Reviewing Medical Education Research Abstracts

Adapted from module sponsored by the Research in Medical Education (RIME) Section of the Group on Educational Affairs of the Association of American Medical Colleges

Disclaimer: The views expressed in this module are those of the authors and do not reflect the official policy or position of the RIME program planning committee, the USUHS, the Department of Defense, or the United States Government.
Module Developers

Paul Hemmer, MD, MPH
Uniformed Services University of the Health Sciences F. Edward Hebert School of Medicine

G. Dodd Denton, MD, MPH
Uniformed Services University of the Health Sciences F. Edward Hebert School of Medicine

Karen Szauter, MD
University of Texas Medical Branch School of Medicine

Brian Mavis, PhD
Michigan State University College of Human Medicine
Objectives of this Module

1. Orient to process of reviewing an abstract;
2. Define major structural abstract components;
3. List key considerations for reviewing an abstract;
4. Identify principles regarding written comments and reviewer etiquette.
Module Content

- Why Review Abstracts?
- Anatomy of an Abstract
- Written Feedback
- Reviewer Etiquette
Why Review Abstracts?

- Contributes to planning committee efforts to develop a conference program
- Supports the work of colleagues
- Part of educational scholarship
- Enhances your own writing
- Engages you with “Community of Practice”
Keep in mind…

- Abstracts are *different* than manuscripts
  - Word restrictions limit the scope
  - Often describe preliminary or pilot studies
  - Excludes supporting documentation such as references, tables, etc.
Anatomy of an Abstract
Introduction

Also known as:
Background
Purpose
Problem Statement
Objectives
The Introduction

- Introduction provides the study background or rationale
  - Reference to what is already known
  - Reference to general knowledge or generally accepted standards of practice

- Introduction must include a purpose or problem statement or hypothesis
Opening sentences should quickly orient reader to the study focus

**Important:** abstract word limits often preclude authors from including any evidence of familiarity with current issues related to the study topic
Problem Statement or Hypothesis

- Introduction must include clear statement of purpose.
  - If not present, Introduction is unacceptable

- Reviewer should clearly understand:
  - what question was being investigated, and
  - why this question is important
Statement of Purpose
- “Our purpose is to describe students’ views on the educational impact of medical school expansion”

Problem statement
- “We investigate the relationship between USMLE Step 1 scores and self-assessed study skills”

Hypothesis
- “We hypothesized that our students would demonstrate improvement in exam scores after participating in on-line remediation”
References to the Literature

- Word limits preclude full reference citations
  - Therefore not an essential element
- Authors might include abbreviated citation, e.g.: Smith, 2008.
- Allows author and reviewer to differentiate between “opinion” and building on another’s work
Is A Pre-Clerkship Examination Equivalent To The Pre-Third Year GPA as a Predictor Of Performance On The Third Year Internal Medicine Clerkship?

- We sought to determine a) whether an early clerkship examination (pretest) correlates with pre-clinical grade point average (GPA) or USMLE Step One scores and, b) whether GPA, USMLE, and pretest scores equally predicted clerkship outcomes.
Methods

Also known as:
- Methodology
- Research Design
General Considerations

- **Clarity**
  - Logical flow; should make sense to reviewer

- **Operationalizes the Research Question**
  - Drives study design

- **Ethics Review**
  - Explicit statement about IRB approval or review is optimal,
    - not considered a “make or break” criterion
Study Design Description

Quantitative Designs
- Randomized control trials
- Matched control studies
- Pre/post designs
- Surveys

Qualitative Designs
- Grounded theory
- Content Analysis
- Focus Groups
- Case studies
- Surveys
Study Sample

- Who was sampled? Was it appropriate?
  - Patients, students, residents, fellows, faculty?

- Was the sample size adequate?
  - Statistical power especially important for “negative” or “no difference” studies

- Was the setting described?
  - What is the context of the study?

