Validating Research: Some Considerations

Marcia McNutt

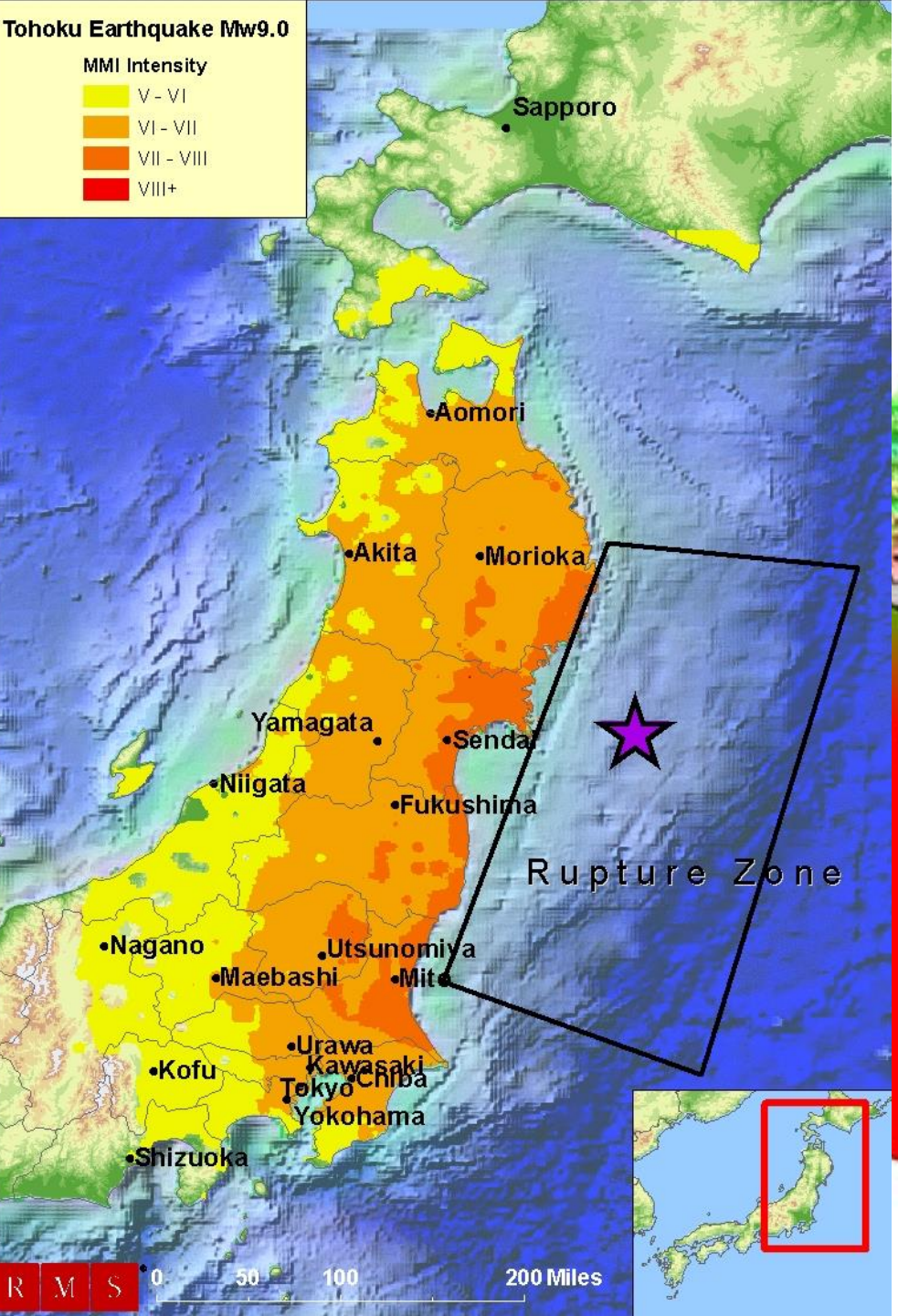
Editor-in-Chief

Science

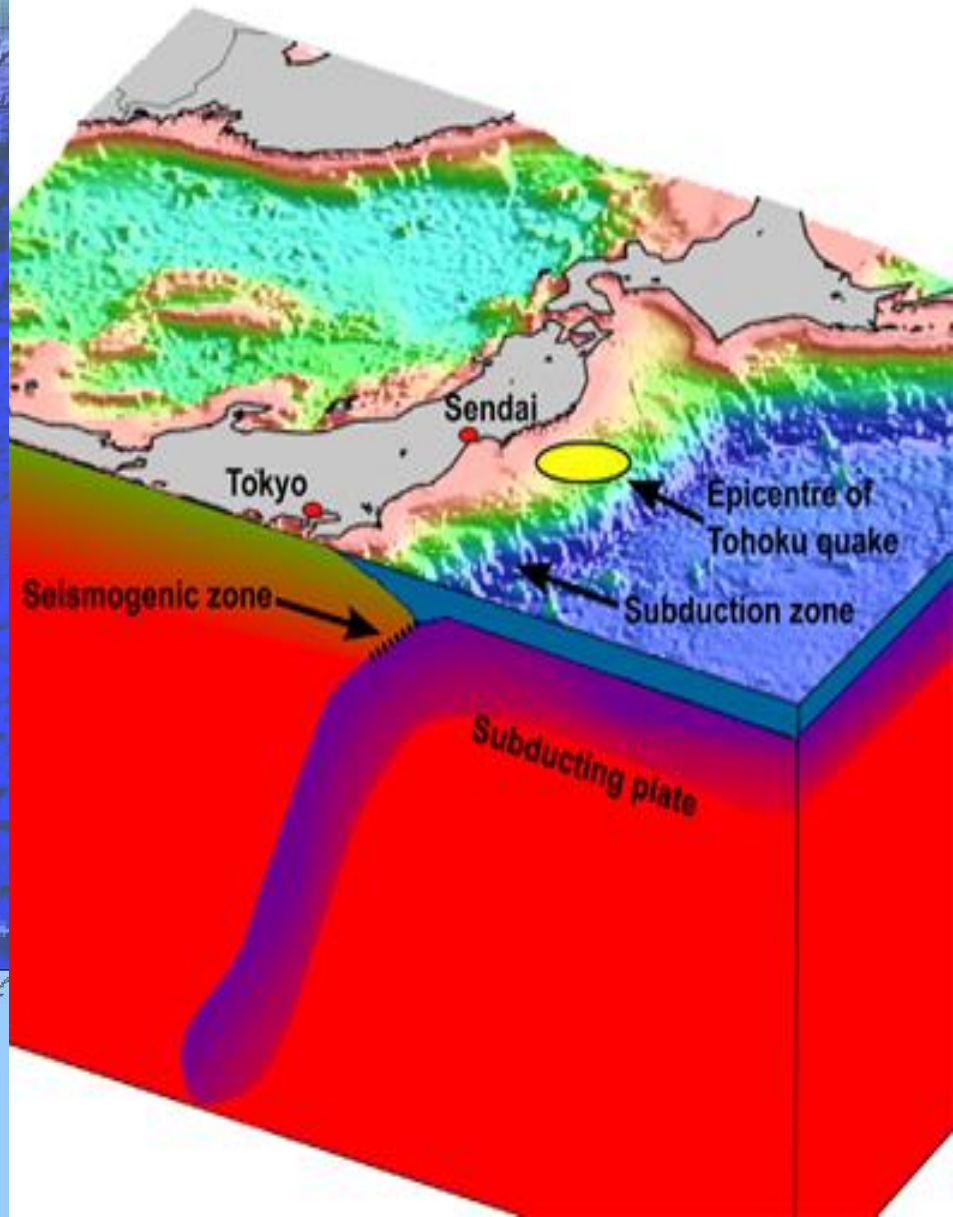
Spectrum of Reproducibility*

- One End Member (minimum standard)
  - Repeatability: Another group can access the data, analyze it using the same methodology, and obtain the same result.
- Other End Member (gold standard)
  - Replication: The study is repeated start to finish, including new data collection and analysis, using fresh materials and reagents, and obtain the same result.

*Ioannidis and Khoury, *Science*, Special Issue on Data Replication & Reproducibility, **334**, December 2011.



