The Good, the Bad and the Ugly: Essential Attributes of a Great Manuscript Review

Steven R. Shackford, MD
Angela Sauaia, MD PhD
E. E. Moore, MD
Steven R. Shackford
Disclosure Information
NOTHING TO DISCLOSE.
OVERVIEW

➢ Shackford:
  ➢ Generic issues: reviewer decisions
  ➢ My process
  ➢ Good review

➢ Sauaia:
  ➢ Review of statistical methodology

➢ Moore:
  ➢ Bad review
OBJECTIVES

- What is the purpose of peer review?
- What do the following really mean?
  - Accept
  - Major Revision/Reconsideration
  - Marginal
  - Reject
- My process
- What are the attributes of a good review?
Purpose of Peer Review

(E A R)

- Ensure the quality of the submitted manuscript
  - Ethical
  - Accurate
  - Relevant to the readership
- Ethical: IRB or IACUC approval
- Accurate: no flaws in design, methodology, or interpretation
- Relevant:
  - Addresses an existing knowledge gap
  - Knowledge gap has relevance to readership
A Good Review

- Improves the Journal
- Opportunity to educate yourself
  - Reading the manuscript
  - Reviewing current literature
- Opportunity to teach/learn
  - Grammar and usage (Strunk and White)
  - Structure of the manuscript
- See Sauaia A et al:
  - Discussion: JTACS: 2013; 74: 1599
  - Introduction: JTACS: 2014; 76: 1322
Reviewer Choices: Definitions

- Accept: “good to go”, no issues with EAR
- Major revision/reconsideration
  - Major flaws, non-fatal
    - Addressed by additional data or analysis
    - Conclusions too sweeping, not justified by data
    - Presentation (too long, rambling)
- Marginal
  - Flaws, nonfatal-even if addressed likely reject
  - Work has already been done and done better
  - Marginally relevant to readership
- Reject
  - Fatal flaws--wrong comparator group, confounders not addressed, not relevant
My Process

- Arrives in my “in box”
- Can I meet the deadline?
  - Schedule time (~2-2.5h)
- Download the PDF: use ADOBE tools
  - Read, highlight, make notes: 1 h
- Wait minimum of 24 hours
  - Review highlighted parts
  - Write review: 1.5 h
Good Review

- Should address each of the following:
  - Novelty and interest (relevance)
  - Background (knowledge gap)
  - Hypothesis/objective
  - Methodology (accuracy, ethical)
  - Interpretation of the data (accuracy)
  - Conclusion
  - Clarity (presentation)
- Confidential comments to editor (justify your decision)
Novelty/Interest

“The research question is very interesting and timely - the authors attempt to shed more light on the important issue of prognostication in elderly TBI. This is particularly timely given increasing attention to the use of non-beneficial care at the end of life in the elderly, and the use of palliative care.”
Background

“The background is very well written. It raises important questions. The review of pertinent literature on TBI and end of life care could be more expansive and include more recent literature, i.e., papers by Turgeon et al on variability of end of life practice in severe TBI.”
Hypothesis/Objective

“The research question in this article is somewhat unclear. Initially stated as an evaluation of the value and validity of the EAST geriatric TBI guidelines regarding responders and non-responders, it also seeks to study end of life care practices in severe TBI, and to determine if EAST guidelines are followed. Both are interesting and pertinent questions, that have been thus far poorly studied in a systematic way.”
Methods

“1. The methods of this study are excellent, and the use of statistical methods appears appropriate for the dataset and research question. One concern is the arbitrary selection of GCS<8 at 72 hours as the definition of non-responders. The authors do not describe why they selected this cut off, and any literature that supports it…”

2. There are important data points and variables missing from the data...

3. Data definitions - define "Change in Code Status”…

4. Sample size - the cohort is too small; raises concern for a Type II error…”
Interpretation

“The interpretation of the data is appropriate, but significantly limited by the issues stated above. End of life Decision Making - authors demonstrate no difference in EOL decision making - what does this mean? Based on the figure and Kaplan Meier curve ALL patients who survived to hospital discharge had no difference in 12 month mortality, suggesting that the responder/non-responder dichotomy is not a useful prognostic tool for long-term survival. The authors should discuss this in the results.”
Conclusion

