statreviewer

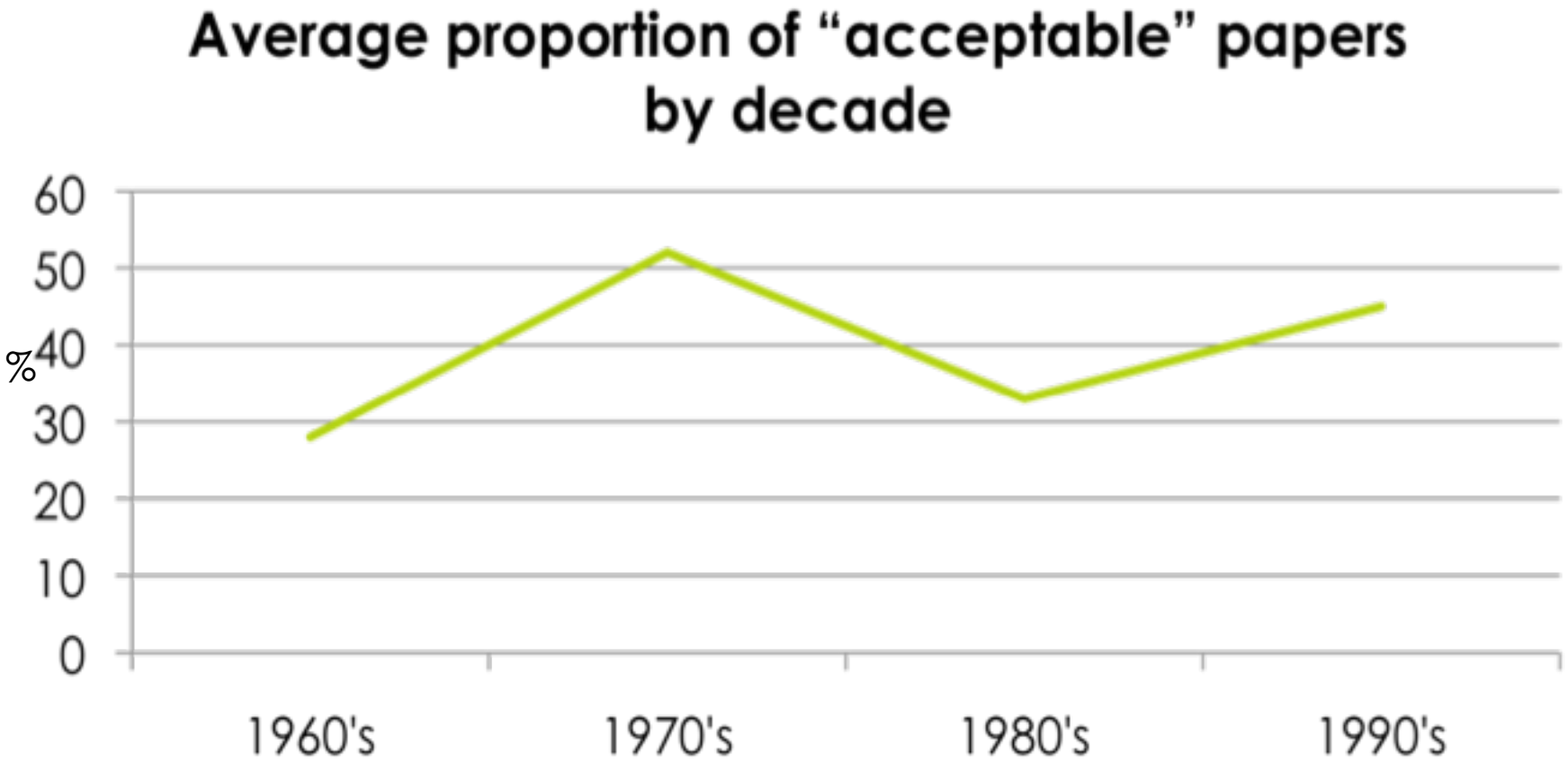
Automated Manuscript Analysis

Timothy T. Houle, PhD
Chief Scientist

The Problem with Statistical Reporting in Journals

- Incorrect choice of statistical techniques
- Flawed application of statistical methods
- Omission of crucial information to replicate study
- Erroneous conclusions
- Choice of statistical methods susceptible to bias or exaggerated claims that will fail to be replicated

How big of a problem is poor statistical analyses in medical journals?



