Thank you for considering our manuscript for publication in Systematic Reviews. We appreciate the opportunity to revise our submission, and to respond to the reviewers’ comments. In addition, we have complied with all editorial requests. Please see below for an itemized response:

Reviewer #1 (Mark Rodgers): Thank you for your comments.

1. We have substantially modified the methods to add clarity.

2. a) The main outcome of interest is the reporting and quality of statistical analysis in observational studies pertaining to interventions in general surgery (ie. “surgical observational studies”). As a secondary analysis, we intend to determine whether significant differences exist between such studies published in leading surgical and medical journals. That being said, the main focus of this review remains surgical observational studies as a whole; we simply hypothesize that there currently exists a substantial difference between the two types of journals. We have clarified this in the abstract, objective, methods, and discussion.

b) We have expanded on the secondary outcome descriptions to include the frequency and type of statistical tests used in medical and surgical journals, the most often reported and missed items within statistically significant results of the study, as well as potential correlation between impact factor and overall/item-wise score.

c) For sensitivity analysis, we propose to remove two medical and surgical journals with the fewest published surgical observational studies and repeat the analysis. We hypothesize that eliminating these journals would lead to improvement of overall reporting quality.

d) We hypothesize that the overall quality of statistical reporting in surgical observational studies will be poor. We also expect that certain types of analyses will be utilized extremely frequently (eg. logistic regression), but that important steps
associated with these analyses will be lacking more often than not. We further hypothesize that our study will show that surgical observational studies published in surgical journals are of lower score in comparison to the medical journals, thus raising caution among readers when appraising observational studies published in surgical journals. This study could be used by journal editors and authors to improve the quality of statistical reporting in observational surgical studies. Some of these points have been added to the manuscript.

3. We have modified the description. All articles will be 1) grouped according to the statistical analysis used, and each quality criterion will be presented as a percentage of articles that fulfilled the criteria given the analysis. 2) Each article will be scored as a percentage of fulfilled criteria in the relevant statistical domain used in study. The score represents a non-weighted score (with each study acting as its own control) and this will be made clear to readers.

Reviewer #2 (Forough Farrokhhyar): Thank you for your comments.

1. Abstract:
   a) This has been corrected (“bias” has been removed).
   b) An objective statement has been added to the text.
   c) This has been corrected (“methodological principles” has been removed)
   d) This has been corrected (“inherent biases” has been removed)

2. Background:
   a) The main focus of this review remains surgical observational studies as a whole; we simply hypothesize that there currently exists a substantial difference in terms of reporting and quality between surgical observational studies published in medical and surgical journals. While this has been shown for trials, we are not aware of such a comparison (pertaining specifically to statistical methods) for observational studies. Because surgeons rely almost very heavily on observational studies published in their own specialty journals, we argue that this is a highly relevant analysis in order to try to affect change. We have clarified this in the abstract, objective, methods, and discussion.
   
   b) Surgeons are largely exposed to observational studies in surgical journals. We believe that if a significant journal-wise difference could be demonstrated, we could caution surgical readership of the potential pitfalls in interpreting the statistical analysis in surgical journals.

3. Methods - Study selection: We have revised the study selection process to include two reviewers. Each reviewer will screen one month for each journal to validate the screening strategy. If there is > 90% agreement, the search strategy will be considered valid.

4. Outcomes:
a) The main objective of the study is to examine the quality of statistical reporting. SAMPL provides a useful backbone to do so. As mentioned in the text, essential elements of SAMPL have been included in Appendix A, in addition to other elements (e.g. propensity analysis). Three authors (including an experienced statistician) have agreed upon a minimum set of criteria to evaluate the quality and statistical reporting of observational studies, and these were included in the current protocol. Given this adapted tool contains items derived from an existing guideline, we argue that the validity of the tool is retained. Moreover, it is worth noting that SAMPL, while authored by prominent statisticians and methodologists, has not been formally validated nor published in a peer-reviewed journal (to the best of my knowledge). Finally, we have also contacted Prof. Altman for additional guidance, but we have not heard back from him.

b) We agree with you that design is closely related to statistical quality. However, the main purpose of this paper is to examine statistical reporting and quality in observational studies. Surgeons who publish observational studies tend to publish large case series, with little true variation in study design. Most attempt to account for confounding using statistical analyses. For this reason, we argue that understanding the quality, reporting, and validity of these analyses is much more useful and will be the focus of our study.

c) We have clarified our planned scoring scheme. All articles will be 1) grouped according to the statistical analysis used, and each quality criterion will be presented as a percentage of articles that fulfilled the criteria given the analysis. 2) Each article will be scored as a percentage of fulfilled criteria in the relevant statistical domain used in study. The score represents a non-weighted score and this will be made clear to readers.

5. Data analysis:

a) We have now clearly stated our hypothesis.

b) Comparisons between the medical and surgical journals will be made using a Chi-square for each domain. In addition, we intend to identify variables associated with high-quality reporting of statistical analysis. To do so, we will dichotomize the cohort of studies on the basis of the 50th and/or 75th percentile of the proportion of fulfilled criteria. We will then construct a logistic regression model to identify variables that are independently associated with high-quality reporting. All variables with a p<0.2 on univariate comparison between high and low-quality reporting will be included in the model. The following minimal set of variables will be compared: journal name, impact factor, medical/surgical journal, continent of origin, sample size, disease category, type of exposure, type of primary analysis. We intend to check for interaction between variables, as well as colinearity.

c) We have modified the sensitivity and subgroup analyses. Subgroup analysis now consists of comparing studies with higher reported strength of association (i.e. relative risk of > 2 or < 0.5) between exposure and outcome. Sensitivity analysis will be done by eliminating two medical and two surgical journals with the fewest published surgical observational studies. We hypothesize this will improve the overall quality of reporting.
6. Discussion: The limitations have been added to the discussion section.

7. Appendices: We have added preliminary analysis. The intent of analysis domain contains elements that are found in the “Reporting hypothesis testing” including “state a priori objective” and “report whether and how any adjustments were made for multiple statistical comparisons”. We have chosen to exclude Bayesian analysis and included propensity analyses instead. This was a group consensus.

8. Title: we have modified the title to reflect reporting quality.

Thank you for your consideration and interest in our manuscript. Yours, sincerely,

Guillaume Martel, MD, MSc (Epi), FRCSC
Associate Scientist, Ottawa Hospital Research Institute
Attending Surgeon, Surgical Oncology Program, Ottawa Hospital – General Campus
Assistant Professor of Surgery, University of Ottawa