“The conclusions are vague, and not entirely related to the results. Authors should discuss their results in the context of the original research question about the utility of the EAST guidelines and basing prognosis on responders/non-responders at 72 hours. Their data suggest that ALL geriatric TBI with GCS<8 have a poor prognosis with respect to long-term mortality - this alone would constitute a trigger for palliative care or end of life care discussions about life sustaining therapy, regardless of "responder” status. This should be highlighted in the conclusions.”
Clarity/Presentation

“The paper is well organized, well written, the tables and figures are relevant and well done. Additional data elements should be added as mentioned above. The discussion and conclusions should be amended to interpret the results at hand, and discuss the original research question.”
Comments to the Editor

“This is a potentially interesting and important paper - little in the literature examines the prognostication in elderly severe TBI, and when palliative care should be considered in this group. However, the hypothesis/research question is not well thought out, and the dataset is incomplete, and too small sample size.

The paper needs MAJOR revision and improvement, but I think if revised would be an important contribution to the literature and should be published. I suggest that the data elements referenced above should be added…”
Angela Sauaia
Disclosure Information
NOTHING TO DISCLOSE.
The Bad

Statistics are not really the BAD, they are WICKED!
The findings should be obvious. Statistics are just frosting.
Learning Objectives

• Study design: appropriate?
• Recognition of common biases in surgery research: survivor bias, intervention bias, selection bias
• Missing data: do they matter?
• Significant and non-significant findings: type 1 error \( (p\text{-value/ confidence}) \) and type 2 error/statistical power
• Multivariate analysis: the essentials
• Propensity score: the hot item of the moment
• Use of reporting standards to review an article
**J Trauma Standardized Statistical Review**

<table>
<thead>
<tr>
<th>Checklist for statistical assessment of general papers</th>
<th>YES</th>
<th>Comments (at the end of document by number)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appropriate study design used to achieve the objective(s)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source of subjects/data appropriately described</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sampling/sample size appropriately described</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entry and exclusion criteria clearly defined</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data exclusions are stated/explained and impact on results are explored</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N reported at the start of the study, for each data set and for each analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discrepancies in value of N between analyses clearly explained/justified</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing data are explained, and impact on findings minimized/explained</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satisfactory follow-up/response rate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adequate uni/bivariate statistical analyses used/described</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adequately multivariate statistical analyses used/described</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confounding and bias explored and minimized</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assumptions of tests applied met (particular attention paid to non-normal data sets or small sample sizes)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjustments made for multiple testing explained</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unit of analysis given for all comparisons</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alpha level given for all statistical tests</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tests clearly identified as one or two-tailed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actual P values are given for primary analyses</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appropriate measure(s) of center (e.g. mean or median)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appropriate measure(s) of variability (e.g. standard deviation or range)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unusual/complex statistical methods clearly explained for JT’s wide readership</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Method of treatment assignment (randomization etc) explained and justified</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any data transformations clearly described and justified</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confidence intervals given for the main results</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conclusion drawn from the statistical analysis is justified</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- This checklist can be used to guide reviews
Appropriate study design
Control group / Comparator

- The major factor in study design is the choice of an adequate control group or comparator group.
- Control group: people similar to the test group in all aspects that affect the outcome except for the treatment/intervention/risk factor of interest.
- Are the groups comparable or do we have apples and oranges?
Are the groups comparable?

- Design methods to have comparable groups:
  - Randomization
  - Matching on selected variables
  - Propensity score matching: matching on a score composed of several confounders

- Analysis:
  - Stratified analysis on specific variables
  - Multivariate analysis

**Important review questions:**

Was confounding sufficiently minimized (by design or analysis?) Were important confounders left out?
Description of Sample

• The description should allow readers to visualize the study population and compare it to theirs.
  – Categorical data: Ns and %
  – Continuous data:
    • Normally distributed data (SBP, glucose, lactate, BD): means and std. deviation
    • Skewed data (transfusions, LOS, ventilation time, often ISS, GCS, TEG LY30, etc.): median and interquartile range (25th percentile and 75% percentile)
    • Cutoffs: were cutoffs (e.g., GCS<8) based on evidence?

