

Reproducible Research: Why and How

Part 2: Design Solutions

Sam Harper



McGill

Department of
**Epidemiology, Biostatistics
and Occupational Health**

SER Pre-Conference Workshop
2020-10-30

2. Design Solutions

2.1 Preregistration

2.2 Pre-analysis plans

2.3 Reporting guidelines

2. Design Solutions

2.1 Preregistration

2.2 Pre-analysis plans

2.3 Reporting guidelines

What is study preregistration?

A detailed
study
proposal that
is:

Time stamped
Records and publicizes time and date.

Read-only
Can't be modified.

Registered prior to data collection/access
Robust to fieldwork, data snooping.

What is preregistration?



Common / required for publishing most RCTs

Controversial for observational studies.

Idea is to help *reduce publication bias*, since registered studies may be followed over time.

No guarantee anyone will publish.

Also can provide intellectual provenance of your ideas and hypotheses.

Good for planning and hypothesizing, **not a straightjacket**.

Why preregistration?

1. It's *not* about minimizing Type 1 errors.
2. It *is* about:
 - Allowing others to transparently evaluate the credibility of the analysis.
 - Assuring that all of the evidence is available for synthesis.

Why not preregistration?

- Observational studies are hard.
- Manuscripts may adhere to registrations rather than reality.
- May discourage innovation/exploration.
- Pre-specification is irrelevant to the credibility of inference.
- Severe tests of hypotheses are more important than pre-specification.

Should Preregistration of Epidemiologic Study Protocols Become Compulsory?

Reflections and a Counterproposal

Timothy L. Lash^{a,b} and Jan P. Vandenbroucke^{c,d}

There is an ongoing debate regarding preregistration of epidemiologic study protocols.^{1–4} We examine the basic idea that preregistration of study protocols and their associated hypotheses would enhance the reliability of observational research. We define instances in which preregistration would be useful, and we support a counter-proposal: a public registry containing descriptions of collected epidemiologic data.

A decision to institute compulsory preregistration of protocols for observational studies—to be enforced by editors and reviewers as sometimes suggested^{1–3}—is not to be taken lightly, and should not be endorsed solely on the basis of an analogous system instituted for randomized trials. Negative reactions toward compulsory registration have been published elsewhere.^{5–12} Note that it is the compulsory preregistration of protocols that is most at issue. There are already mechanisms by which epidemiologists can voluntarily preregister their protocols,¹³ if they feel preregistration is advantageous. The open question is whether such preregistration should be required in order for observational research to be published in leading journals (assuming the same “enforcement” mechanism would be adopted as for clinical trials).

We examine the validity of the analogy between randomized trials and observational studies with regard to the value of preregistering protocols. We then examine the idea that prespecification of a hypothesis enhances the credibility of results, and that avoidance of “false positives” should always be a primary concern. We discuss research settings when preregistration of an observational study protocol might be of value. Finally, as a

Where can you pre-register your study?



- The [AEA registry](#) includes the option to upload PAPs. Search under "advanced options" for studies which include PAPs.
- [EGAP](#) is a registry for political science studies, some of which include pre-analysis plans.
- [3ie](#) also has a database (RIDIE) of ongoing international development impact evaluations.
- The Open Science Framework ([OSF](#)) also invites pre-registered studies and PAPs.
- Many clinical trials in the US must be registered with [Clinicaltrials.gov](#).

Where can you find templates for preregistration?

See the [page](#) at Open Science Foundation

Writing up pre-registered studies

1. Include a link to the registration
2. Report *all* pre-registered results.
3. Explain and justify deviations.
4. Non-registered analyses appropriately described as "exploratory" or "hypothesis generating".

Why does preregistration matter?

