

XXIV Congresso Nazionale Siset
Abano 9-12 Novembre 2016

Come scrivere un articolo scientifico

P. Bucciarelli - A. Tosetto

Behind every great paper there is...

- A relevant study question
 - Study hypothesis vs null one
 - Sample size
- A good design
 - Sources of errors
 - Confounding

Relevant Clinical Questions

- Background - “What is it?”
 - General information on a condition or disease: prevalence, incidence, association with risk factors
 - “Etiology” setting: look for causes

Relevant Clinical Questions

- Background - “What is it?”
 - General information on a condition or disease: prevalence, incidence, association with risk factors
 - “Etiology” setting
- Foreground – “What do I do for this patient?”
 - Intervention/Investigation
 - Comparison Intervention/Investigation
 - Outcome (Patient-Oriented)
 - Therapy/Prognosis: look for effects

«Good» study designs

For therapeutic interventions

Experimental designs

- RCT
- Absolute risks

For etiology or prognosis:

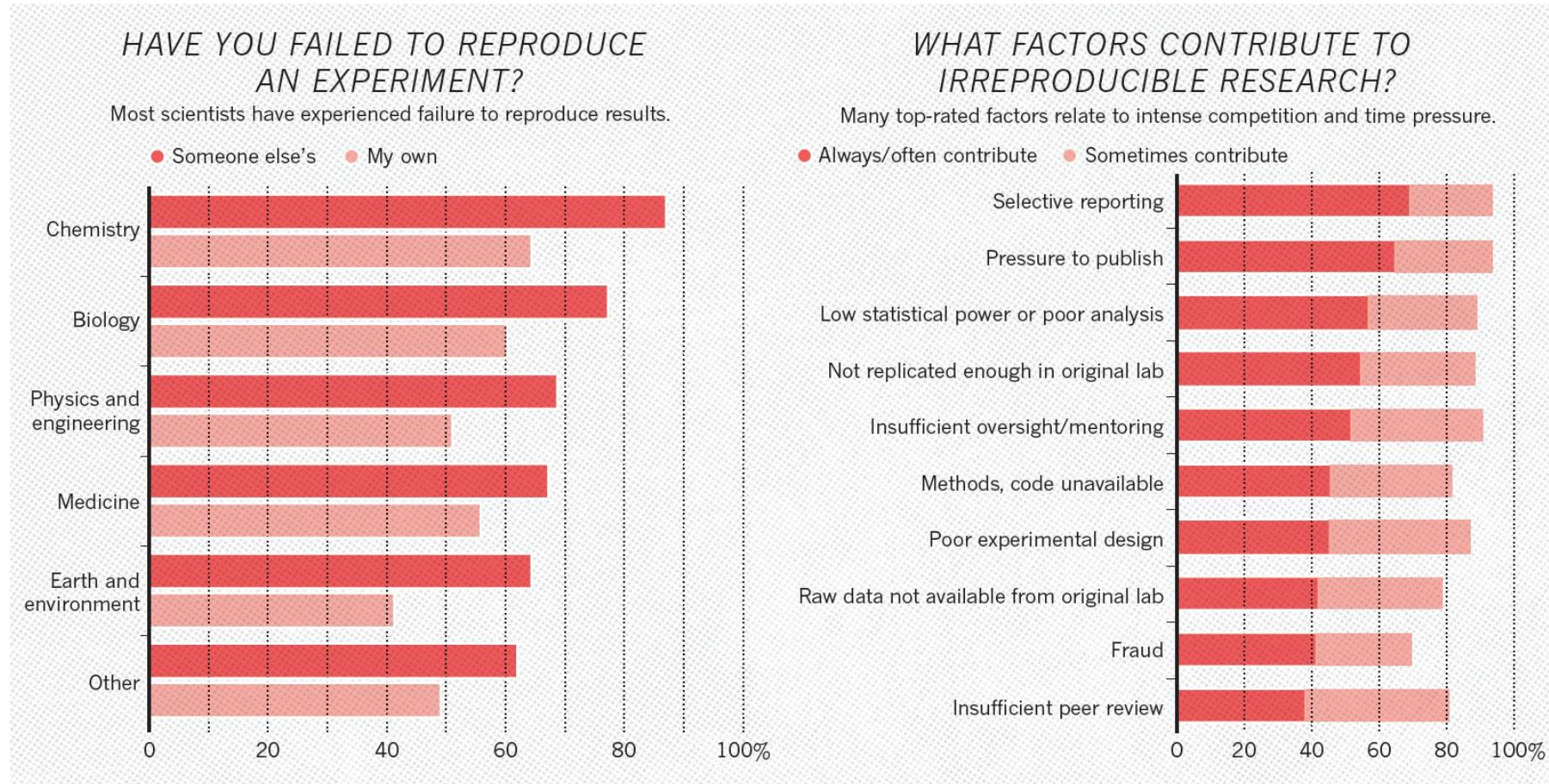
Observational designs

- Retrospective/Cohort
- Relative risks

Some common problems

- Few “good” retrospective designs
 - Case and controls from same area/time frame
- Uncontrolled Bias in (retrospective) cohorts
 - Handling of missing is missing
 - Patients’ specific covariates NA (e.g., administrative-driven databases)
- Missing study hypothesis (... and sample size!)
 - Fishing expeditions
 - *Post-hoc* hypothesis
 - Convenience sampling: pilot testing \neq random sampling

Irreproducible research?



Rules for Effective Statistical Practice

1. Shift from statistical techniques to scientific questions
2. State why you performed a specific statistical tests
3. Keep it Simple
4. Careful design & data collection can greatly simplify analysis
5. Provide Assessments of Variability
6. Check assumptions
7. Make Your Analysis Reproducible

Reproducible research

- A study can be truly reproducible when it satisfies at least the following three criteria.
 - All methods are fully reported.
 - All data and files used for the analysis are (publicly) available.
 - The process of analyzing raw data is well reported and preserved.
- Therefore:

Same data + Same script = Same results

Start with good data collection

- Prefer an EDC over Excel
- Commercial solutions available
- Open-source
 - REDCap
 - OpenClinica

Keep together data, analysis and reports

- Data should be “ready” to be shared to other researchers
 - IPD meta-analyses, study replication
 - Non-proprietary data formats
- Stat scripts should be referenced and available in a non-proprietary format
- Reports drafted directly from scripts

Some open-source tools

- RStudio + R
 - Import data directly from Web-based EDC
 - Data management & analysis
- Markdown, Knit
 - Directly outputs R results to PDF, Word, HTML
- Shiny
 - Build web sites to display HTML pages built with Knit
- GitHub
 - Collaborative data-analysis tracking version changes