Residents' Perspective

Clare Atzema, MD

From the Royal College Emergency Medicine Residency Training Program, University of Toronto, Toronto, Ontario, Canada.

Section Editor

Clare Atzema, MD
University of Toronto, Royal College Emergency Medicine Residency Training Program
Toronto, Ontario, Canada

Presenting at Journal Club: A Guide

[*Ann Emerg Med.* 2004;44:169-174.]

INTRODUCTION

Many residents count down the months of residency training with relief, grateful for each passing journal club for which they have not been asked to present a paper. With the invitation comes the work involved in article review and critique, followed by the public display of research prowess, or lack thereof. Unless one reviews scientific papers frequently, often one can recall only a few of the nuances that constitute good research, such as a large "n" and randomization of subjects. The resident invariably makes a frenetic search of either the library or his or her own collection of research methodology papers. What follows is an article that consolidates those details in one piece of text. It is based on the instructional CD provided to peer reviewers for *Annals of Emergency Medicine*[1] and is not comprehensive, but provides enough guidance to finish the journal club with a sense of satisfaction.

0196-0644/$30.00

doi:10.1016/j.annemergmed.2004.03.035

ANALYZING THE PAPER

When reading the article for the first time, write your thoughts in the margins to help solidify a general impression of the work. On this initial read, devote particular attention to tables and charts (editors do not like to repeat table results in text, and tables convey a lot of information in a succinct format). Put the article away for a day, or even for 30 minutes if that is all the time that you can spare, to let your thoughts percolate.

The goal of the second reading is to forage details. Consider the following questions, and actually consult the article text to find the specific answer. Starting with the Introduction, ask yourself the following: (1) Does this Introduction justify the importance and originality of the paper? (2) Does this Introduction have a specific research question?[1]

The research question depends on the study type, because only certain types of research can answer certain types of questions. If the study is testing a hypothesis, the research question should include who and what is being compared, what the outcome is, and what the authors consider to be an important difference. The absence of any of these components is a significant flaw.[1] A hypothesis-testing study will often report the results using statistical tests and either *P* values or, even more relevant to clinical medicine, confidence intervals.[2,3] If the study is descriptive in nature, the aim is to convey detailed information to generate a specific research hypothesis for future research. The results in these studies should not be reported using statistical tests, and there does not need to be a specific a priori hypothesis, just a research question.[1]

The Methods section is the most important section, because poor methodology alone can invalidate the entire study.[1] The first detail to ascertain is the study design. Is it a primary study, with original data, or is it a systematic review, which summarizes primary studies to answer a question? Next, the study types may be broken down into either experimental or observational. The experimental study or clinical trial is either randomized or nonrandomized, and the intervention is under the control of the investigator. The observational study is either descriptive, such as in a case report or a case series, or it is analytic (allowing for hypothesis testing), and it is either taken at one time (ie, a cross-sectional design, such as a survey) or

over time, such as in a retrospective (ie, case-control) or prospective (ie, cohort) design.[4,5] A chart review is a method of data collection, like interviewing, and it is usually done in a case-control format. If you cannot ascertain the study design from the information provided in the article, this is indicative of a major flaw.

A Methods section should not be terribly short; if it is, this may be a warning of a poor quality study.[1] Ask yourself if this is a recipe that you could follow if you had to replicate this study exactly; take the time to work through the details in your mind. The second question to ask in this section concerns validity: internal and external. Are the results going to be internally valid using these methods, or will the study design produce results that are valid in the group it was conducted on? Is there bias, the introduction of systematic, nonrandom error that has affected the results? Several major areas of bias that can confound results are defined in Table 1. Will the results be externally valid, or can these results be applied to the population they are supposed to apply to? This was the reason the work was done, after all. Next, ask yourself if these patients are similar to the patients that you see—if you aren't sure, there is not enough detail provided to apply the results appropriately, and this is a problem.