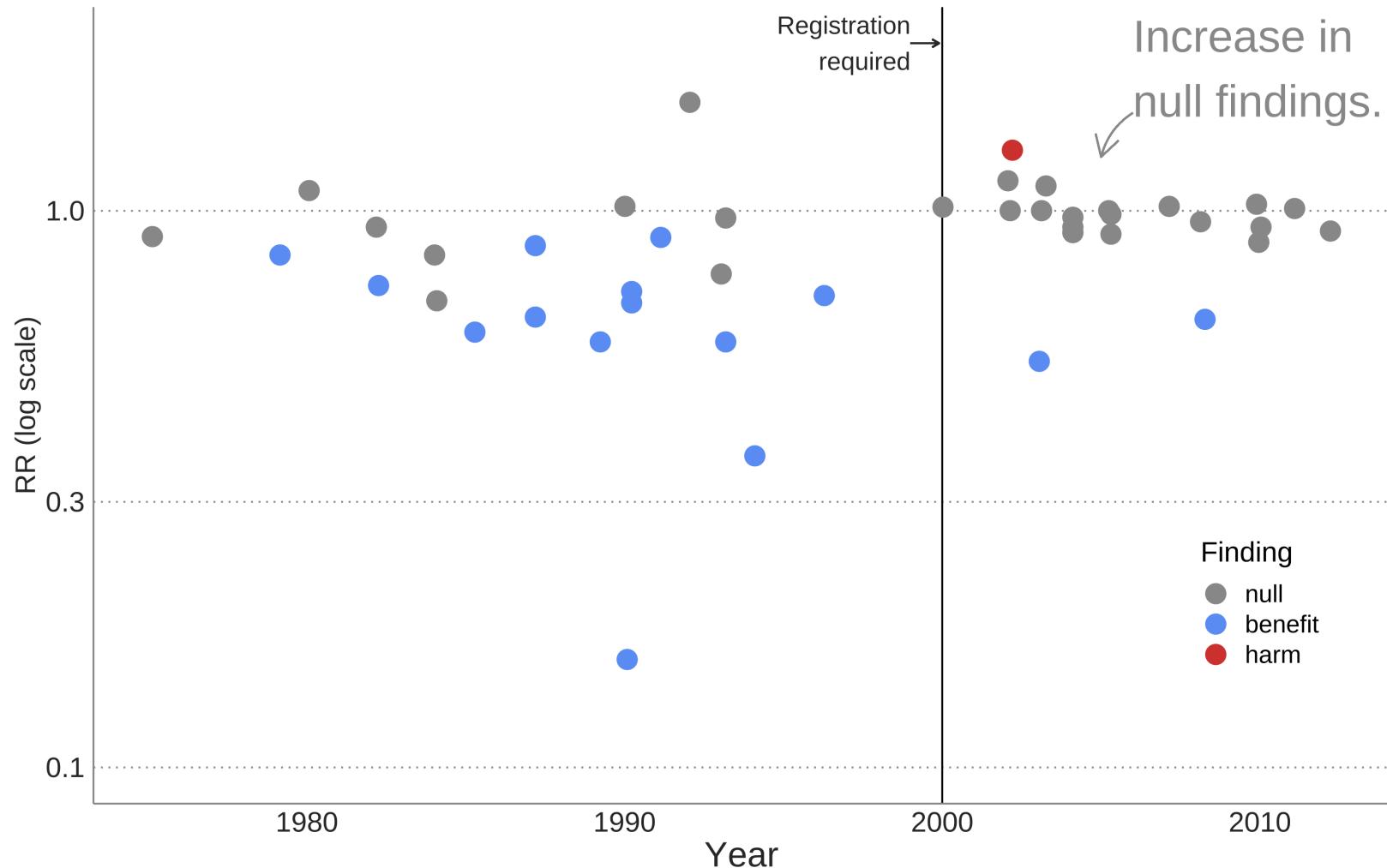
Evidence synthesis should be on *all* the evidence.

Distorts planning of future studies.

Unethical and wasteful.

Registration is useful

In 2000 NHLBI required the registration of primary outcome on ClinicalTrials.gov for all their grant-funded activity.



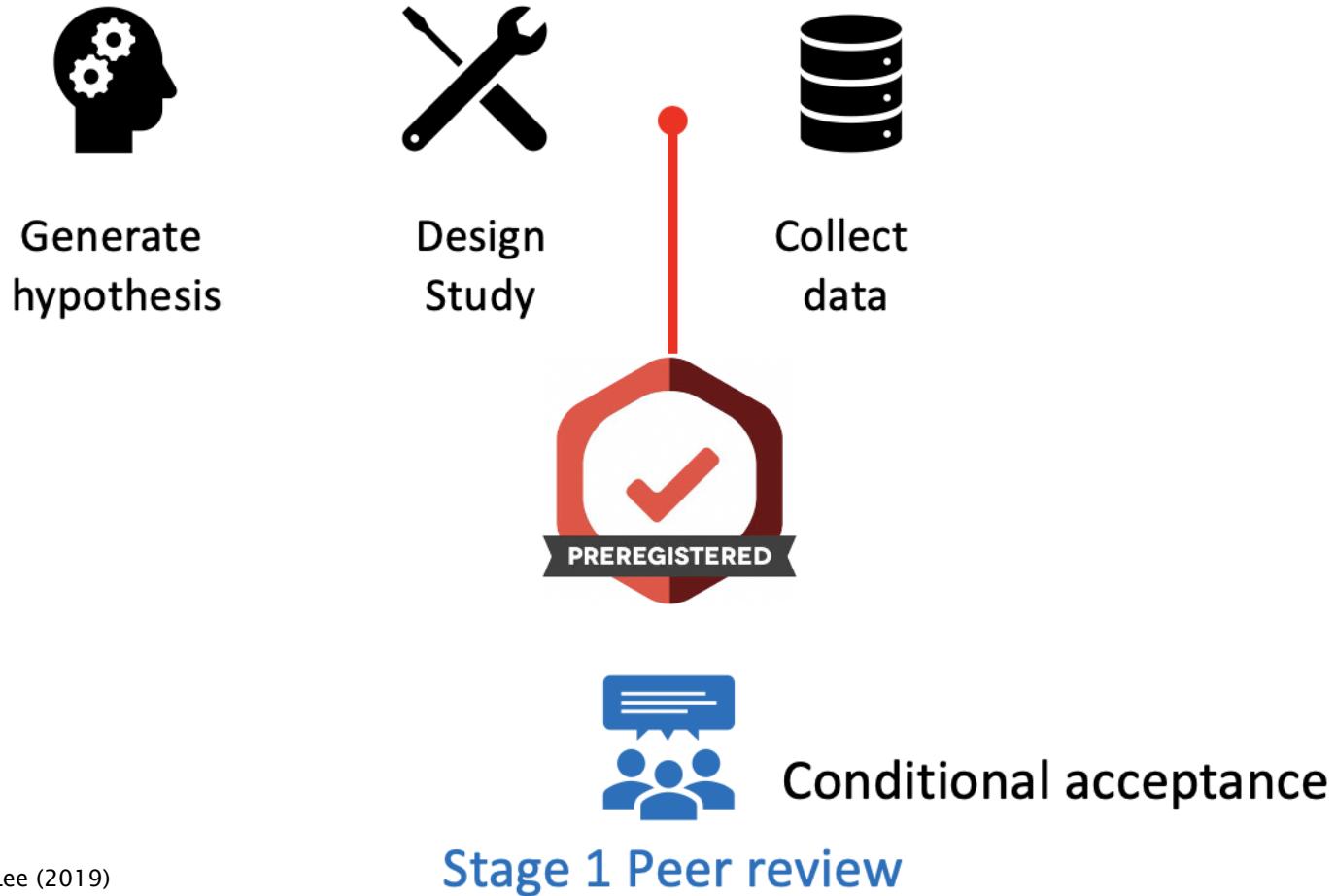
redrawn from Kaplan and Irwin (2015)

