Review article
What a reviewer wants

ANDREW J. DAVIDSON MBBS MD*†‡ AND
JOHN B. CARLIN BSc (Hons) PhD†
*Department of Anaesthesia, Royal Children’s Hospital, Melbourne, Australia, †Clinical Epidemiology and Biostatistics Unit, Murdoch Childrens Research Institute, and University of Melbourne Department of Paediatrics, Melbourne, Australia and ‡Department of Pharmacology, University of Melbourne, Melbourne, Australia

Summary
Understanding what a reviewer wants helps authors write better papers and design better research studies. Central to any research study is the research question. This is the first thing the reviewer looks for. The question, or research aim, should be relevant, original and very clearly defined. The paper should also be scientifically valid; the methods should achieve the aim by addressing the question, and the conclusions should match the results. The statistical analysis often attracts criticism, especially the relatively simple problems arising with the use of $P$-values and the concepts of statistical and clinical significance. There are several agreed guidelines describing exactly how studies should be reported. Following these is a great help to reviewers.

Keywords: review; statistics; pediatric anesthesia

Introduction
A fundamental part of an academic profession is the communication of research findings from one to another. The better we do this, the stronger our profession and the better care we give to children. Journals such as this one, and the papers that make up this journal, are an important part of this knowledge transfer. Papers change clinical practice, and a salient article in this journal will be read or downloaded well over a thousand times per year. The quality of papers is therefore not only important for the journal, it is important for the children under our care.

One way of ensuring quality is peer review. Editors ask reviewers to assess a paper because of the reviewers’ particular knowledge and to provide a more balanced assessment of the paper’s worth. While good reviews can substantially improve papers, we acknowledge that peer review is not a perfect process, and at times can be frustrating. The aim of this paper is to highlight what reviewers look for. There is no universal formula for reviewing, and reviewers all have their own particular nuances and styles. While this paper cannot be exhaustive, it will hopefully help future authors design better studies, write better papers and facilitate their smooth and speedier passage through the review process.

The essentials of a publishable paper
For a scientific paper to be publishable it must fulfil three criteria. It must be relevant, original
and scientifically valid. This basic trio of requirements recurs in many aspects of clinical research. A paper must also be clear and well written. However, while poor English can be corrected in the review or editing process, relevance, originality and validity cannot be corrected. If any are in substantial doubt, a paper cannot be published. When summarizing their opinion for the editor, the first things a reviewer addresses are relevance, originality and validity (1).

Relevant
There is no point in answering a question that nobody is asking. In clinical research the most relevant, or important, papers are those that change clinical practice. For basic science papers, the most relevant papers are those that add significantly to knowledge in areas of current scientific interest, open up new areas of investigation or move basic science closer to answering important clinical questions. The more important the related clinical questions, the more likely the basic science will be relevant. Relevance and importance are linked to a paper’s impact. A high impact paper concerns health issues that affect large numbers of people, have a large effect on quality of life, have an appreciable mortality or incur a large economic cost.

Original
The research must be novel. There is no point in answering a question that has already been answered. If a study is only marginally different to a previously published study then there must be something in particular that justifies publication: for example a different population, intervention, or outcome measure. There must also be a plausible argument that the difference would produce different results. There is a trade-off between originality and relevance. When originality is marginal there is a better chance of publication if the importance is great. Indeed in some circumstances the importance may be so great as to warrant publication of a near-identical or replication study. For example, awareness under anesthesia is an important issue so several studies investigating incidence have been published.

Valid
A valid paper is one where there is a logical, justifiable and clear progression from research question to conclusion. Do the methods produce results that can answer the research question? Are the conclusions drawn from the results justifiable? To assess validity, the reviewer will assess the methods, including where relevant selection of subjects, randomization, blinding, use of controls and interventions, outcome measures and analysis.