Lastly in the Methods, ask yourself if it provides the "who, what, where, when, and how."[1] This is your recipe card, which should include the inclusion and exclusion criteria, as well as who did the intervention, the measurements, and the analysis, in order to answer the question of "who?" The "what" is the intervention, the diagnostics, and the measurements, along with an outcome definition and a power calculation if a hypothesis is being tested. The "where" should give you enough information to decide if the patients tested are similar to your own, and the "when" applies to the study entry and interventions, which may have occurred a significant amount of time before publication, so practices may have changed since then. In addition, ask yourself if you think the interventions and measurements were conducted at an appropriate time relative to one another in order to answer the study question. The "how" should be a detailed description of the analytic strategy and the data analysis, along with the methods used for measurements. The reason to ask for this recipe card is to confirm the internal and external validity of the study.

The most important aspect of the Results section is that it should parallel the Methods section. If they don't match, this is a problem.[1] Introducing new findings in the Results that the study wasn't designed to detect using the selected methods reflects poor methodology. Similarly, check that all the questions posed in the Methods were answered. Take the time to do some basic addition to see that the numbers add up, and that there are no disappearing subjects. If 2 groups are being compared, look for a table that compares the baseline characteristics of the groups, to check that their composition is similar. Unless you have a degree in clinical epidemiology, leave the statistical analysis to the statisticians, because physicians who don't

Table 1.
Several common sources of bias.

Bias	Description
Selection bias	Occurs when selection of study subjects is systematically distorted, which may predetermine the study outcome.[5,14]* For example: • A hospital study of diarrhea will overestimate severity because mild cases will not seek medical attention. • If certain types of patients refuse to take part in a study, the results may be biased.
Measurement or information bias	Bias in classifying disease, exposure, or both. For example, assessment of a patient may be influenced by the interviewer's knowledge of the patient's disease status, or knowledge of disease status may influence assessment of exposure. Leading questions can bias responses, and patients with a disease tend to remember past exposures differently than those without diseases.[14]
Confounding variables	A factor that distorts the true relationship of the variables of interest because it is related to the outcome but not to the study question and is unequally distributed between the groups being compared.[5,†]
Workup or verification bias	This bias is introduced when patients who have a positive (or negative) test result are preferentially selected for testing by a criterion standard. This results in missed patients, who were not selected for testing because of a milder form of disease. May result in erroneously high sensitivities and negative predictive values.[15]

*If the selected group was not representative of the individuals meant to be studied, this does not by itself constitute selection bias; the factor that makes the group different from the intended population must be able to affect the results as well.[5]
†A confounding variable does not invalidate the study if the problem is recognized and taken into account in the analysis.[5]

have statistical expertise tend to evaluate the use of statistical tests incorrectly. Your goal could be to understand if the right type of statistical test was used in this analysis.

The point of the Discussion is to evaluate the robustness of the study results and to address the limitations and caveats of the work (because all studies have limitations). It is not a literature review and should only place the

Table 2.
Criteria specific to a randomized controlled trial: The CONSORT checklist.[16]