Science
AAAS



Reasons for Lack of Replicability

- Information withheld (not enough space, not deemed important, etc.)
- Tacit knowledge the practitioner doesn't even know he/she possesses
- System not sufficiently known (not all independent variables controlled)
- False positives (or negatives…)

No Substitution for Experimental Groups Observing Each Other's Protocols

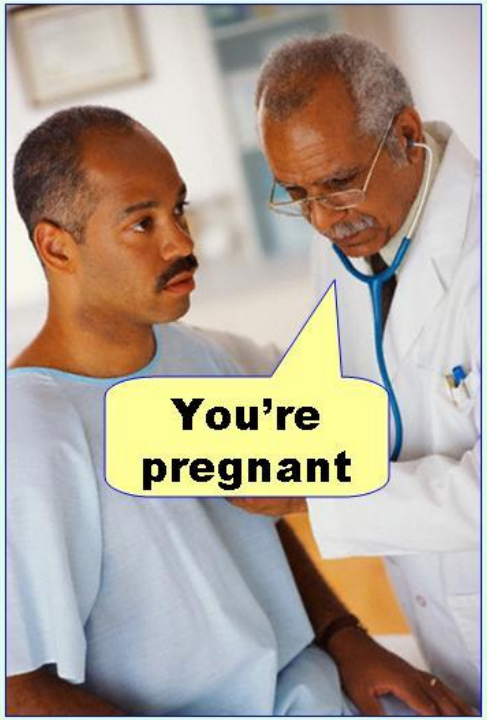
Pilot Sullenberger ditched US Air jet in the Hudson after double bird strike. He described how he smelled 'burning birds' as both engines shut down.

Not all independent variables have been identified as such…

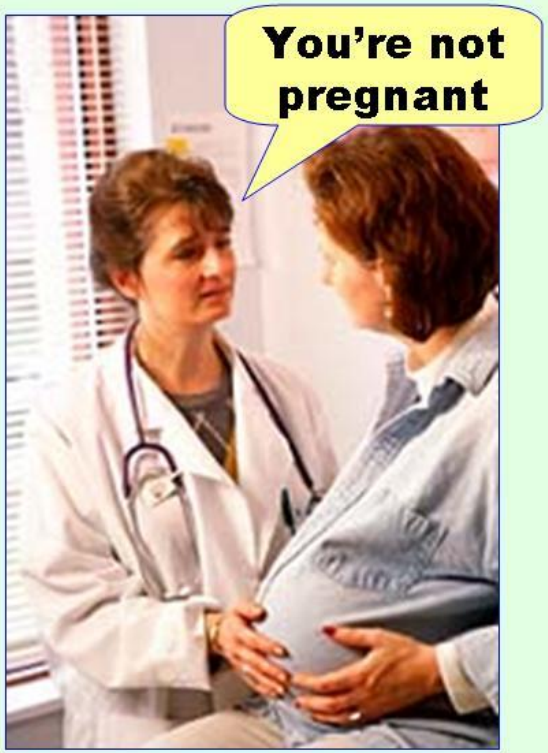
The smell of a man makes mice more stressed than the smell of a woman

The Problem of Bias

Type I error
(false positive)



Type II error
(false negative)



Much Attention to Fraud/Misconduct

- Note that much current effort does not bear on fraud
  - Requiring posting raw data in public place
  - Transparency in analysis methods and approaches
- Replication (not repeatability) will uncover true fraud
- Many times whistle blowers reveal fraud
- At least most agree on path for best actions in case of fraud

Mistakes are more common, more difficult

- Repeating analysis can uncover weakness in reported result
- The culture of science can work to our advantage
- The realities of the low probability of obtaining science funding/publication in top journals can work to our disadvantage
- Authors, institutions, funders have an aversion to retractions for honest mistakes. Overkill?

The Right Incentives

- Want to encourage PIs to check results of others (how can funders encourage repeating studies?)
- Need to encourage technical comments to correct the record on results that are not repeatable.
- Reward PIs who consistently produce high-quality results

Role of Funding Agencies

- Make panels alert to criteria for reproducibility at proposal stage as it needs to be part of the experimental plan and will have budget implications.
- Consider whether reproducing key experiments is worth funding.
- Preferentially support PIs who produce reproducible research.