The research question
The central component of any paper is the research question. The first thing a reviewer asks themselves is ‘what is the question?’ A good question is a question that is clearly defined and whose answer is unknown, relevant and knowable. A question can be defined in terms of ‘PICOT’ – population, intervention or exposure of interest, comparator (if applicable), outcome and timeframe. The question forms the basis for the aim or objective of the paper. A paper that appears confusing or difficult to follow is often a paper than does not have a clearly appreciable research question. Nearly every paper presented for review has a research question somewhere; sometimes it is just hard to find it.

Structure
Papers, except for narrative reviews, are usually divided into abstract, introduction, methods, results and discussion. Case reports have a case description rather than a methods and results section. Systematic reviews should also have a methods and results section. In a systematic review the author seeks to answer a particular question by searching the literature and synthesizing the findings. The methods describe the method of search and systematic evaluation.

There are agreed general guidelines on how to prepare a paper (http://www.icmje.org) and specific guidelines for how to report studies of various types (Table 1). These guidelines give specific checklists to ensure that important areas are included. Increasingly, journals request that authors follow these guidelines and use the checklists accompanying these guidelines.
The title

The title of the paper should be concise and yet accurately describe the content of the paper. The words in the title need to be carefully chosen to ensure they are caught by search strategies and the syntax carefully chosen for accurate meaning. The title should be a label and not a sentence stating the result. Such sentences tend to be excessively long and may sound dogmatic.

Abstract

Many readers will only ever read the abstract. The wording should be perfect and particular care should be taken not to make unwarranted conclusions. Exceeding the word limit (usually 250 words) may result in the end being chopped when the abstract is captured in some electronic databases.

Introduction

In many cases the introduction should be short; about 250–500 words. It should not be an in-depth literature review. The purpose of the introduction is to clearly and simply explain what the research question is, why it is relevant, why it is original and very briefly how it will be answered. There are various ways this can be performed; one way is to write what we know, what we do not know, what the gap in knowledge is, why it is important to fill that gap and then lastly how the study will fill that gap. The introduction must include the study aim or objective. If there are several aims then the primary one should be defined. Remember that by definition there can only be one primary aim. An excellent paper has an obviously important and original question, and therefore needs only brief introduction. If the introduction is long, requiring many paragraphs to justify the relevance or explain the originality, then the chances are that the relevance or originality is marginal.

Methods and statistics

This section should describe exactly how the study was performed. Ideally it should avoid excessive detail but contain sufficient information to allow the reader to repeat the study.

The methods section must describe the statistical analysis, in sufficient detail to enable the reader to follow what was performed. Most deficiencies in statistics are not in the complex analyzes but simple errors or omissions. The author should include how summary statistics are to be presented (for example mean or median, standard deviations or interquartile ranges, frequencies, proportions, etc.). For inferential statistics the reader should be able to understand what analysis was performed on which data.

Registration

Many randomized trials are registered with a central registry before they start. Indeed in the European

---

Table 1

Guidelines for reporting studies

<table>
<thead>
<tr>
<th>Study type</th>
<th>Guideline</th>
<th>Acronym</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomized trials</td>
<td>Consolidated standards of reporting trials</td>
<td>CONSORT</td>
<td>Moher et al. (4) and Altman et al. (5)</td>
</tr>
<tr>
<td>Meta-analyses of randomized trials</td>
<td>The quality of reporting of meta-analyses</td>
<td>QUORUM</td>
<td>Moher et al. (6)</td>
</tr>
<tr>
<td>Diagnostic studies</td>
<td>The standards for reporting of diagnostic accuracy</td>
<td>STARD</td>
<td>Bossuyt et al. (7)</td>
</tr>
<tr>
<td>Observational studies in epidemiology</td>
<td>Strengthening the reporting of observational studies in epidemiology</td>
<td>STROBE</td>
<td>von Elm et al. (8)</td>
</tr>
<tr>
<td>Meta-analyses of observational studies</td>
<td>Meta-analysis of observational studies in epidemiology</td>
<td>MOOSE</td>
<td>Stroup et al. (9)</td>
</tr>
</tbody>
</table>