Adapted from: Altman D. Statistical reviewing for medical journals. Statistics in Medicine, 1998; 17:2661 - 2674

Statistics and Peer Review

"…the majority of statistical analyses are performed by people with an inadequate understanding of statistical methods. They are then peer reviewed by people who are generally no more knowledgeable"

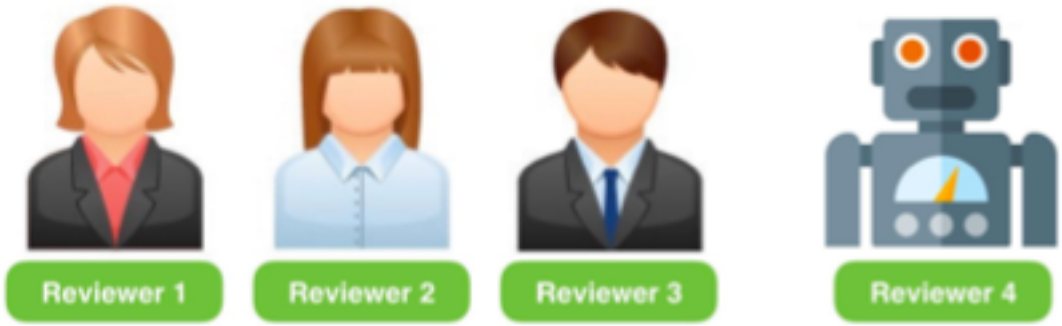
-Douglas Altman

- Lack of experts with requisite skills
- Reliance on non-experts – prone to reinforce common errors and misconceptions
- Statistical consultants can be costly ($75-$200 per review)
- Consultants often overstretched
  - Rushed reviews
  - Inconsistent application of standards

Our Solution

Add Another Reviewer

Always Available. Lightning Fast Response.