Important review question: can you visualize the sample? Can you compare it to your own?
## Statistical Inference

<table>
<thead>
<tr>
<th>Result of statistical test</th>
<th>Truth</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Null hypothesis false</td>
</tr>
<tr>
<td>Statistically significant</td>
<td>True positive</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Not statistically significant</td>
<td>False negative</td>
</tr>
<tr>
<td></td>
<td>Type II error (β)</td>
</tr>
</tbody>
</table>
If the results are statistically significant (p<0.05)

- Sample size is small:
  - Statistical power/type 2 error is irrelevant
  - Study may have internal validity (significant finding) YET Study may lack external validity a.k.a. generalizability

Important review question: does the small sample raises issues around generalizability and reproducibility?

- Sample size is large:
  - Statistical significance vs. Clinical significance

Important review question: Would you change your practice to achieve that difference?
Non-significant findings (p>0.05)

• Negative results regarding important outcomes must be ALWAYS accompanied by a power analysis
• NO significant difference DOES NOT equate to EQUIVALENCE

• NON-INFERIORITY TRIALS:
  – Similar outcomes but better adherence, lower costs, etc.
  – Difference in outcomes established based on evidence: shown to not increase harm and not decrease benefits
  – Power directed at detecting this established difference (or bigger)
  – Sample size usually much larger than superiority trials

Important review question: if results were non-significant, did authors report statistical power? If a non-inferiority trial, was difference pre-established based on evidence?
P-hacking

• Definition: trying multiple things until getting the desired result

• How it is done:
  – Exclusion of patients
  – Dropping variables
  – Variable manipulation
Common Biases In Trauma Research

Seven deadly sins in trauma outcomes research: An epidemiologic post mortem for major causes of bias

Deborah J. del Junco, PhD, Erin E. Fox, PhD, Elizabeth A. Camp, MSPH, Mohammad H. Rahbar, PhD, and John B. Holcomb, MD, on behalf of the PROMMTT Study Group, Houston, Texas

RESULTS:
Sources of bias in trauma research include ignoring (1) variation in patients’ indications for treatment (indication bias), (2) the dependency of intervention delivery on patient survival (survival bias), (3) time-varying treatment, (4) time-dependent confounding, (5) nonuniform intervention effects over time, (6) nonrandom missing data mechanisms, and (7) imperfectly defined variables. This list is not exhaustive.

CONCLUSION:
The mitigation strategies to overcome these threats to validity require epidemiologic and statistical vigilance. Minimizing the highlighted types of bias in trauma research will facilitate clinical translation of more accurate and reproducible findings and improve the evidence-base that clinicians apply in their care of injured patients. (J Trauma Acute Care Surg. 2013;74: S97–S103. Copyright © 2013 by Lippincott Williams & Wilkins)

RECOMMENDED
Common biases in trauma research

Selection bias: selected patients differ from excluded patients in systematic ways. Types:

- **Survivor bias:** death or d/c before receiving intervention
- **Missing data:** missing data must ALWAYS be reported. Important when >10% and **not at random**. If missing not at random, must be addressed in analysis

- **Intervention bias:** groups defined by receiving an intervention related to the outcome or to the risk factor (e.g., massive transfusion)

**Important review questions:** Was bias sufficiently minimized? Was missing data reported and addressed?
Multivariate Model Reporting
Both for Propensity Score Matching
and Multivariate Adjustment

Important questions for reviewers:

— Was the sample large enough for a multivariate model: >7-10 patients WITH THE OUTCOME per variable included in the model.

— Were all major confounding variables accounted for? Was confounding sufficiently minimized? How were they selected?

— Were all measures of risk (odds ratios, risk ratios, hazard ratios) in the expected direction? If not, did the authors justify/discuss it?

— If multicenter, were cluster effects by facility accounted for? (i.e., patients within facility are more correlated than patients of different facilities)

— Were interactions and polynomial terms (e.g., U-shaped association) tested for?

— Did the authors report measures of model performance (e.g. AUROC, R-square, etc.) Were these measures discussed and interpreted?
If a polynomial term (e.g., $X^2$) is not included, a straight line will be fitted (red line in graphs above) and one misses the different risk associated with the middle categories. The stats software will not warn you about this.
Survival Analysis
When time (to event) is the essence

A great analytic tool when:
1. Time to event is of interest
2. There is censoring (patients die before having the “opportunity” to experience the event, loss to follow-up)
3. Time-varying covariates: intervention/risk factor varies over time

Important review question: if any of the above three criteria were present, was survival analysis used?