---

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd, Pediatric Anesthesia, 18, 1149–1156
Union, Australia and New Zealand this is a requirement. There are now several registries available (Table 2). What needs to be registered? Sponsored drug trials certainly need registration but increasingly other types of study are also being registered. If in doubt it is better to err on the side of registration. Regardless of local requirements, some journals will not publish a trial that has not been registered. Although registration was originally intended to increase awareness of unpublished sponsored negative trials, registration has also provided an opportunity for the reviewer and reader to check that the researchers performed the study in the way originally planned. This is particularly useful to ensure that the reported primary outcomes and analysis were as intended, and that analysis plans have not been changed to deliver preferable findings.

Results

This section should start by indicating how many subjects were recruited and how many successfully completed the study. All subjects should be accounted for including those enrolled but not studied, those withdrawn and protocol violations. Traditionally, demographic data to describe the subject population is presented next. If there are randomized groups, descriptive statistical comparisons of demographic data should be presented. Hypothesis testing should not be applied to randomized baseline data. The main outcome data should be described in the text, with the primary outcome before secondary outcomes. Figures and tables provide more detail. The text may bring attention to key data in the figures or tables but there should not be extensive duplication between figures, text and tables. Reviewers are particularly annoyed when numbers in tables do not add up, or if there are inconsistencies between text, tables and figures. Symmetrically distributed data should be summarized with a mean and standard deviation. Data with a skewed distribution should generally be summarized with a median and inter-quartile range, although summaries based on a logarithmic scale (e.g. geometric mean) may be useful. For important outcomes the raw data should be presented if possible. Continuously distributed outcome measures may be displayed by group using ‘dotplots’ (sometimes known as ‘bubble-plots’), or box and whisker plots for larger sample sizes.

There is one area relating to statistical analysis where reviewers may be particularly critical: the use of P-values and the concept of statistical significance. Precise use of methods and terms in this area is important because seemingly minor errors can easily lead to unjustified conclusions and hence wrong clinical decisions. The term ‘significant difference’ often causes confusion, and if possible authors should avoid it. There are two important aspects to a difference – the magnitude of the difference and the uncertainty surrounding the reported difference as an estimate of the actual difference in the larger population. The magnitude of the observed (‘estimated’) difference may be clinically significant or not (for example a difference in recovery time of 2 min is of little clinical significance). However, a confidence interval for the difference is essential to appreciate the uncertainty due to chance variation. The difference observed would not be replicated precisely in another study and may not be a reliable indication of the true difference that would be found in a much larger study.

P-values are widely misinterpreted. A P-value can only be correctly interpreted in relation to the null hypothesis underlying the corresponding test. The P-value is the chance that a difference as large as (or larger than) that observed in the study sample could have arisen if the null hypothesis is true. This is often paraphrased as the probability that the difference was ‘purely due to chance’ but this loose interpretation can mislead. A P-value of less than a particular predetermined value, usually 0.05, is generally referred to as ‘statistically significant’.