- What was the source of data?
  - Survey, papers, blogs, examinations, etc.
### Quantitative Outcomes

<table>
<thead>
<tr>
<th>Type of Outcome</th>
<th>Focus</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. REACTION</td>
<td>satisfaction, usefulness, motivation</td>
<td>“I asked students to write about critical incidents in a blog and they liked it”</td>
</tr>
<tr>
<td>2. LEARNING</td>
<td>Acquisition of knowledge, skills, attitudes</td>
<td>“Students who wrote in a blog significantly improved their knowledge of critical incidents”</td>
</tr>
<tr>
<td>3. TRANSFER</td>
<td>Change in knowledge, skills, attitudes</td>
<td>“Students who wrote in the blog led a tutorial project about critical incidents for second year medical students”</td>
</tr>
<tr>
<td>4. RESULTS</td>
<td>Changed KSA lead to changes in real world outcomes</td>
<td>“Students who wrote about critical incidents in blogs had fewer medication errors than prior interns.”</td>
</tr>
</tbody>
</table>
What was the plan for analysis of the data?

Were the statistical methods described so that you understand how it was done?

Were the methods appropriate for the research question?

If you need help, ask the meeting planners if you can review the abstract with others.
Questions for Qualitative Studies

- Was sampling appropriate to the question?
  - What were the sources of information?

- Was data collection appropriate?
  - Interviews, focus groups, survey responses, written comments

  - Triangulation may be utilized
Questions for Qualitative Studies

- Is the analysis described sufficiently and clearly?
  - What was done?
  - By whom?
  - When?
- How were coding differences resolved?
- Was iterative checking done?
Abstracts are limited in length, so don’t expect to find:

- Controlling for confounding variables
- Addressing validity threats
- Characterizing subgroups within sample
Methods Example

IS A PRE-CLERKSHIP EXAMINATION EQUIVALENT TO THE PRE-THIRD YEAR GPA AS A PREDICTOR OF PERFORMANCE ON THE THIRD YEAR INTERNAL MEDICINE CLERKSHIP?

Methods: On clerkship day one, students take a faculty-developed, 100 question multiple-choice examination. Teacher evaluations are converted into total clinical points. The NBME subject examination in medicine is given in the 12th week. From 1997-2001, data from 585 students (88%) were available. Multiple linear regression (MLR) was used with a standardized z-NBME score and total clinical points score as separate outcome measures and the z-pretest score, USMLE Step One, and GPA as exposures. Correlations used Pearson's coefficient.
Results

Also know as:

- Findings
- Analyses
Results: Essential Elements

- Should flow from and relate to other sections
- Clearly organized
- Key findings are presented

- Statistical analyses/findings are presented
  - Actual data (numbers) should be presented
  - Congruent with Methods section
  - Descriptive data usually present
  - Inferential analyses, with p values, should be present
Results: Essential Elements

- Descriptive results characterize dataset:
  - Response rate to a survey
  - Participation rate in cohort studies
  - Numbers in control and intervention groups
  - Numbers and/or reasons for exclusion
  - Demographic information
Inferential Statistics

- Issues to consider:
  - Are all the important results presented?
  - Are there any omissions you expected to see?
  - Do the numbers add up? Are they consistent throughout the abstract?
  - Do the statistical tests used make sense with the data presented?
  - Do the authors present measures of functional significance?
Inferential Statistics

- Data and p values should be reported

  - Example: in the case of a correlation coefficient, the author should state the actual correlation coefficient and p value:

    - “r = .25, p<.05” vs. “the correlation between X and Y was significant (p < .05”).
What Influences Statistical Significance?

- **Sample size**
  - As sample size increases, smaller differences needed for significance

- **Multiple comparisons**
  - As number of statistical tests increases, likelihood of at least one statistically significant finding by chance alone increases
    - the p value can and should be adjusted for multiple comparisons
Functional Significance

- Statistical analyses with *significant* p values ideally should be accompanied by a measure of functional significance.
  - Examples: Effect size, confidence intervals, proportion of variance explained.

- Such measures help to show the size or importance of the statistical difference.
“Negative” studies

- Negative findings can be important for learning what works (or doesn’t work) in medical education

- However, need to consider if negative or non-significant findings due to lack of statistical power
  - Most common reason is insufficient sample size
  - With negative findings, author should include a comment about power in the abstract
    - If absent, this is a feedback point in the written comments
Results: Qualitative Studies

- Often, a description of the findings of the analysis
  - Themes and sub-themes that emerged
  - Differences among participants, groups may be discussed
  - Quantification of the results often not included or relevant
Results Common Problems

MIXES RESULTS & DISCUSSION

“Statistical analysis supports that our educational intervention should be broadly implemented as students improved their OSCE communication scores (65.5 vs. 58.5, p < .05, Effect size 0.8)”