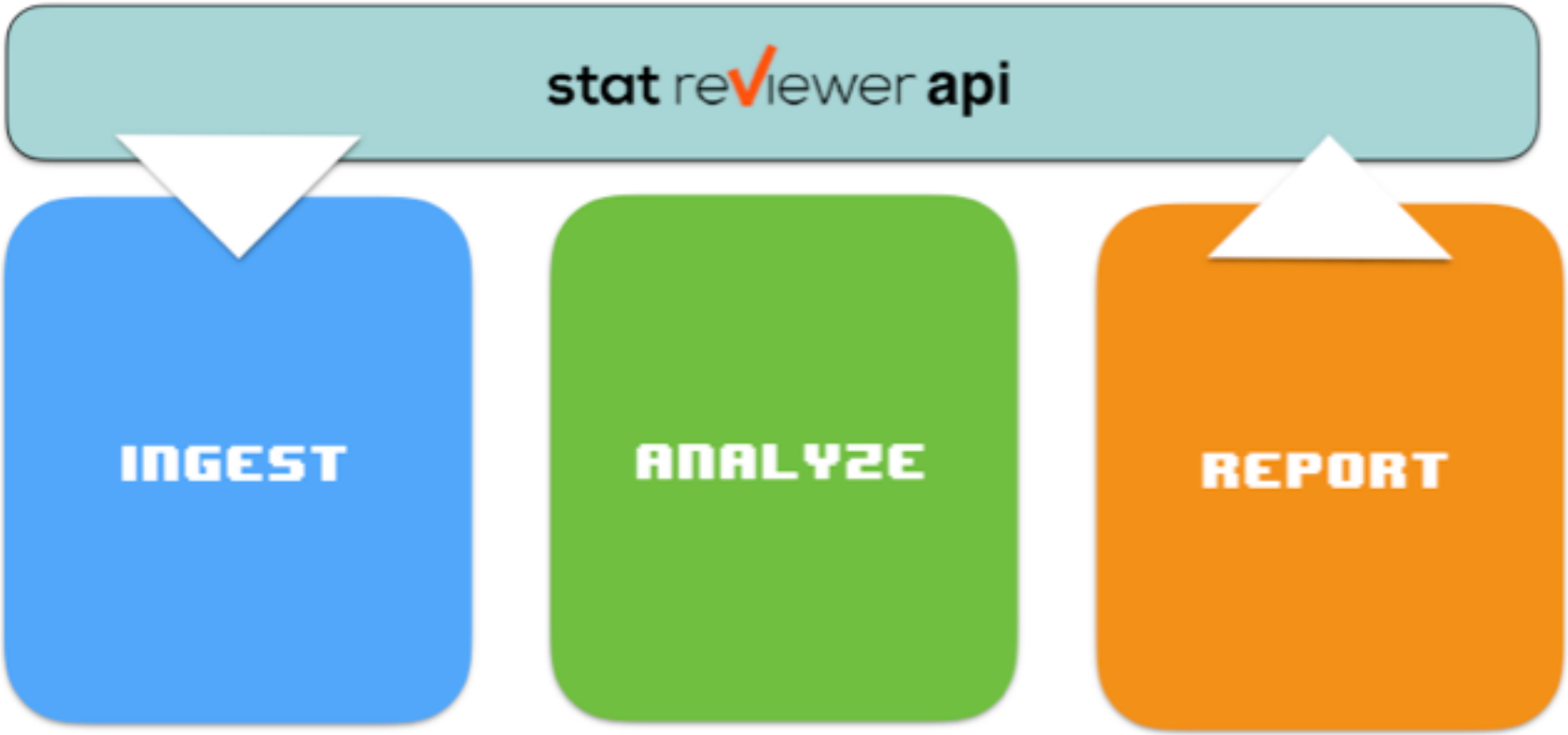


Hint: This is StatReviewer

statreviewer...

- Identifies problems with scientific manuscripts
- Creates a statistical and methodological issue report
- Is instantaneous and fully automated

What's Inside



INGEST

- ✓ Document files or plain-text accepted
- ✓ API-driven
- ✓ Document sections are extracted and prepared for scan

- ⊘ For best results, key sections should be present and identifiable using common titles

  Such as “Abstract”, “Introduction”, “Methods”

ANALYZE

Thousands of Checks

A robust arsenal of algorithms are employed to identify everything from large document patterns to small markers

Expert-Informed Artificial Intelligence

We use industry standards as a basis for Machine Learning, not blind unsupervised learning

Always Improving

Continuous feedback loops hone and advance our platform

ANALYZE

What We Find

- Reporting style (e.g., precision of estimates, confidence intervals, etc.)
- Error checking (e.g., inferential statistics, proportion checks, proper confidence interval coverage)
- Inferential statistics (e.g., types of tests, assumptions, reporting of values)
- Descriptive statistics (e.g., style of reporting, appropriateness of statistic, avoidance of SEM).
- Ethical approvals (e.g., IRB, ACUC, ethical board approvals, consent)
- Type-I error control
- Statistical power analyses
- Interpretation of results (e.g., misinterpretation of statistical significance, confidence intervals).
- Many, many more…

ANALYZE

Statistical / Methodological Review

Reporting	vs	Design
What was done?		Are these tests appropriate?
What was found?		Are the estimates biased or incorrect?
How to interpret?		Are the conclusions appropriate?

REPORT

4 Types of Reports

Classic "Peer" Review

Natural language algorithms create a human-looking review containing identified issues and links to help & education.

Editorial Review

Details relevant to editorial decision can be withheld from the shared report and presented as a separate report to only the editorial staff.

Checklists

When only an overview of guideline adherence is needed, a checklist style report can be generated.

Scores (in development)

Quality or Risk scores can be generated and used as a quick glance tool - but are not intended to replace an editorial decision.

Title

1 The subject of the study is very likely to be a keyword in a search engine. To make your study more relevant to your readers, include the subject of the research in the actual study title.

More Information About This Issue

Most scientific topics cover a diverse set of subjects. Readers often wish to know if a particular study was done using animals or humans. Within human research participants, there is a vast array of subgroups such as age, racial, diagnostic, etc. To help your readers identify the relevance of your study to their interests it is helpful to report the research subject(s) directly in the title.

Examples

This title: "Tumor size and lymph node status on survival"
Can be revised to this: "Tumor size and lymph node status on 5-year survival in women with breast cancer"

Additional Resources

For more information, the University of Southern California Libraries has a list of excellent suggestions to assist authors in devising their titles:
http://libguides.usc.edu/writingguide/title

Results

> 10 To properly interpret the study, a reader must be able to evaluate potential bias due to lost, missing, or excluded data. In that regard, please report if any data were missing or lost for any reason. It is important to report the extent of missing data using frequency counts (e.g., "6 animals did not survive and were excluded from the analysis"). If there were no missing data, this fact should be stated (e.g., "There were no missing data").

More Information About This Issue

Missing data, or data that is incompletely observed, is very common in all types of research designs. In bench research there can be failed experiments, corrupted or flawed measurements, or human error in collecting measurements. In research involving direct interaction with human participants, missing data can occur for the same reasons or for omitted responses by the participants. In systems research, missing data may occur due to a host of reasons beyond the investigators control. No matter the reason for the missing data, it is crucial to report the existence of missing data, and the extent to which observations were missing in the analysis.

All statistical analyses involve the selection of observations from a larger population. Failure to consider the biasing effect of missing observations can often lead to erroneous conclusions.

Examples