INVITED FEATURE: STATISTICS PRIMER

The review process from a biostatistics perspective

Claire Cameron, Ari Samaranayaka, Robin Turner

We would like to provide an insight on how biostatisticians might review a journal article. We are writing this for two reasons:

1. You may intend to submit an article to a journal for publication. We hope this primer helps you avoid common pitfalls and mistakes in your write-up.
2. You may be asked to review an article yourself. We hope this article gives guidance on things to look for and avoid.

Most journal articles are peer reviewed (this primer has undergone editorial review). That means at least one expert reads over the article to assess its quality and suitability for publication. Those experts may be someone from the same discipline, a related discipline, or (because of the challenges of finding reviewers) a person not closely tied with the discipline. Sometimes the expert may be a biostatistician who is, often, not a content expert but is an expert in research methods and analyses. The reviewer can then recommend the article be (a) accepted for publication as is, (b) revised (either major or minor) or (c) rejected. If there has not been a statistical review, there may be an option to recommend one. Ideally, from a biostatistical viewpoint, all quantitative studies should have a *good quality* statistical review. A paper from the British Medical Journal (BMJ) in 2002 highlighted some problems of peer reviewing that are still relevant.[1] These included reviewers unjustifiably criticising statistical methods, making criticism for its own sake, and imposing a reviewer's personal statistical dogma in a manner inappropriate to the context.[1]

Some publishers provide guidance for their reviewers. For example, Wiley provides extensive advice to reviewers (aimed at content experts).[2] It may be interesting and useful to look at these guidelines as an author for information on how to structure and present your paper.

What is the difference between a content review and a statistical review? Very simply, the person providing the content review will know how important the research question is in that field of study and make a judgement on how well that question has been answered. The statistical reviewer will look at how appropriate the study design and statistical methods were for answering the question and how well they were applied. Consequently, they too can judge how well the research question has been answered. Understanding this difference in focus is useful for both an author and a reviewer. It encourages authors to describe all aspects of their study clearly, which consequently helps content reviewers decide when to send articles for statistical review.

When undertaking any type of review, understanding the overall design of the study is critical. Is there a standard design that you can classify the study into — for example, a case-control study, a randomised control trial (RCT), or a cross-sectional study? The EQUATOR network[3] has a website that is a great resource, bringing together reporting guidelines for different study designs. If a study falls into one of these designs, the article should adhere to the appropriate reporting guidelines and checklist (if applicable). For example, RCTs are unlikely to be published in high-impact journals if they do not adhere to the Consolidated Standards Of Reporting Trials (CONSORT) statement and associated checklist. Be careful though; we have seen reviewers misunderstand the study design and make demands of authors that cannot be addressed because they are not applicable.

We recommend using the reporting guidelines when designing a study. Guidelines clarify what information should be included when reporting study results and they help to make sure that nothing important is missed at the design stage. Table 1 shows some of the common study designs with their reporting guidelines.

Table 1: A summary of the study guidelines contained in the EQUATOR Network.[3]

Common study designs	Reporting guidelines
Observational Studies:	
Cross-sectional Survey	STROBE
Case-control	STROBE
Cohort	STROBE
Experimental Studies:	
Randomised Control Trial (RCT)	CONSORT statement
Feasibility/Pilot (for trials)	CONSORT extension for pilot and feasibility studies
Other Types of Studies:	
Animal Pre-clinical Studies	ARRIVE
Systematic Reviews	PRISMA
Study Protocol	SPIRIT
Diagnostic/Prognostic Studies	STARD, TRIPOD
Case Reports	CARE
Clinical Practice Guidelines	AGREE
Qualitative Research	SRQR
Quality Improvement Studies	SQUIRE
Economic Evaluations	CHEERS

In this next section, we will look at the different elements of a paper and discuss what we expect to see as biostatistical reviewers.

Abstract

Looking carefully at the Abstract is the starting point for forming a view on the quality of the study:

- What is/are the objective(s) of the study? There should be a research question or a collection of research questions.
- Are the results and conclusions consistent with the research questions?

› Is the study design clearly identified? Importantly, the study design should be suitable for the stated research question(s).

If the Abstract is not clear or logical, it is possible that the paper that follows will not be either.