Paper Section and Topic	Item	Description	Reported on Page #
Title and abstract	1	How participants were allocated to interventions (eg, "random allocation," "randomized," or "randomly assigned").	
Introduction			
Background	2	Scientific background and explanation of rationale.	
Methods			
Participants	3	Eligibility criteria for participants and the settings and locations where the data were collected.	
Interventions	4	Precise details of the interventions intended for each group and how and when they were actually administered.	
Objectives	5	Specific objectives and hypotheses.	
Outcomes	6	Clearly defined primary and secondary outcome measures and, when applicable, any methods used to enhance the quality of measurements (eg, multiple observations, training of assessors).	
Sample size	7	How sample size was determined and, when applicable, explanation of any interim analyses and stopping rules.	
Randomization: sequence generation	8	Method used to generate the random allocation sequence, including details of any restriction (eg, blocking, stratification).	
Randomization: allocation concealment	9	Method used to implement the random allocation sequence (eg, numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned.	
Randomization: implementation	10	Who generated the allocation sequence, who enrolled participants, and who assigned participants to their groups.	
Blinding (masking)	11	Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to group assignment. When relevant, how the success of blinding was evaluated.	
Statistical methods	12	Statistical methods used to compare groups for primary outcome(s); methods for additional analyses, such as subgroup analyses and adjusted analyses.	
Results			
Participant flow	13	Flow of participants through each stage (a diagram is strongly recommended). Specifically, for each group report the numbers of participants randomly assigned, receiving intended treatment, completing the study protocol, and analyzed for the primary outcome. Describe protocol deviations from study as planned, together with reasons.	
Recruitment	14	Dates defining the periods of recruitment and follow-up.	
Baseline data	15	Baseline demographic and clinical characteristics of each group.	
Numbers analyzed	16	Number of participants (denominator) in each group included in each analysis and whether the analysis was by "intention-to-treat." State the results in absolute numbers when feasible (eg, 10/20, not 50%).	
Outcomes and estimation	17	For each primary and secondary outcome, a summary of results for each group and the estimated effect size and its precision (eg, 95% confidence interval).	
Ancillary analyses	18	Address multiplicity by reporting any other analyses performed, including subgroup analyses and adjusted analyses, indicating those prespecified and those exploratory.	
Adverse events	19	All important adverse events or side effects in each intervention group.	
Discussion			
Interpretation	20	Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision, and the dangers associated with multiplicity of analyses and outcomes.	
Generalizability	21	Generalizability (external validity) of the trial findings.	
Overall evidence	22	General interpretation of the results in the context of current evidence.	

results of the study in the context of the current medical literature.[1] Problems covered in the Limitations should be those that the authors believe have the ability to bias the results, including lack of confidence resulting from small sample sizes. If there is a confounding variable that affects the results, have the investigators identified it and taken it into account in their analysis?[3]

Figure.
Criteria specific to a survey.[1,18]

1. Was the instrument used reliable and valid? In other words, were the study questions validated, or were they made up by the investigators? Are their development and validation described?
2. Was the survey instrument self-administered or interview based? If interviews were conducted, who conducted them, were they trained, and were they blinded to the research question?
3. Was the appropriate population surveyed?
4. Were the results collected anonymously, and if not, was this likely to affect the results?
5. Is the survey instrument displayed, and are the questions clear? Is a Likert or other validated scale used?

Table 3.
Criteria specific to chart review.[17]

Criteria	Description
Training	Train chart abstractors to perform their jobs. Describe the qualifications and training of the chart abstractors. Ideally, train abstractors before the study starts, using a set of "practice" medical records.
Case selection	Use explicit protocols and describe the criteria for case selection or exclusion. Do "checks" of patient logs and other tools to see if eligible patients are being missed.
Definition of variables	Define important variables precisely.
Abstraction forms	Use standardized abstraction forms to guide data collection. Ensure pre-established rules for handling of data that are conflicting, ambiguous, missing, or unknown.
Meetings	Hold periodic meetings with chart abstractors and study coordinators to resolve disputes and review coding rules.
Monitoring	Monitor the performance of the chart abstractors.
Blinding	Blind chart reviewers to the etiologic relation being studied or the hypothesis being tested. If groups of patients are to be compared, the abstractor should be blinded to the patient's group assignment.
Testing of interrater agreement	A second reviewer should reabstract a sample of charts, blinded to the information obtained by the first correlation reviewer. Report a κ statistic, interclass coefficient, or other measure of agreement to access interrater reliability of data.

Finally, the Conclusion should answer the research question(s) posed in the introduction. That's it.

