

**SURGICAL RESIDENTS' CLUB
UNIVERSITY OF SOUTH FLORIDA**

Reading the medical literature requires some sophistication. This scoring sheet points out some important issues to consider when reading medical papers. Considerations raised herein are not to be taken as absolute, but instead as strong guidelines. Furthermore, this scoring sheet can be used in preparing manuscripts as it is a valuable “checksheet” to ensure all bases are covered. The more points a manuscript gets, the better written. However, how interesting it is, how meaningful it is, or how important it is may be another issue. It is impossible to write a strong paper from a bad data set. A poor paper can be written from a great data set. A great paper can only be written from a well designed study and a great data set.

A) WHY?

Why was the work undertaken? The authors should provide sufficient justification as to why this study was undertaken. This justification should be found in both the Introduction and, in a non-repetitive manner, in the Discussion sections of the article. Bias in undertaking the study should be sought and may negatively impact the scoring of the article. In judging whether or not there is justification for undertaking the study, use the following criteria.

1) CLINICAL RELEVANCE (0-10):

The work addresses a clinical problem of relevance: Score:0-10. In assigning this score you must apply your own judgment as to the clinical relevance and magnitude of the clinical problem which the paper addresses. Every paper should be important, not just a “so what, who cares” paper. An excessively esoteric paper (e.g., Wilson’s disease on Mars) scores no points. You need to know the body of medical literature to accurately score this. Also, consider the forum this paper is presented in (“newspaper” vs trade journal).

2) ORIGINALITY (0-5):

Ideally, the study proposes work not previously done to address a clinical problem. In assigning the score, you must determine not only the innovation of the project but also its value in solving the clinical problem. The authors should make the case for their work in the Discussion and the Introduction sections.

If the work is not original, it must validate, confirm, refute, or provide another approach to work previously done, thus, lowering the potential score for originality. Another paper about a common problem (bile duct injuries after laparoscopic cholecystectomy) scores no points unless it is somehow unique (very large size, etc.). In assigning a score, you must judge how important it is to confirm or refute the previous work(s) in this area and how relevant or important such a confirmation would be. An example of work which confirms previously done work might be a study done with a larger sample size including different types of patients, or a study which prolongs the experimental period, or new data which suggests that a previous study was flawed. Be critical, especially of “copycat” papers. A “copycat” paper scores no points.

3) *PURPOSE (0-5):*

Is the purpose clearly stated? In assigning this score, look for the phrase “we undertook this study to” or “the purpose of this study was to...”. If you do not find such explicit wording, then the purpose is implied and only 1-2 points can be given. If you can find an implied purpose, assign the value of 1 or 2 depending upon how much work you have to do to uncover the purpose. The purpose of the study should be clear. Don’t hesitate to score zero if you have to work hard to figure out why the study was undertaken. Look for bias. Bias can be acceptable, if the design of the study allows for an independent outcome. That is, bias can not affect study outcome.

4) *HYPOTHESIS (0 or 5):*

Is the hypothesis clearly stated? Yes, score 5; no, score 0. This is one of the most important parts of a scientific paper. The authors must state that they have formulated a hypothesis, which the work is intended to prove or disprove. The hypothesis should be explicitly stated. Look for phrases such as “we hypothesized that...” or “our hypothesis was...”. Implied hypotheses get no score. Always look for a potential inappropriate bias (e.g., funding source). The hypothesis stated should not be, though can be, a null hypothesis.

B) HOW?

How did the authors do their work? You should be critical in assessing the methodology and study design.

If the paper describes animal work in which variables are controlled, you should expect a well-designed and controlled study. You should study the following issues.

DESIGN (0-10)

Is the design of the study appropriate to test the hypothesis? Are the variables that the authors propose to measure, appropriate to test the hypothesis? Score: 0-10. In assigning a value, determine if there are control animals that receive similar treatment in all aspects but the intended therapy or insult. Is the duration of the study long enough to observe the outcome or long enough to be clinically applicable? How relevant are the variables to the hypothesis? Is the end point which the authors propose appropriate and relevant to the clinical situation? This is a very important point. For example, in shock research the “effectiveness” of a given therapy is often measured in terms of survival. Clearly, the effectiveness of any therapy can be judged in terms of survival, but survival is clearly dependent on *multiple* variables. Rather, therapeutic regimens should also be judged in terms of their “efficacy” in treating relevant physiologic variables such as oxygen delivery, oxygen consumption, blood pressure, and cardiac output. Simply stating that a given therapy does not improve survival is not necessarily an indictment of a therapy, which is intended to improve physiology. Any critical analysis of a therapy should look at both efficacy and survival.