Introduction/Literature/Background

This section of the paper is to provide the rationale for the study. A biostatistician would like to see the research question(s) and objectives listed clearly here (usually in the last paragraph, probably as a repeat of what is in the Abstract with more details, along with any supplementary objectives) and find that they follow logically from the description of the background. These objectives are critical — they determine the study design, the outcome measure(s), other measures of interest, and the appropriate analysis plan. These questions must be answered through the analysis, in the Results section and in the Discussion and Conclusion.

Methods

In an earlier statistics primer,the authors talked about sample size estimation and where it fits in the research process.[4] At the beginning of that article, they presented a diagram showing, in part, that the study design is entirely dictated by the research question. The sample size calculation depends on what the study design is and what the analysis is going to be. This may lead back into refining the research question. All these elements should have been decided upon before any data were collected.

STUDY DESIGN

The study must be designed to answer the research question. For example, if assessing the effect of a treatment, an RCT may be the best way to answer this question. Estimating the prevalence of a particular disease at one point in time will probably require a cross-sectional survey (to determine the proportion of the population that has the disease). The specifics of the design must be outlined in the Methods section. Ideally, another researcher should be able to replicate the design if they wanted to do a similar study.

SAMPLE SIZE

The power calculation (if there is one) should be clearly described including all assumptions used with appropriate referencing. A biostatistician should be able to replicate the sample size estimate from reading this statement. This calculation must happen *before* the study begins as part of the design. Some study designs require a sample size estimation, for example RCTs. Some studies may not be able to control sample size and therefore sample size estimation is irrelevant (for example, some analyses of large routinely collected datasets). Common sense is needed here.

ANALYSIS PLAN (THE STATISTICAL METHODS SECTION)

Whilst this will have been drafted at the study design stage as an analysis plan (with some studies publishing this in a protocol or even separately as a paper in its own right), the Statistical Methods section is often (re)written for publication and should describe the statistical methods used to obtain the results. This should be in the same order as the results are presented so that the Results are easy to follow. This section must show that the analysis is appropriate for (1) answering the research question, (2) the design that has been described, (3) handling unintended data issues experienced (if any), and (4) explaining any subjective decisions taken.

Problems that we commonly see in this part:

1. The analysis is not appropriate for the design. For example, a study design may include the sampling of schools and then further sampling of children within those schools. An analysis at the child level, not allowing for the clustering within schools, would not be appropriate for the design.
2. The analysis does not provide an answer to the research question. For example, the objective is to measure size of a treatment effect, but hypothesis testing is undertaken with only p-values reported (p-values do not measure the size of the effect).
3. The analysis described is not what the sample size estimate was based on. For example, a t-test is used to estimate sample size, but the analysis uses a complex longitudinal model with no simple t-test at the primary time of interest.
4. The analysis is not described clearly. For example, lack of clarity as to what the outcome is, what specific model has been used, or which covariates are included in which models. Other researchers should be able to replicate the analysis.
5. The analysis includes elements that have come up as interesting (because the data is there) but is not answering the research question. For example, reporting subgroup analyses because the data has been collected, not because it answers the research question.
6. Providing an ambiguous description of the methods when a simple clearer description is possible. For example, saying "multivariate models were used to test the significance" tells us nothing about the type of model used nor the hypothesis tests used. An example of a more specific version that is clearer might be along the lines of "the difference between the treatment group and control group was estimated in a regression model with BMI at 12 weeks as the outcome, with potential confounders (list them) as covariates in the model."
7. Including results in the Methods section. It can be difficult to separate the methods from the results but 'how things were done' should be in the Methods section and 'what was actually found' should be in the Results section.

The Statistical Methods section of a paper should directly address the objectives and explain how the presented results were obtained. They are critical to the reproducibility of the study and should not leave out key information. The statistical reviewer should be able to follow the subjective steps taken in implementing the statistical methods and the reasons for them; they are likely to query any lack of clarity.

Results

If the Methods have been clearly written, the Results should be easy to follow as it simply explains everything that was found. Importantly, the Results section should only contain observed results, not their interpretation (that belongs in the Discussion).

Do the results reflect the analyses that have been done? Do the numbers look sensible? This is bearing in mind that studies do throw up quirky results at times. Even so, there should be some face validity to them or discussion of the quirkiness later in the Discussion section. When reporting estimates, is there an estimate of precision (for example, standard deviation, standard error, or confidence intervals)? Are the results simply based on p-values whilst ignoring effect sizes and their clinical significance?[5]

If reporting guidelines are used, there is often a checklist that can provide guidance on what information should be included. There are plenty of study design-specific items to look for within the results which we have not covered here.