What if my results are null?

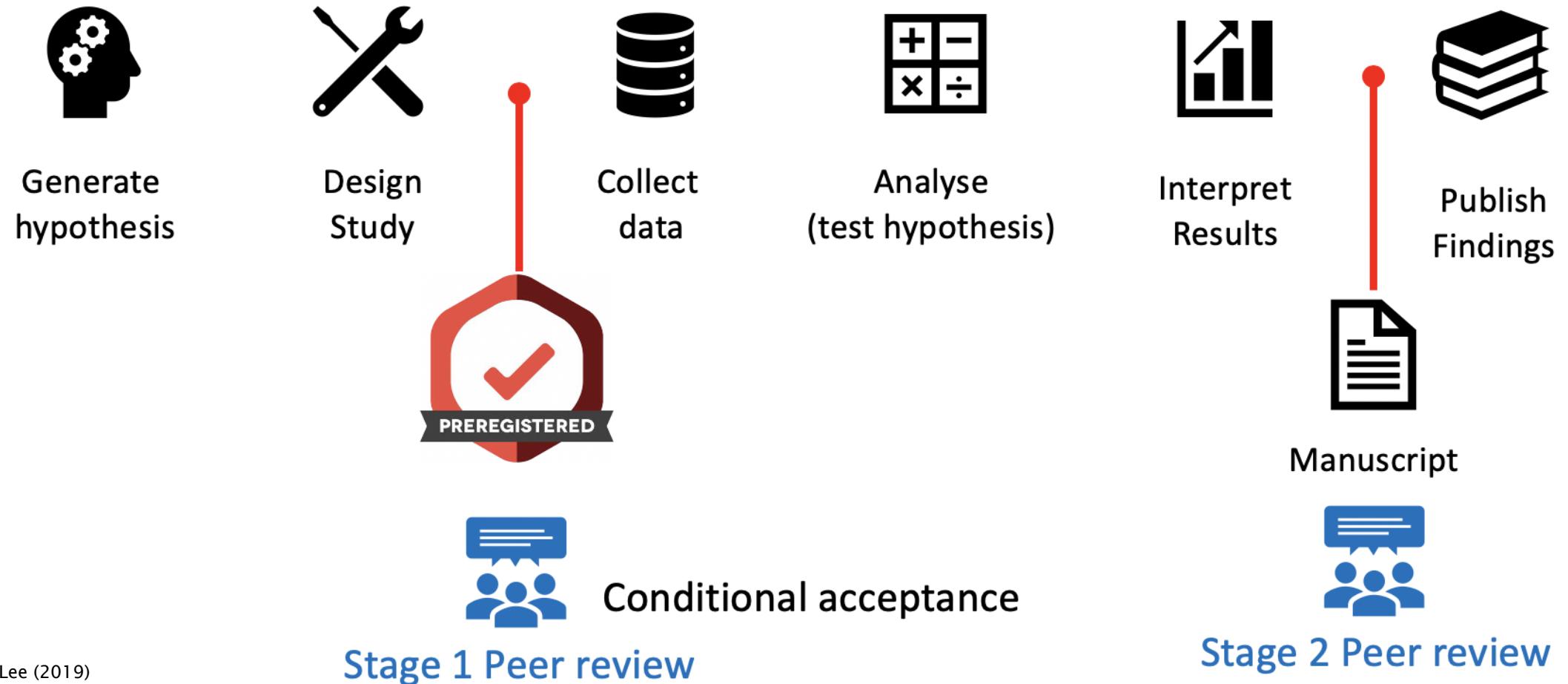
You showed us that they won't get published!