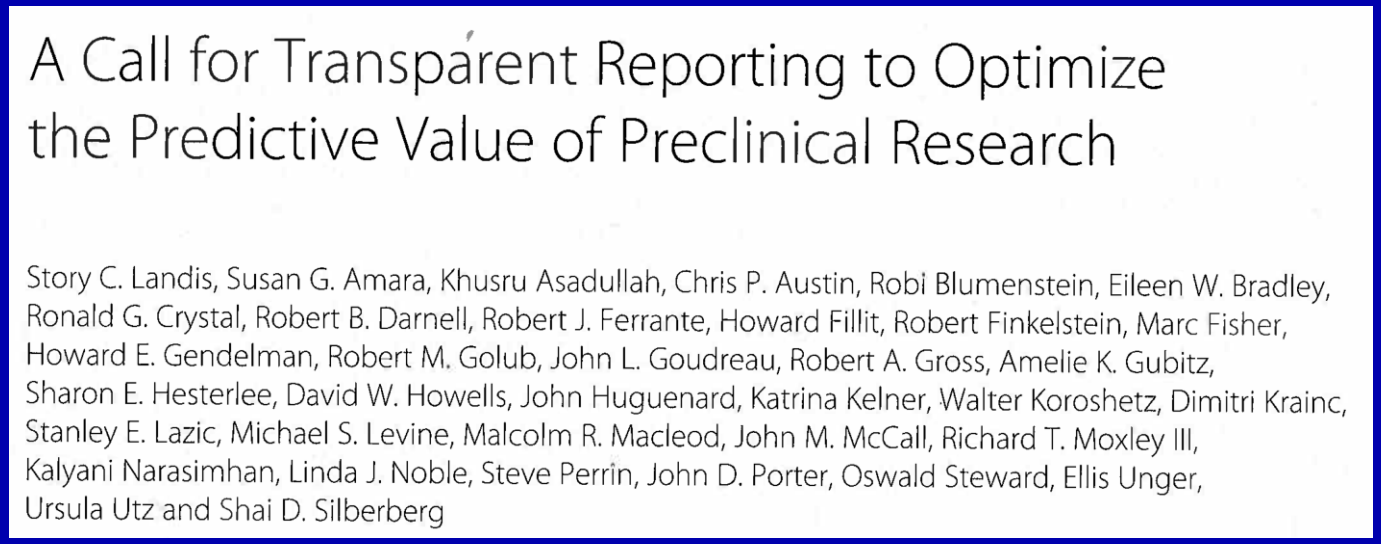
The Role of Journals

- Prestigious journals have some role in adopting standards because scientists want to publish there.
- But traditional journals are facing more competition from new publishing models (OA, preprint servers), not all of which have same requirements on authors for reproducibility.
- Journals are likely the first to know when research they published is not reproducible; have obligation to alert the scientific community.

What *Science* is Doing

- Recently announced a set of new initiatives to increase reader and reviewer confidence in studies published in *Science*
- Adding additional members to the BoRE (Board of Reviewing Editors) from the statistics community with the help of the American Statistical Association

Reproducibility: Preclinical Studies



A Call for Transparent Reporting to Optimize the Predictive Value of Preclinical Research

Story C. Landis, Susan G. Amara, Khusru Asadullah, Chris P. Austin, Robi Blumenstein, Eileen W. Bradley, Ronald G. Crystal, Robert B. Darnell, Robert J. Ferrante, Howard Fillit, Robert Finkelstein, Marc Fisher, Howard E. Gendelman, Robert M. Golub, John L. Goudreau, Robert A. Gross, Amelie K. Gubitz, Sharon E. Hesterlee, David W. Howells, John Huguenard, Katrina Kelner, Walter Koroshetz, Dimitri Krainc, Stanley E. Lazic, Michael S. Levine, Malcolm R. Macleod, John M. McCall, Richard T. Moxley III, Kalyani Narasimhan, Linda J. Noble, Steve Perrin, John D. Porter, Oswald Steward, Ellis Unger, Ursula Utz and Shai D. Silberberg

*

- A pre-experiment plan for handling data (not on the fly)
- Sample-size estimation to ensure appropriate S/N
- Randomization in sample treatment
- Blind conduct of the experiment

Nature*, **490, 187, 2012.

Reproducibility: All Studies

- Upon acceptance, ask reviewers/editors to select papers with unusually excellent treatment of data and samples to volunteer to write up their approach in as general terms as reasonable
- Collect a compendium of treatments across all fields of science that will provide input for NINDS-style workshops later in 2014 selected areas

Reproducibility

- First of three workshops sponsored by the Arnold Foundation scheduled for November 3-4, 2014 at the Center for Open Science in Charlottesville, VA
- Focus on the Social and Behavior Sciences
- Attendees to include researchers, journal editors, funding agency reps
- Next up: Field studies, spring 2015.