Discussion

An interpretation of the main findings should appear at the start of the Discussion. The interpretation needs to be justified by the results and it needs to answer the research question. The interpretation cannot go beyond the meaning of the analysis of the data that is available. A well-structured Discussion section is described by Skelton and Edwards.[6]

Investigations are likely to be imperfect in most situations. For that reason, it is critical that the limitations are clearly explained, along with the impact of those limitations on the findings. Good research is done the best possible way (within practical constraints), uses appropriate methods, clearly states the study limitations, and draws conclusions that allow for those limitations.

General Comments

We have tried to provide guidance for authors and reviewers. As a reviewer, asking questions is a useful strategy to highlight issues, particularly when it is unclear what the authors have and have not done. As an author, questions from reviewers are there to highlight areas that might need further clarification. If a reviewer misunderstands something, the author has not been clear in their writing. The reviewers are never asking for a personal explanation of the work; rather, the authors usually need to change the manuscript in response to the reviewers' questions.

As a reviewer, your questions and suggestions to authors should help improve the manuscript by encouraging clear reporting of the methods and results. Consider whether they would help you to improve the manuscript if you received them as the author. Remember that the purpose of the review process is to improve the quality of published research.

When reviewing a manuscript, you are dealing with people so you should be courteous and polite. This is regardless of whether the review process is open or not. As suggested above, asking questions is a good way to engage in the process and will likely get the best out of the authors. We have also found it useful to think about whether your concerns with the paper are major or minor. Are the concerns so major that you do not believe the results anymore or are they more minor where you feel the results still stand but require more fine tuning? The balance of major versus minor concerns will help you in rating the article as suitable for publication (after revision) or not.

No reviewer should be assessing work beyond their expertise. We notice this acutely, as biostatisticians, where we cannot fully assess the content area but are focussed on the biostatistical elements. We can ask common sense questions, but it would be difficult for us to make strong statements on the content. We could, for example, critique whether the Introduction has explained the gap adequately, but not if there was relevant literature missing nor if an intervention was novel. If the area of work is outside your expertise, you can decline the invitation to review. If you feel the statistical aspects are beyond your understanding, you should recommend a statistical review.

In summary, we have provided some insight into what biostatisticians are looking for in a review. We hope this is helpful to you as an author and as a reviewer. Sometimes, a manuscript looks quite good, but by the time you have made all your suggestions, the recommendation becomes a major revision. Hopefully, though, the revision improves the paper, thereby improving the standard of health-related research, and that is a good thing for the world.

References

1. Bacchetti P. Peer review of statistics in medical research: the other problem. BMJ. 2002 May 25;324(7348):1271-3.

2. Wiley Author services [Internet]. Step by step guide to reviewing a manuscript [cited 2022 Sept 2]; Available from: https://authorservices.wiley.com/Reviewers/journal-reviewers/how-to-perform-a-peer-review/step-by-step-guide-to-reviewing-a-manuscript.html

3. The Equator Network [Internet]. Oxford (UK). Reporting Guidelines [cited 2022 Sept 2]; Available from: https://www.equator-network.org/reporting-guidelines/

4. Samaranayaka A, Cameron C, Turner R. Sample size in health research. New Zealand Medical Student Journal. 2021 Apr 22(32):52-4.

5. Cameron C, Samaranayaka A, Turner RM. P values: what is their significance? New Zealand Medical Student Journal. 2020 Sep 11(31):48-9.

6. Skelton JR, Edwards SJ. The function of the discussion section in academic medical writing. BMJ. 2000 May 6;320(7244):1269-70.

About the authors

› Associate Professor Claire Cameron, BSc(Hons), DipGrad, MSc, PhD, is a Biostatistician in the Biostatistics Centre, Division of Health Sciences, University of Otago.
› Dr Ari Samaranayaka, BSc, MPhil, PhD, is a Senior Research Fellow and Biostatistician in the Biostatistics Centre, Division of Health Sciences, University of Otago.
› Professor Robin Turner, BSc(Hons), MBiostat, PhD, is the Director of the Biostatistics Centre, Division of Health Sciences, University of Otago.