Updates on Requirements for Scientific Rigor and Reproducibility

Lique M. Coolen, PhD
Associate Dean, College of Arts and Sciences
Professor, Biological Sciences

January 22, 2020
Objectives of workshop

Discuss efforts to increase scientific rigor and reproducibility

- Reproducibility issues
- Grant writing guidelines
- Publication requirements
- Open data sharing
- Pre-registration, pre-prints
- Data recording and management

Audience Discussion
What does scientific rigor mean?

Scientific rigor is the strict application of the scientific method to ensure robust and unbiased experimental design, methodology, analysis, interpretation and reporting of results. This includes full transparency in reporting experimental details so that others may reproduce and extend the findings.

(Definition: NIH)
Growing awareness: need to enhance rigor and reproducibility

https://www.nih.gov/research-training/rigor-reproducibility
Why did this become such a discussion?

Failures to replicate published studies
- Between labs
- Within labs

NIH-funded projects specifically to test reproducibility
- Pre-clinical
- Clinical
Consequences of replication failures in research

- Endangers the foundational “building block” nature of scientific knowledge
- Impedes progress in treating and curing diseases
- May expose patients to unnecessary or potentially harmful agents in clinical trials based on faulty preclinical evidence
- Wastes limited time and resources
- Erodes public confidence in and support for scientific research
What underlies to many failures to replicate?

- NOT intentional scientific fraud or fabrication of results
- Flawed experimental design
  - Lack of positive and negative controls
  - Bias, biological variability
- Flawed data analysis
  - Underpowered (hence increased risk for false positives (Type I error) and false negatives (Type II error))
  - Failure to distinguish between pilot/discovery studies and final/hypothesis-driven studies
- Flawed reporting
- Publication of positive results
  - Omission of results that do not fit the story
Begley’s “Six red flags for suspect work”

  - Were experiments performed blinded?
  - Were basic experiments repeated?
  - Were all the results presented?
  - Were there positive and negative controls?
  - Were reagents validated?
  - Were statistical tests appropriate?
Best practices for experiment planning

- Minimizing unintended bias
- Pre-determination of sample sizes (power analysis)
- Distinguishing discovery experiments vs. hypothesis-testing
- Pre-identification of stopping points
- Pre-defined procedures for data handling
- Transparent reporting
- Pre-registration of study protocol and analysis plan
Best practice: minimizing unintended bias

- “Bias is unintentional and unconscious. It is defined broadly as the systematic erroneous association of some characteristic with a group in a way that distorts a comparison with another group…The process of addressing bias involves making everything equal during the design, conduct, and interpretation of a study, and reporting those steps in an explicit and transparent way.”
  - Ransohoff & Gourlay, (2010), J Clin Oncol

- Best Practices:
  - Random assignment to groups
  - Blinding
  - Outcome expectation
    - “testing” a hypothesis vs. “proving” a hypothesis
  - Pre-registration of study protocol and analysis plan
Best practice: Pre-determination of sample size

• Routinely done in clinical trials
  o Needs to be more routinely implemented in pre-clinical studies

• Requires consideration of:
  o Endpoint sensitivity, and thus required “N” of subjects
  o Expected data variability
  o Possible effect size
  o Desired level of confidence
Common practices: P-hacking

- **P-hacking** = reanalyzing data in many different ways to yield a target result [Insel (2014) *NIMH Director’s Blog]*

- Common examples of P-hacking:
  - Finding a trend in the data and performing additional experiments to increase “n”
    - “testing to a foregone conclusion”
  - Carrying out multiple analyses/comparisons and reporting only the one(s) that significantly differ
    - “cherry picking results”
  - Finding a trend in the data and removing “outliers” so comparisons become significant
  - Conducting multiple analyses on a group of data without accounting for multiple comparisons
Summary: Rigor in Experiment Design and Analysis

- Blinding
- Randomizing
- Distinguish between pilot/discovery study and final/hypothesis-based study
- Rigorous experimental design
  - Online platforms being generated (eXperimental Design Assistant - Tietronix Software Inc)
- Automation of analyses
- Proper use of Statistics
  - Sample size
  - P-Hacking and removal of outliers
Many Recent Resources for Reading

Defense of the Scientific Hypothesis

From Reproducibility Crisis to Big Data

Bradley Alger

- Includes original data from surveys of scientists regarding their experiences with, and opinions of, the hypothesis, as well as of their actual use of the scientific hypothesis in their published work
- Includes in-depth critiques of published works (books and articles) that have cast aspersions on the place of the hypothesis in modern science
- Includes a thorough analysis of the hypothesis, including detailed examination of the ideas of Karl Popper and John Platt to which few scientists are ever exposed
Grant proposals: Describing Scientific Rigor

- Reviewers are instructed to review rigor of:
  - previous research on which proposal is based
  - proposed studies

What is scientific rigor?