<table>
<thead>
<tr>
<th>Registry</th>
<th>Web address</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Australian Clinical Trials Registry (ACTR)</td>
<td><a href="http://www.actr.org.au/">http://www.actr.org.au/</a></td>
</tr>
<tr>
<td>The US National Library of Medicine Randomized Controlled Trial Number Registry</td>
<td><a href="http://www.clinicaltrials.gov/">http://www.clinicaltrials.gov/</a></td>
</tr>
<tr>
<td>The International Standard Randomized Controlled Trial Number Registry</td>
<td><a href="http://www.controlled-trials.com/">http://www.controlled-trials.com/</a></td>
</tr>
<tr>
<td>The National (UK) Research Register</td>
<td><a href="http://www.update-software.com/national/">http://www.update-software.com/national/</a></td>
</tr>
<tr>
<td>European Clinical Trials Database</td>
<td><a href="http://eudract.emea.eu.int/">http://eudract.emea.eu.int/</a></td>
</tr>
</tbody>
</table>
but it is important to be aware that the choice of the threshold 0.05 is entirely arbitrary. It is highly misleading to use $P < 0.05$ as a benchmark indicating proof or not (2). A better approach is to interpret the $P$-value as indicating the strength of evidence against the null hypothesis, implying evidence that a real difference exists in the larger population. For example, $P = 0.01$ indicates stronger evidence for a true difference than $P$ around 0.05, which might be referred to as weak or marginal evidence for a difference. The choice of ‘strong’ or ‘weak’ is still arbitrary but avoiding the term ‘statistically significant’ helps the reader to understand that all is not black or white and it also avoids confusion with clinical significance. Using this terminology may appear cumbersome at first, but with practice most researchers and readers will prefer it, and as it more faithfully describes the findings, find it far more satisfying. The word ‘trend’ is occasionally used to describe findings where $P$-values are slightly higher than 0.05. However, the term ‘trend’ should be reserved for a monotonic function of time or dose, and should not be used to describe a $P$-value.

The pitfalls of statistical significance are particularly apparent when authors interpret findings where there is little or no evidence for a difference. In particular, finding no evidence against the null hypothesis by no means implies it is true, i.e. does not demonstrate equivalence. For this reason, a confidence interval around an estimate should be reported. The confidence interval (conventionally using the 95% level) gives the reader information on the range of values within which the true difference may lie, taking account of sampling variability, and so it should be interpreted as a whole with respect to clinically meaningful differences (Table 3). Note that the choice of 95% ‘confidence’ is still an arbitrary cut off, analogous to the traditional $P < 0.05$. Note also that the choice of what is clinically significant may also be subjective; what is clinically significant to one anesthetist may not be clinically significant to another. Reporting the 95% confidence intervals helps the reader decide for themselves the clinical meaning of the result.

A further common problem in statistical analysis is the reporting of results from multiple statistical tests. The more tests that are performed, the more likely that evidence for difference will emerge purely by chance. To avoid this, authors should plan exactly what single test will be performed on the primary outcome before starting the study. Adherence to this can be checked by the reviewers by going to the trial registration website. Other tests and other outcomes can of course still be reported, but less weight should be attached to these findings. If there are multiple outcomes and tests, with no single a priori defined primary outcome, then the conclusions will be less compelling, especially if it appears that the presentation has focussed on one or two results that ‘stand out’ posthoc (3). Note that a primary outcome cannot be defined after the study has been completed.

**Discussion**

The discussion is perhaps the least important part of the paper, but often the hardest to write. There is no set format and the discussion must suit the type of study and findings. In general the discussion should start with the answer to the research question: in other words, a very concise indication of what was found and how this fits with the study aim. Results

<table>
<thead>
<tr>
<th>$P$-value</th>
<th>95% Confidence interval (CI) for the difference</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evidence for a difference</td>
<td>$&lt;0.05$</td>
<td>The interval does not cross zero</td>
</tr>
<tr>
<td>Evidence for equality</td>
<td>$&gt;0.05$</td>
<td>The interval crosses zero but the limits of the interval do not include what would be regarded as clinically relevant values</td>
</tr>
<tr>
<td>Inconclusive</td>
<td>$&gt;0.05$</td>
<td>The interval crosses zero but its range includes what would be regarded as a clinically relevant value</td>
</tr>
</tbody>
</table>
do not need to be repeated in detail and new results cannot be introduced in the discussion.

The discussion should focus on the primary outcome and should include comments on the relevance of the findings. Authors should resist the temptation to exaggerate the significance of their findings and be careful when claiming to be the first to do something. The discussion should include how the results are original and how they relate to other studies. If there are numerous other studies, only salient points from salient studies need be included. The authors should openly discuss the strengths and limitations of their study when compared to similar studies. The study may change the way people practise so authors should honestly indicate any reservations or limitations. No study is perfect and identifying limitations increases the credibility of the paper.