I have to make rent, you know.

Emphasis on design: Registered Reports



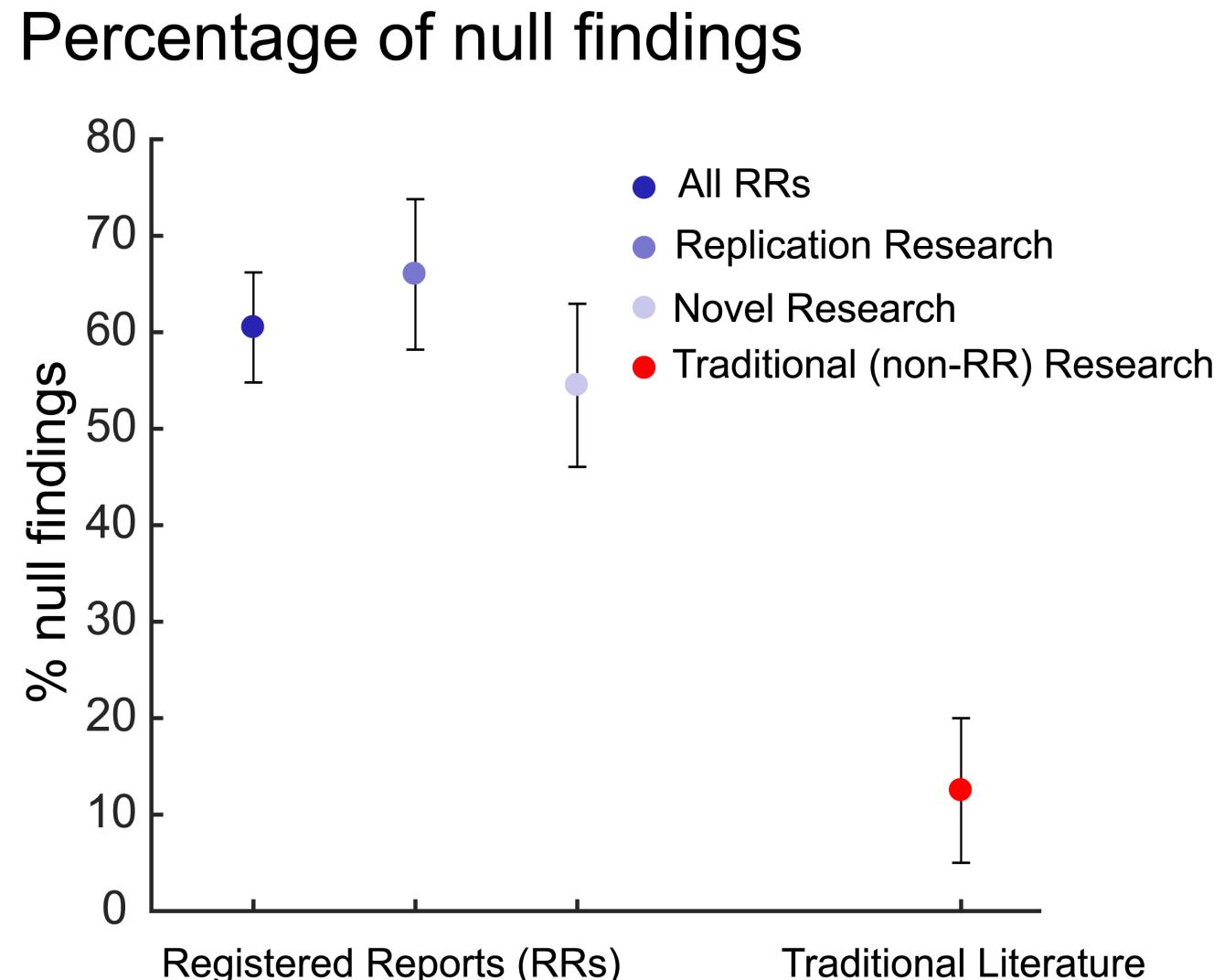
Emphasis on design: Registered Reports



RRs in Psychology

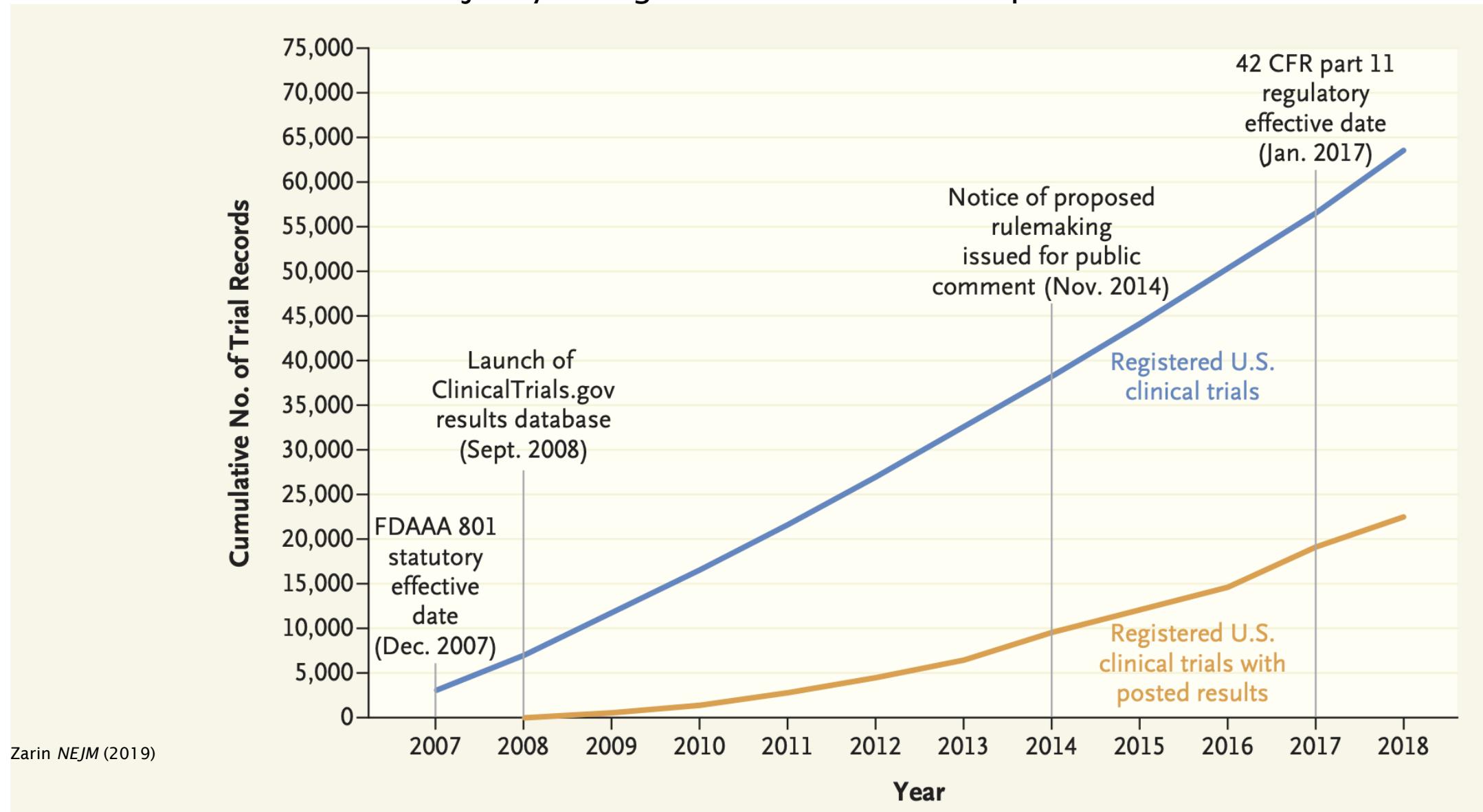
Little difference between 'replication' studies and 'novel' studies.

Big difference from non-registered studies.



Registration is useful
but not sufficient

A majority of registered RCTs still not reported.



But is preregistration enough?

- Still many differences between registration and published reports.

RESEARCH

Open Access



COMPare: a prospective cohort study correcting and monitoring 58 misreported trials in real time

Ben Goldacre^{1*} , Henry Drysdale¹, Aaron Dale¹, Iloan Milosevic¹, Eirion Slade¹, Philip Hartley¹, Cicely Marston², Anna Powell-Smith¹, Carl Heneghan¹ and Kamal R. Mahtani¹

Methods

We set out to prospectively identify all trials published in five leading medical journals over a six-week period, identify every correctly and incorrectly reported outcome in every trial by comparing the published report against the published pre-trial protocol (or, where this was unavailable, the pre-trial registry entry), write a correction letter to the journal for publication on all misreported trials, and document the responses from journals.¹ We used mixed methods combining quantita-

Academic journals are not helping

Summary statistics on correction letter publication

	<i>Annals</i>	BMJ	<i>JAMA</i>	<i>Lancet</i>	<i>NEJM</i>	Total
Letters required	5	2	11	20	20	58
Percentage of letters required	100.00%	66.70%	84.60%	83.30%	90.90%	86.6% (95% CI 78.4–94.7%)
Letters published	5	2	0	16	0	23
Percentage of letters published	100%	100%	0%	80%	0%	39.7% (95% CI 27.0%–53.4%)
Mean publication delay for published letters	0 days (online)	0 days (online)	n/a	150 days	n/a	104 days (median 99 days, range 0–257 days)

Abbreviations: *BMJ* British Medical Journal, *CI* confidence interval, *CONSORT* Consolidated Standards of Reporting Trials, *JAMA* Journal of the American Medical Association, *n/a* not applicable, *NEJM* New England Journal of Medicine

Preregistration is not a panacea

Preregistered \neq correct/sensible/useful

Transparency helps, but cannot fix terrible design or methods.