- The structured and controlled application of the scientific method using the highest standards in the field

  - Includes considerations in:
    - Experimental design
    - Minimizing unintended bias
    - Data analysis
    - Transparent reporting
    - Reproducibility
Peer reviewed publications: Transparent Reporting

- Detailed methods section
  - No page limits
  - No reference to methods “as in” previous papers without providing details
- Detailed description of statistical methods and findings
  - Effect sizes
- Resource validation and authentication (also in NIH grants):
  - https://scicrunch.org/resources (RRID)
  - http://antibodyregistry.org
Peer reviewed publications: Transparent Reporting

Open data sharing

- All data made available with publication
  - Through journal-websites

- Online data sharing
  - Open Science Framework
  - https://osf.io
Pre-print

Benefits of Preprints

We see preprints as an important step toward a more open and transparent peer review process — one that brings with it tremendous benefits for both individual authors and the broader scientific community.

**RAPID DISSEMINATION OF YOUR RESULTS**

Preprints allow you to share your results when you’re ready — whether you’re researching an emergent disaster, applying for a grant, or just excited to broadcast your work to a wider audience.

**ESTABLISHING PRIORITY**

It’s common for researchers to achieve a similar advance at around the same time but the publication process can artificially delay one paper or favor another. Posting preprints allows researchers to publicly date stamp their discoveries.

**INCREASED ATTENTION (AND CITATIONS!)**

The sooner research becomes available, the sooner it can begin to receive views and citations. In this case, common sense is backed up by evidence. Research shows that public posting increases the number of times papers are viewed and cited.

**CAREER ADVANCEMENT**

Preprints enable you to showcase your latest work for grant, hiring, or tenure committees. A link to a publicly posted preprint is more illustrative and compelling than a title on a CV with the annotation “in development” or “under review.”

**COMMUNITY**

Preliminary feedback helps authors improve manuscripts. Collegial discussion can lead to new ideas, follow-up studies, or collaborations with other research groups. Plus, you can cite your preprint in your letters of inquiry.

**UNLIMITED AND TIMELY UPDATES**

From the moment a preprint appears online to the day that the article is published in a peer reviewed journal, you can make as many updates as you want or need. Each version is numbered and incorporated into the preprint record.

https://www.plos.org/why-preprint
Pre-registration of experimental methods

• Pre-registration of experimental design
  • Peer reviewed
  • Not peer reviewed

• Also platform to deposit data sets to be shared with private groups
Updates in Practices for Data Recording and Managing

• Electronic formats for:
  • Pre-registration of experimental protocols
  • Complete logs of all activities
  • Complete logs of all data entry
  • Complete records of all data edits and analyses
  • Easy to share with collaborators across the world
  • Easy to share with large community once published
Data recording, management and repository tools: suggested by audience

- Google docs: allows for version tracker (shows times of edits and people who performed the edits)
- Github: through Kent State:
  - [https://github.com/kent-state-university-libraries](https://github.com/kent-state-university-libraries)
Cultural factors outside of the actual conduct of research that undermine rigor in laboratory practices

- Even with the best intentions and implementation of best practices, rigorous research practices are often undermined by the fabric of the sociology and culture of the modern research enterprise.
- This can include time pressures such as grant deadlines, promotion and tenure reviews, and the overall pressure to publish work of “broad interest,” which often means multi-lab collaborations.
- Time pressures can also lead search as well as promotion and tenure committees to use proxies of excellence, such as journal impact factor, and to under-value strong contributions that advance the field.
- These issues affect scientists from every career stage and path, in all areas of their work — research practices, laboratory management, publication, grant and manuscript peer review, mentoring and recruiting, promotion and tenure considerations, and more.
Additional thoughts:

• Does the current scientific training properly prepare the next generation?
  • In the lab? In graduate and undergraduate programs?

• Do faculty and trainees receive sufficient support from their institutions?
  • If not: what is needed?

• Will all traditional manuscript disappear?
  • All data reporting and sharing will be electronic
Thank You!
Email: jcoolen@kent.edu