• All reporting standards have checklists.
• We encourage authors to use these but reviewers can also use them to guide reviews.
Manuscript Review: The Ugly

EE Moore, MD Denver CO

Disclosure Information:

Haemonetics ... Research support / patents
TEM... Research support

Editor, Journal Trauma and Acute Care Surgery

Friday Sept 11, 2015
Journal Review : No Information

<table>
<thead>
<tr>
<th>Custom Review Question(s)</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do you have any conflicts of interest relating to this manuscript?</td>
<td>No</td>
</tr>
<tr>
<td>Do you agree with the authors’ level of evidence rating for this study?</td>
<td>N/A</td>
</tr>
<tr>
<td>Do you have reason to believe that this manuscript (in whole or part) has published before?</td>
<td>No</td>
</tr>
<tr>
<td>Would you be willing to write an editorial critique to appear with this paper, if accepted?</td>
<td>No</td>
</tr>
<tr>
<td>Should this manuscript be reviewed by the Journal’s biostatistician?</td>
<td>No</td>
</tr>
<tr>
<td>Overall novelty/interest of the research question:</td>
<td></td>
</tr>
<tr>
<td>Coherence and completeness of background:</td>
<td></td>
</tr>
<tr>
<td>Hypothesis/study objective clarity:</td>
<td></td>
</tr>
<tr>
<td>Adequacy of methods/approach:</td>
<td></td>
</tr>
<tr>
<td>Interpretation of data:</td>
<td></td>
</tr>
<tr>
<td>Conclusion/translatable message:</td>
<td></td>
</tr>
<tr>
<td>Clarity of presentation/organization:</td>
<td></td>
</tr>
<tr>
<td>Do you wish to earn Continuing Education credits for this review? Your review will be evaluated and must be approved by the editor. If approved, your CME certificate will be e-mailed to you.</td>
<td>No</td>
</tr>
<tr>
<td>How long did it take for you to complete this review?</td>
<td>30 minutes up to 1 hour</td>
</tr>
<tr>
<td>The next 2 questions are required in order to receive CME credit. Performing this review has improved my knowledge and ability to assess the scientific literature in order to make informed decisions in my practice.</td>
<td>No</td>
</tr>
<tr>
<td>Performing this review has improved my critical thinking and writing skills within my area of expertise.</td>
<td>No</td>
</tr>
</tbody>
</table>

Comments to Author:

Additional comments to author:
none
# Journal Review: No Information

<table>
<thead>
<tr>
<th>Custom Review Question(s)</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do you have any conflicts of interest relating to this manuscript?</td>
<td>No</td>
</tr>
<tr>
<td>Do you agree with the authors' level of evidence rating for this study?</td>
<td>Yes</td>
</tr>
<tr>
<td>Do you have reason to believe that this manuscript (in whole or part) has published before?</td>
<td>No</td>
</tr>
<tr>
<td>Would you be willing to write an editorial critique to appear with this paper, if accepted?</td>
<td>Yes</td>
</tr>
<tr>
<td>Should this manuscript be reviewed by the Journal's biostatistician?</td>
<td>No</td>
</tr>
<tr>
<td>Overall novelty/interest of the research question:</td>
<td></td>
</tr>
<tr>
<td>Coherence and completeness of background:</td>
<td></td>
</tr>
<tr>
<td>Hypothesis/study objective clarity:</td>
<td></td>
</tr>
<tr>
<td>Adequacy of methods/approach:</td>
<td></td>
</tr>
<tr>
<td>Interpretation of data:</td>
<td></td>
</tr>
<tr>
<td>Conclusion/translatable message:</td>
<td></td>
</tr>
<tr>
<td>Clarity of presentation/organization:</td>
<td></td>
</tr>
<tr>
<td>Do you wish to earn Continuing Education credits for this review? Your review will be evaluated and must be approved by the editor. If approved, your CME certificate will be e-mailed to you.</td>
<td>No</td>
</tr>
</tbody>
</table>

## Comments to Author:

This is an excellent manuscript that covers a major problem in care of the elderly particularly after trauma. It is an incredible amount of material to go over. If this manuscript is accepted, I would be willing to submit an editorial.
## Journal Review: Selecting Verbiage