PRESENTS RESULTS

“Students who received intervention scored higher on the end of clerkship OSCE communication station compared to controls, 65.5 vs. 58.5, p < .05, Effect size 0.8”
Results: Other Common Problems

- Inadequate sample size/response rate
- Statistical analyses inappropriate for type of data (e.g., nominal, ordinal, continuous)
- Numbers don’t add up throughout section
- Insufficient, inadequate, or absent data presented for interpretation
Results: More Common Problems

- Methods & Results incongruent
- p values
  - Excluded
  - Presented without data or test statistic
- “Results will be presented at the meeting”
- Results are disorganized, incomplete
Results: Example

Pre-clerkship Examination

- The pretest and GPA were moderately correlated ($r = 0.57$, 95% CI 0.50 - 0.63), while the USMLE correlated with both the pretest and GPA ($r = 0.66$, 95% CI 0.60-0.72; and 0.72, 95% CI 0.66-0.78, respectively). By univariate analysis, GPA, USMLE, and pretest scores were significant predictors of both outcomes. For teachers' points, the MLR model showed GPA and pretest explained 19% of the variance and remained significant predictors (pretest $p<0.04$ and GPA $p<.001$), but USMLE did not contribute. For the z-NBME MLR model, GPA, pre-test, and USMLE explained 43.7% of the variance (all with $p<0.001$).
Discussion

Also know as:

- Conclusions
Discussion/Conclusion

- This section provides space for:
  - Interpretation of results
  - General reaction to findings
  - Statements of limitations
  - Thoughts on study’s impact
  - Proposals for next steps
Discussion/Conclusion

- Related to problem statement/hypothesis
  - should pull all of the information together

- Interpretation of study’s findings
  - What do the results mean?
  - Were they as expected?
  - Do they answer the original hypotheses or problem statement?
Fair and Balanced?

- Have the authors provided a reasonable interpretation of what is presented in the Results?
- Are obvious limitations of their findings stated or balanced within the Discussion?
- Does the Discussion make assertions beyond the data?
Overall

- While ideal, word limits might prevent the author from:
  - providing a framework for application of the findings or next steps for study
  - discussing information that might be useful to share with the conference audience
- These omissions are acceptable
Discussion: Example

Pre-clerkship Examination

- A faculty developed pretest given on the first day of the internal medicine clerkship can identify students at risk of poor performance and may offer advantages over USMLE Step One and preclinical GPA.
Practical Do’s & Don’ts
Conforming to Submission Requirements

- As a reviewer, it is important to be familiar with the submission requirements.

- Some submission systems reject an abstract that exceeds the word count:
  - If the abstract seems long, check the word count and comment on it in the review if excessive.
Writing Style

- Writing style is an individual choice, however optimally using every word is important.

- Despite word limitations, correct grammar and spelling are essential.
General Rules

- Always write comments
  - Engage authors in feedback

- Your chance to convey:
  - Concerns/suggestions for improvement
  - Compliments
  - Reasons for ratings of abstract elements

- Consider writing comments FIRST, then use the rating form
Written Comments

- Be respectful
- Be specific
- Be constructive
- Be focused
Example of Comments

Poor

- “The worst abstract I’ve ever read!”
- “Data presentation was pitiful.”
- “Wonderful study!”

Improved

- “The authors need to attend to several areas before resubmitting; these include:
- “The authors need to present the actual data and statistical analyses.”
- “This was an excellent study for the following reasons:”
Final Recommendation

Definitely Include
Highly rated in most/all categories; highly relevant

Probably Include
Few concerning areas; can be improved before presenting. Relevant

Probably Exclude
Several areas of concern, will require significant revision to remedy

Definitely Exclude
Many areas of concern; fatal flaws present; needs substantial revision
Reviewer Etiquette
Are you Available to Review?

- Most reviews must happen within a discrete (and often short) time period

- Look at your calendar before you agree to review

- If the timing is bad, politely DECLINE!
Comfort with the Content?

- Are you comfortable addressing the specific topics of each submission?

- If you are not an appropriate reviewer, return it immediately so other reviewers can be found.
Time is of the Essence!

- Read and score abstracts as soon as possible after receipt
- The “due date” is the LAST DAY that program planners hope to receive reviews
- If you will not be able to meet the deadline – contact the program planners ASAP so that alternate plans can be made