Post-hoc analysis can be worthwhile

Probing surprising results or mechanisms generates knowledge.

May also lead to 'halo' effects

Preregistered research deserves equal opportunity interrogation.

2. Design Solutions

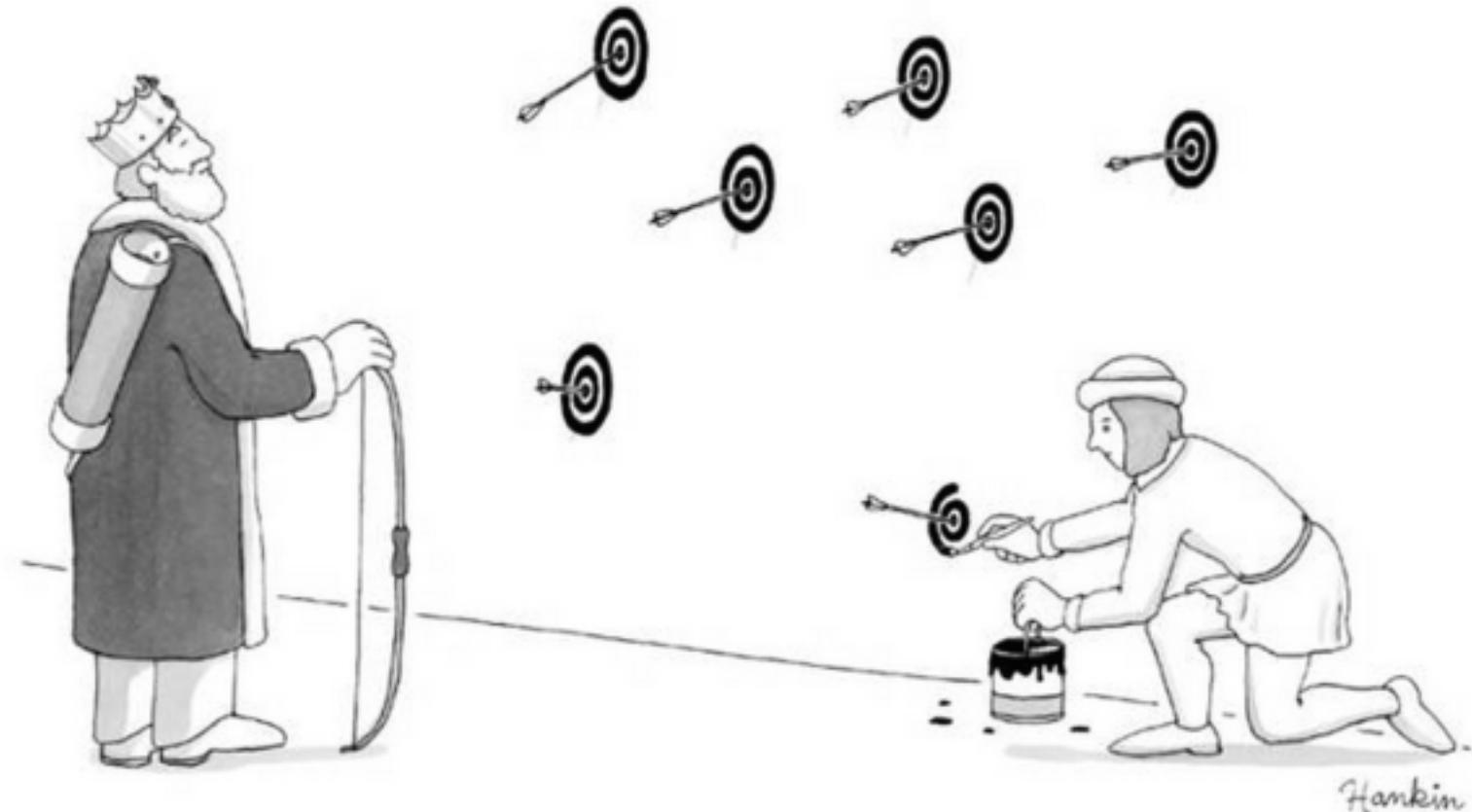
2.1 Preregistration

2.2 Pre-analysis plans

2.3 Reporting guidelines

Hypothesizing After the Results are Known (HARKing)

- Pretending what you found was what you were looking for.
- Easy to "find" theory / biological evidence consistent with results.



What is a pre-analysis plan?



- Detailed description of research design and data analysis plans, submitted to a registry before looking at the data.
- Helps to tie your hands for data analysis (address researcher degrees of freedom, etc.).
- Distinguish between confirmatory and exploratory analysis.
- Increases the credibility of research.
- Transparent methods make it easier for others to build on your work.