After reviewing the criteria specific to the study design (Tables 2 and 3; Figure), consider some larger questions. Are there any fatal flaws?[1] Fatal flaws are usually a lack of scientific importance, originality, or validity. If the topic is not scientifically important or if the article makes no original contribution (ie, either a new finding, a new negative result, or confirmation of either), there is no benefit for mankind to gain from the results. Most importantly, if the methods of the article are not scientifically valid, the results aren't useful, whatever they are. Two other overarching questions to ask include the following: What are the strengths and weaknesses? Will it change your practice in any way?

The final steps before presentation are to check your review conclusions against someone else's, and to add any points that you may have missed. Do this by checking for editorial comments on the article (often provided as an article link on the journal Web site), which are usually made by an expert in the field, and check the review literature for opinions, such as the *Emergency Medicine Abstracts* series. Other references that may be useful are included in Table 4.

Table 4.
Useful references.

Reference	Source
JAMA: Glossary of Methodologic Terms	http://jama.ama-assn.org/content/vol290/issue1/images/data/125/DC5/auinst_term.dtl
JAMA: User's Guides to the Medical Literature	http://www.cche.net/principles/main.asp
BMJ: "How to Read a Paper"	http://bmj.com/collections/read.shtml
Cochrane Library	http://www.update-software.com/Cochrane/default.HTM
Meta-analysis	Moher D, Cook DJ, Eastwood S, et al. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUORUM statement. *Lancet.* 1999;354:1896-1900.
	Olson CM. Understanding and evaluating a meta-analysis. *Acad Emerg Med.* 1994;1:392-398.
Randomized, controlled trial vs observational trial	Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med.* 2000;342:1887-1892.
Diagnostic tests	Mower WR. Evaluating bias and variability in diagnostic test reports. *Ann Emerg Med.* 1999;33:85-91.

THE PRESENTATION

When sitting before your audience, start by summarizing the article, expressing concern at problem areas without interposing your opinion yet. Keeping some sort of outline at hand will limit any unintended omissions. Next, offer your global opinion, based on the major strengths and weaknesses of the article, any unanswered questions that you have, and the presence of any fatal flaws. Remember that it is easy to criticize, and that research is time-consuming and difficult for most people; acknowledging this provides a more objective analysis. Try to be constructive rather than destructive. Note whether you think the authors achieved their stated objective, and whether you feel this objective was appropriate to begin with. Lastly, state whether the results will affect your practice in any way, as well as any other points that you think are important.

Keeping in mind the goals of journal club at your institution will guide the emphasis of your presentation. In general, there are 2 major goals of journal club: the primary goal is to learn how to critically read a scientific article, and the secondary goal is to keep abreast of current medical literature on a specific topic.[6-10] I agree that the acquisition of critical appraisal skills should be the primary goal of journal club, because residency programs contain other educational sessions that teach the content material of medicine, whereas journal club is one of the few components of residency training where residents learn how to read critically, a skill they will need throughout their careers.

Most journal clubs review no more than 2 to 3 articles[6,7]; this may avoid superficial treatment of the studies. The presence of an attending staff person who is an expert in critical appraisal is optimal; ideally, this staff person would be a dedicated faculty director[11] for the year or more, would act as a resource for residents, and would provide a synopsis of learning points at the end of each journal club meeting.

Surprisingly, close faculty supervision during resident preparation time did not improve critical appraisal skills in one study on the structure of journal clubs.[7] A topic expert may provide clinical context to the study findings, and/or a separate resident may be assigned to provide a brief review (no longer than 2 to 3 minutes) of the topic and the state of the literature before the article presentations.