If the study involves humans, the variables are difficult to control and scoring can be a little more liberal. First, was IRB approval obtained? This will provide some assurance of proper design. Then determine the study design. High score should be given for prospective randomized and double-blinded studies. Note how patients are randomized. This should be clear. Lower scores are obtained for those studies that are retrospective chart reviews, “clinical series” or case reports. In general, these latter designs score few points, though the paper may have considerable clinical interest (but that is a different issue—note A.1.).

Are the clinical variables explicitly defined? This is critically important in any retrospective or prospective study. The criteria for entry of a patient into the study should be explicitly defined, otherwise, “garbage in, garbage out.” Clinical end points should also be explicitly defined. For example, if the work is a study of “head injury”, how is “head injury” defined? How is “infection” defined? How is “pancreatic cancer” defined and which cell types are included? Do the authors use ICD 9 codes, etc.? Who coded the diagnosis? What is the database and how good is it? If the study is comparing treatments based on morbidities, how well are the morbidities defined? For example, if “pneumonia” is an end point, how is “pneumonia” defined? Be critical.

METHODOLOGY (0-10)

Are the methodology and techniques for measurement standardized and accepted? Score 0-10. In assigning a value, do the authors employ a standard methodology, which has been proven to effectively measure or quantitate the variable they are using? Is the methodology used clinically? Is it available to you? Are the authors proposing a new technology or assay to measure a variable? Are there better ways to measure than what they use? If they are proposing a new variable, how accurate or valid is the method and do they provide accuracy or validation data? Be critical and thoughtful.

ANALYSES (0-10)

Were statistical analyses undertaken? Was the sample size sufficient to detect a clinically significant difference? Were the analytical methods appropriate to the study design and were corrections made for multiple analyses or comparisons? Was the data parametric or non-parametric and were appropriate statistical analyses appropriate? Score 0-10. In assigning a value, judge whether analysis was done or whether only descriptive statistics (mean age, mean injury severity score, number of males or females, etc.) were presented. If only descriptive statistics are presented assign a value of 0. If the data set or study design only requires descriptive statistics, this is not a strong paper, though possibly interesting (that is another issue: note A.1). The authors should also explicitly state the level of significance for their study. Generally, authors will accept a 5% chance of a type one error (rejecting the null hypothesis when it is, in fact, correct). Data should be presented as mean \pm STD, not mean \pm SEM, though the latter looks better.

C) WHAT?

Results section. Do the authors clearly and succinctly present all the data that they described in their methods section? Make sure data is presented fairly, openly, and without bias. The data of the study should be so clearly presented that you should be able to analyze the data. Do they

provide measures of variance, such as standard deviation? Were animals or subjects eliminated (do all the numbers add up)? If animals and subjects were eliminated, do they provide the sample sizes for each of the analyses performed? Are the data easily retrievable from the result section? Score 0-35. Be critical. In assigning a score, determine how much work you had to do to find the relevant data to either prove or disprove the hypothesis. Look at the tables and graphs and decide if they are clear and add to the presentation. The more work you had to do, the lower the score should be. If you perceive that there are confounding variables which might affect the results, do the authors provide data about those confounding variables? In sum, the data should be presented so that you could redo their analyses and test their conclusions.

D) SO WHAT!

This is the Discussion section in which the authors provide us with the relative importance of their work compared to previously done work and compared to the literature. Here, they present the importance of the work in addressing the clinical problem.

The first paragraph of the discussion should succinctly and clearly state why this paper should be published. If the authors can't state a strong case, can you? This should be strongly and clearly stated. This is very important.

The subsequent paragraphs should discuss all data, putting the results of the study in some perspective.

Is clinical relevance explicitly stated? Do the author's data justify their conclusions? Do they compare their work to relevant and important or landmark studies in the field? Are the references current (within the last five years), adequate in number, and from respectable journals? Are there important references missing. Is there a summary or conclusion, which clearly relates to the work they did and which clearly refers to the hypothesis? Are the authors critical enough of their work? Score: 0-10. A poor discussion gets 0 points.

E) SUMMARY

- Don't help the authors. Don't work too hard in reading a paper.
- Be Critical
- Answer "Why should this paper be published?" It should be easy to answer.
- Is this paper important? Will anyone care?
- 80-100: Strong paper
- 60-80: OK
- 40-60: Caution
- < 40: A waste of paper

Note: I recognize Dr. Steven Shackford for his efforts in this manuscript grading system.