<table>
<thead>
<tr>
<th>Interpretation of data:</th>
<th>The Results section opens with a sentence describing the Methods. Is this the way in which the authors propose to communicate in a scientific paper?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conclusion/translatable message:</td>
<td>The conclusions are wildly overstated and not supported by the data.</td>
</tr>
<tr>
<td>Clarity of presentation/organization:</td>
<td>The overall presentation is terrible and appears that this paper was not reviewed by any senior scientist...or anyone who actually cares.</td>
</tr>
</tbody>
</table>
Journal Review: Selecting Verbiage

Comments to Author:

Additional comments to author:
This paper suffers from fatal structural flaws. It is unpublishable. It starts with a provocative title, but then gives NO data supporting this titled hypothesis. Worse, the hypothesis in the title is different than the hypothesis in the Background. It goes on jumping around in a disorganized fashion. It is not well or even properly written. In the end its a tiny series of n=2 ureteral injuries that got fixed, n = 7 that were missed, and a few more that were watched. It almost literally adds nothing to the literature.

This review sounds unnecessarily harsh, but it is 100% accurate. It is the review YOU should have given yourself before bothering to submit, or the review that a kindly expert whose advice you should have sought out might have given you before you submitted. The Journal of Trauma is the pre-eminent trauma journal in the world. It would help you going forward to better understand what sorts of studies generally appears within its covers, to allow you to better prepare your manuscripts, and to allow you to better fit the quality of your submissions to this very high quality journal. "People learn from one another, just as iron sharpens iron." Proverbs 27:17
Journal Review: Mixed Messages

<table>
<thead>
<tr>
<th>Custom Review Question(s)</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do you have any conflicts of interest relating to this manuscript?</td>
<td>No</td>
</tr>
<tr>
<td>Do you agree with the authors' level of evidence rating for this study?</td>
<td>N/A</td>
</tr>
<tr>
<td>Do you have reason to believe that this manuscript (in whole or part) has published before?</td>
<td>No</td>
</tr>
<tr>
<td>Would you be willing to write an editorial critique to appear with this paper, if accepted?</td>
<td>Yes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Clarity of presentation/organization:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>The paper is clear, concise, well-written, well-organized and well edited. The thought process is easy to follow and the entire process from hypothesis to conclusion flows smoothly and logically. References are abundant and address the background for and pre-existing knowledge for the study. The Tables are easy to interpret and clearly present important additional data not given in the text.</td>
<td></td>
</tr>
</tbody>
</table>
Journal Review: Mixed Messages

Additional comments to author:

The authors have done a very nice job collecting data on the value of MRI to exclude spine injury in the setting of a negative CT after trauma. They collected data on more than 1700 patients from multiple studies and found exactly 5 patients in whom abnormal MRI studies led to surgical intervention (0.003%).

I don't think that they have adequately proved from their data that a cervical collar can be discontinued in an obtunded patient after just a negative CT. I especially do not know what they are referring to when they say to "reserve MRI for those situations in which the surgeon has additional concerns requiring further imaging". They did not delineate those concerns in the manuscript.
Journal Review: Mixed Messages

<table>
<thead>
<tr>
<th>Hypothesis/study objective clarity:</th>
<th>Endovascular repair for transected proximal descending thoracic aortic injuries from blunt injury is bests accomplished today by endovascular techniques. (NOW MANY MANY papers have said just that.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adequacy of methods/approach:</td>
<td>Not novel. Same review as I have read by many different authors.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coherence and completeness of background:</th>
<th>Solid review of the literature and no glaring deficiencies.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypothesis/study objective clarity:</td>
<td>Again very solid.</td>
</tr>
<tr>
<td>Adequacy of methods/approach:</td>
<td>No methods sections to describe their search methods and inclusion of this review which is usually standard in ost literature review papers.</td>
</tr>
<tr>
<td>Interpretation of data:</td>
<td>Nicely done on the selected manuscripts.</td>
</tr>
</tbody>
</table>
Journal Review : Systematic

Comments to Editor:

Additional comments to the editor: see attached biostats review. I am appalled that the other reviewers did not realize how poorly analyzed the study is. They have the data, so they can do better, but the sample is quite small, which weakens the study.

<table>
<thead>
<tr>
<th>Review Question Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Original Submission</strong></td>
</tr>
<tr>
<td>Reject</td>
</tr>
<tr>
<td>Provisional Accept</td>
</tr>
<tr>
<td>Reject</td>
</tr>
</tbody>
</table>