Role of Universities

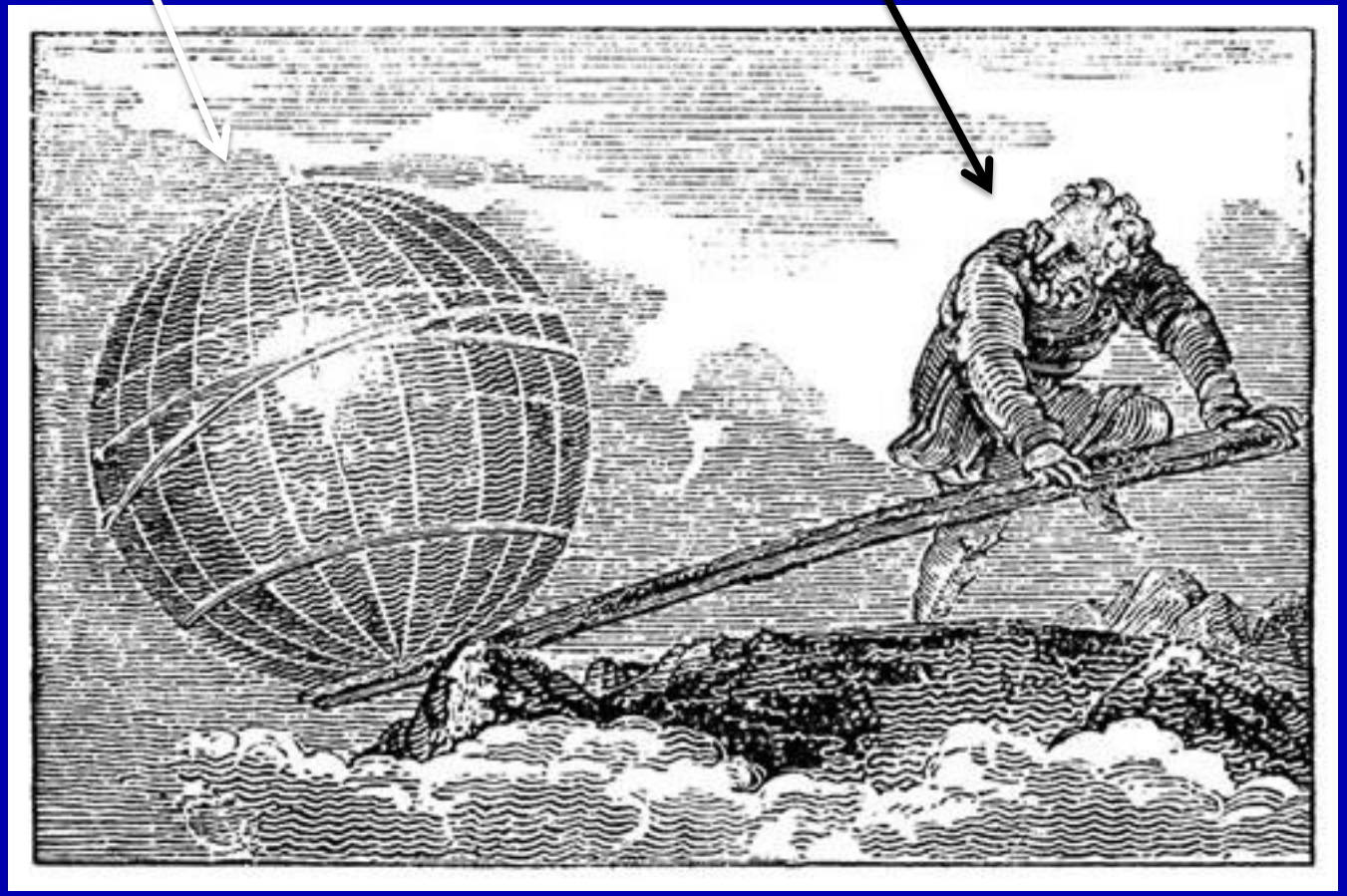
- Responsible for training future and current researchers in the scientific method and best practices to improve reproducibility.
- Can reward researchers who produce reproducible results and withhold rewards from researchers who produce non-reproducible research.

Role of Scientific Societies

- Consider honoring those who consistently produce reproducible research.
- Devote special sessions at scientific meetings to the topic of best practices in reproducibility.
- Adopt reproducibility guidelines for society publications.

This needs to be a team effort.

Science
AAAS



I have heard it said that scientific journals use leverage to promote reproducible research from the research community.

However, in my experience the better analogy for the relationship is that of a well choreographed pair of dancing partners.