Authors should stand by their data. A fundamental aspect of scientific method is that theory fits observation and not that observation should fit preconceived theory. If the results are different to what was expected from previous studies and the reason for this is unclear, then the authors should just say so and not speculate with implausible and convoluted explanations. The reasons for inconsistency (which will always include a component of chance variation) may not yet be known, and provided the methodology is sound, posterity will clarify the true effects.

If a conclusion is included it must be one or two succinct sentences restating how the findings answered the research question. New ideas should not be first introduced in the conclusion and it is not helpful to finish with the empty comment ‘more studies need to be done’!

References

References should be up to date and accurately support the statements made. A smart author knows that inevitably some of the references will have been written by the reviewers. Misquoting them does not help!

Acknowledgements

Authors should acknowledge all substantial assistance for the study. It is particularly important that all financial assistance is detailed. Authors should also indicate any conflicts of interest.

Housekeeping

Authors must always carefully follow the journal’s instructions to contributors, including the formatting of the references. Failure to do so delays publication, irritates the reviewers and editors, and indicates a lack of attention to detail that undermines the credibility of the paper. If the authors cannot be bothered to follow the instructions, then perhaps they cannot be bothered with the meticulous approach needed for good research.

Ethics and privacy

For publication, a research study must meet both local ethical standards and the standards set by the journal. The journal is guided by the Declaration of Helsinki (http://www.wma.net/e/policy/b3.htm). For research involving humans there must be a clear statement outlining the Human Research Ethics Committee or Institutional Review Board approval, and, where applicable, that subjects or their guardians gave their consent. In general, children are regarded as vulnerable with limited ability to give informed consent. This heightens the need for equipoise and the minimization of harm. Legal requirements and public expectations are increasing the need to maintain subject privacy and confidentiality. This includes both the publication of potentially identifiable photographs and medical or personal details.

Writing style

All authors have their own particular style of writing, but for scientific publications there are some generally agreed rules. Occasional use of the first person may be warranted, but in general the third person is preferable to many ‘I’s and ‘we’s. To sound neutral and impartial, scientific writing often uses the passive voice. Sentences should generally be short and clear. Jargon and abbreviations should be avoided and only generic names used for drugs. Many readers will not be native English speakers.

Although reviewers will not alter a particular author’s style, they will expect correct grammar,

Typographic errors annoy reviewers. Multiple errors can affect meaning, and reduce the reviewer’s ability and inclination to give a helpful and thorough review. Greater leniency is given to non-English speakers, while for native English speakers errors of expression can undermine the reviewer’s confidence in the paper’s credibility. Errors can be reduced by proof reading before submission. Similarly, with many drafts the authors become increasingly familiar with the paper and can lose sight of what may not be clear to others. Asking a colleague to proof read can help.

Too long

The instructions to authors indicate roughly how long a paper should be. It is common for a reviewer to comment that the paper is too long. To shorten a paper the author can check that results are not duplicated, trim excessive literature review and remove any tangential discussion about outcomes not central to the research question. Similarly, speculation about the possible mechanisms underlying a result, or possible consequences of results should be brief and limited to the important outcomes.

The response

Reviewers are not perfect. It does not help getting angry or exasperated by odd or contradictory reviews. Authors should look carefully at what a reviewer has written. Do they want clarification of a particular point so that they can make a decision, or are they suggesting a particular addition or change be incorporated into the manuscript? Reviewers should also indicate the importance of their points: what is essential, and what are merely helpful suggestions? The editor may highlight important points or try and reconcile differing reviews. Addressing the editor’s comments is particularly important. Authors do not have to make all the changes suggested by reviewers, but all comments must be addressed in the response letter. If the authors disagree with a reviewer they must clearly explain why. Ignoring a point annoys reviewers and gives the impression that something is being hidden.

Final word

Most relevant, original and valid papers will eventually make their way into print. The fastest way is to submit a carefully and clearly written paper. Knowing what reviewers are looking for should help.

Further reading


References


*Accepted 22 August 2008*