Confirmatory and exploratory studies have different aims

Confirmatory

- Well-theorized.
- Plausible mechanisms.
- Minimize false positives.
- Hypothesis *testing*.

Exploratory

- Pushes new ideas.
- Hypothesis *generating*
- Minimize false negatives.
- Testing irrelevant.

What goes into a pre-analysis plan?

- General info (Title, PIs, Staff)
- Introduction and Summary
- Study Design:
 - Hypotheses
 - Main variables
 - Study setting.
 - Intervention components.
 - Data collection methods.
 - Treatment assignment mechanism.
 - Power calculations.
- Analytic decisions
 - models
 - derived variables
 - clustering
 - multiple testing
- Threats/mitigation/robustness checks.
- Dissemination plans

Example from development economics

RESHAPING INSTITUTIONS: EVIDENCE ON AID IMPACTS USING A PREANALYSIS PLAN*

KATHERINE CASEY
RACHEL GLENNERSTER
EDWARD MIGUEL

Despite their importance, there is limited evidence on how institutions can be strengthened. Evaluating the effects of specific reforms is complicated by the lack of exogenous variation in institutions, the difficulty of measuring institutional performance, and the temptation to “cherry pick” estimates from among the large number of indicators required to capture this multifaceted subject. We evaluate one attempt to make local institutions more democratic and egalitarian by imposing participation requirements for marginalized groups (including women) and test for learning-by-doing effects. We exploit the random assignment of a governance program in Sierra Leone, develop innovative real-world outcome measures, and use a preanalysis plan (PAP) to bind our hands against data mining. The intervention studied is a “community-driven development” program, which has become a popular strategy for foreign aid donors. We find positive short-run effects on local public goods and economic outcomes, but no evidence for sustained impacts on collective action, decision making, or the involvement of marginalized groups, suggesting that the intervention

Conclusions:

Turning to empirical methods, this paper underscores the importance of PAPs to limit data mining and generate appropriately sized statistical tests, and discusses some of the practical trade-offs we faced in implementation. We confront the fundamental tension between researcher discretion versus commitment and argue that flexibility to explore questions that arise as the research and project unfold is sometimes desirable yet should only be exercised in tandem with complete transparency over deviations from the *ex ante* specifications. In the context of a PAP, limited flexibility with full transparency allows the scholarly community to make its own assessments about the credibility of different results. We show how misleading an undisciplined interpretation of treatment effects can be in the absence of a PAP by constructing two opposing and equally erroneous narratives based on our data.

Example from epidemiology

Note the time-stamp, which provides credible evidence of *when* you had your brilliant ideas.

Pre-analysis plan_2020-Jan-27_FINAL.pdf (Version: 1)

Check out Delete Download View Revisions

The screenshot shows a file listing for "Pre-analysis plan_2020-Jan-27_FINAL.pdf". On the right, there are buttons for "Check out", "Delete", "Download", "View", and "Revisions". Below the file name, the version number "Version: 1" is shown. On the left, there's a sidebar with "Writing" and "OSF Storage (Canada - Montréal)" sections, and the file itself is listed. The main area is titled "Revisions" and contains a table with columns: Version ID, Date, User, Download, MD5, and SHA2. The first row shows Version ID 1, Date 2020-01-30 09:13 AM, User Sam Harper, 0 downloads, MD5 c1f3c508af41b1eb69d6, and SHA2 b67f1cc672dda474c3b. The "Date" column is highlighted with a red box.

Version ID	Date	User	Download	MD5	SHA2
1	2020-01-30 09:13 AM	Sam Harper	0		c1f3c508af41b1eb69d6 b67f1cc672dda474c3b

Example from epidemiology

Can be challenging for observational studies or secondary data analyses.

Can you prove when you obtained data access?

Pre-analysis plan for “Short term benefits but long term harm? Assessing the consequences of antenatal corticosteroid administration for child neurodevelopment”

Jennifer A Hutcheon¹, Sam Harper², Amanda Skoll¹, Myriam Srour³, Jessica Liauw¹, Erin Strumpf^{2,4}

¹Department of Obstetrics & Gynaecology, University of British Columbia

² Department of Epidemiology, Biostatistics, and Occupational Health, McGill University

³ Department of Pediatric Neurology, McGill University

⁴ Department of Economics, McGill University

Purpose

This document describes a pre-analysis plan for a study examining the child health consequences of antenatal corticosteroid administration in a population-based cohort of linked administrative and clinical records from British Columbia, Canada. We use a regression discontinuity design that exploits the pronounced change in antenatal corticosteroid administration practices based on a clinical practice guideline that recommended administration up to 33 weeks, 6 days of gestation (33+6 weeks), but not at or beyond 34+0 weeks. This pre-analysis plan was written after the individual datasets had been received and some descriptive statistics calculated for key variables, but prior to the linkage of the datasets or analyses linking exposure with longer-term child health outcomes.