THE EVIDENCE FOR JOURNAL CLUB

Lastly, in the spirit of critical appraisal, one might ask if there is evidence for the effectiveness of journal club. Studies in the area are small. One randomized controlled trial found an improvement in residents' knowledge of epidemiology and biostatistics,[12] whereas another cohort study did not.[13] Residents who attended journal club reported being more skeptical of conclusions in the medical literature,[10,12] but there was no difference in their critical appraisal scores when compared with controls.[12] There are other goals to consider: journal club can stimulate interest in research, which may motivate residents to learn more about reading critically and to conduct their own research. Indeed, one study found that journal club was a powerful motivator of house staff reading behavior and, by self-assessment, their ability to critique methods of medical articles improved.[12] Journal club also demonstrates to residents how difficult it is to critically evaluate an article; it serves to sensitize residents to the limitations of their own skills as readers of the medical literature. We argue that, as one does not expect to become an expert on a topic after reading a few articles, so one should not expect to become an expert in research methodology after reading several scientific articles. If this were the case, no one would complete a masters or PhD degree in clinical epidemiology. The other, aforementioned benefits suffice to make journal club a worthwhile enterprise.

I thank Jerome R. Hoffman, MD, for his invaluable input and guidance on this article.

The author reports this study did not receive any outside funding or support.

Reprints not available from the author.

Address for correspondence: Clare Atzema, MD, Sunnybrook & Women's College Health Sciences Centre, 2075 Bayview Avenue, Room BG-13, Toronto, Ontario M4N 3M5, Canada; E-mail catzema @sympatico.ca.

REFERENCES

1. Callaham M, Schriger D, Cooper RJ. *An Instructional Guide for Peer Reviewers of Biomedical Manuscripts.* Dallas, TX: American College of Emergency Physicians; 2001.
2. Braitman LE. Confidence intervals assess both clinical significance and statistical significance. *Ann Intern Med.* 1991;114:515-517.
3. Guyatt G, Jaeschke R, Heddle N, et al. Basic statistics for clinicians: 2. Interpreting study results: confidence intervals. *Can Med Assoc J.* 1995;152:169-173.
4. Grisso JA. Making comparisons. *Lancet.* 1993;342:157-160.
5. Riegelman RK. *Studying a Study and Testing a Test.* 4th ed. Philadelphia, PA: Lippincott Williams & Williams; 2000.
6. Alguire PC. A review of journal clubs in postgraduate medical education. *J Gen Intern Med.* 1998;13:347-353.
7. Bazarian JJ, Davis CO, Spillane LL, et al. Teaching emergency medicine residents evidence-based critical appraisal skills: a controlled trial. *Ann Emerg Med.* 1999;34:148-154.
8. Crank-Patton A, Fisher JB, Toedter LJ. The role of the journal club in surgical residency programs: a survey of APDS program directors. *Curr Surg.* 2001;58:101-104.

9. Linzer M. The journal club and medical education: over one hundred years of unrecorded history. *Postgrad Med J*. 1987;63:475-478.
10. Woods JR Jr, Winkel CE. Journal club format emphasizing techniques of critical reading. *J Med Educ*. 1982;57:799-801.
11. Inui TS. Critical reading seminars for medical residents. A report of a teaching technique. *Med Care*. 1981;19:122-124.
12. Linzer M, Brown JT, Frazier LM, et al. Impact of a medical journal club on house-staff reading habits, knowledge, and critical appraisal skills. A randomized control trial. *JAMA*. 1988;260:2537-2541.
13. Langkamp DL, Pascoe JM, Nelson DB. The effect of a medical journal club on residents' knowledge of clinical epidemiology and biostatistics. *Fam Med*. 1992;24:528-530.
14. Sitthi-Amorn C, Poshyachinda V. Bias. *Lancet*. 1993;342:286-288.
15. Mower WR. Evaluating bias and variability in diagnostic test reports. *Ann Emerg Med*. 1999;33:85-91.
16. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Lancet*. 2001;527:1191-1194.
17. Gilbert EH, Lowenstein SR, Koziol-McLean J, et al. Chart reviews in emergency medicine research: where are the methods? *Ann Emerg Med*. 1996;27:305-308.
18. Fink A, Kosecoff J, eds. *How to Conduct Surveys: A Step by Step Guide*. Beverly Hills, CA: Sage; 1998.