2. Design Solutions

2.1 Preregistration

2.2 Pre-analysis plans

2.3 Reporting guidelines

"Most publications have elements that are missing, poorly reported, or ambiguous"

Reducing waste from incomplete or unusable reports of biomedical research

Paul Glasziou, Douglas G Altman, Patrick Bossuyt, Isabelle Boutron, Mike Clarke, Steven Julious, Susan Michie, David Moher, Elizabeth Wager

Abstract Trials: missing effect size and confidence interval (38%); no mention of adverse effects (49%) ⁷²
Methods Trials: 40–89% inadequate treatment descriptions ^{11,13} fMRI studies: 33% missing number of trials and durations ³ Survey questions: 65% missing survey or core questions ²⁵ Figures: 31% graphs ambiguous ⁴⁵
Results Clinical trials: outcomes missing: 50% efficacy and 65% harm outcomes per trial incompletely reported ⁶ Animal studies: number of animals and raw data missing ¹⁷ (54%, 92%); age and weight missing (24%) Diagnostic studies: missing age and sex (40%) ¹⁵
Discussion Trials: no systematic attempt to set new results in context of previous trials (50%) ⁶⁹
Data Trials: most data never made available; author-held data lost at about 7% per year

Figure 3: Estimates of the prevalence of some reporting problems (see publication column, figure 1).
fMRI=functional MRI.

Importance of intervention details

Want decision-makers to act on your evidence?

Can they actually understand what you did?

Poor description of non-pharmacological interventions: analysis of consecutive sample of randomised trials



OPEN ACCESS

Tammy C Hoffmann *associate professor of clinical epidemiology*, Chrissy Erueti *assistant professor*, Paul P Glasziou *professor of evidence-based medicine*

- Of 137 interventions, only 53 (39%) were adequately described;
- The most frequently missing item was the “intervention materials” (47% complete);

Missing due to:

- copyright or intellectual property;
- absent materials or intervention details;
- unaware of their importance.

Reasons given for study intervention materials being unavailable

Category of reason (number of authors providing a response in this category) and illustrative quotes from authors:

Materials not publicly available (9)

"Due to legal copyright restrictions at my university I am unable to send"
"Not publicly available because we based them on materials provided by our local government"
"Not publicly available—only to our trainers"
"Not yet—they will be made publicly available within two years"
"No it is not. Attached is a table of contents"
"The training materials from the trial are not online—we had no real reason to do that"

Corresponding author did not have copy of materials to send or could not provide further details about intervention (8)

"People originally in the position have moved on"
"I am unable to find . . . my old computer files"
"I'm afraid I no longer have access to those materials"
"I do not have it"
"I am not able to answer most of your questions. I was not involved with running the trial, only analysing and reporting on the QOL results after the data was collected"
"I can't provide these"

Other (3)

"You will have to read the literature"
"No, is in Dutch"
"The [materials] are tailored, thus it is difficult to disseminate. We could send an example"

Materials were previously publicly available but no longer are (2)

"URL doesn't exist anymore"
"We had been making it previously available, but need to update it, so are no longer"

Reporting guidelines exist for entire research lifecycle

Question and approach

Systematic review

👉 PRISMA/PROSPERO

Pre-intervention

Research protocol/preanalysis

👉 SPIRIT

Research report

Trials/Observational studies

👉 CONSORT/STROBE

Cost-effectiveness

Benefits and costs of interventions

👉 CHEERS



Your one-stop-shop for writing and publishing high-impact health research

find reporting guidelines | improve your writing | join our courses | run your own training course | enhance your peer review | implement guidelines



Library for health research reporting

The Library contains a comprehensive searchable database of reporting guidelines and also links to other resources relevant to research reporting.

-  [Search for reporting guidelines](#)
-  [Not sure which reporting guideline to use?](#)
-  [Reporting guidelines under development](#)
-  [Visit the library for more resources](#)



Reporting guidelines for main study types

Randomised trials	CONSORT	Extensions
Observational studies	STROBE	Extensions
Systematic reviews	PRISMA	Extensions
Study protocols	SPIRIT	PRISMA-P
Diagnostic/prognostic studies	STARD	TRIPOD
Case reports	CARE	Extensions
Clinical practice guidelines	AGREE	RIGHT
Qualitative research	SRQR	COREQ
Animal pre-clinical studies	ARRIVE	
Quality improvement studies	SQUIRE	
Economic evaluations	CHEERS	

[See all 442 reporting guidelines](#)

How to describe the placebo used in a trial?
Damiao Alves, Unsplash

Use the **TIDieR-Placebo** reporting guideline!

● ● ● ●

(Some) evidence that it might matter.

- Some evidence that item reporting has increased.
- Consistent with revised CONSORT (2001).
- Non-adopting journals report fewer items.

The quality of reports of randomised trials in 2000 and 2006: comparative study of articles indexed in PubMed

Sally Hopewell, senior research fellow,¹ Susan Dutton, senior medical statistician,¹ Ly-Mee Yu, senior medical statistician,¹ An-Wen Chan, assistant professor,² Douglas G Altman, director¹

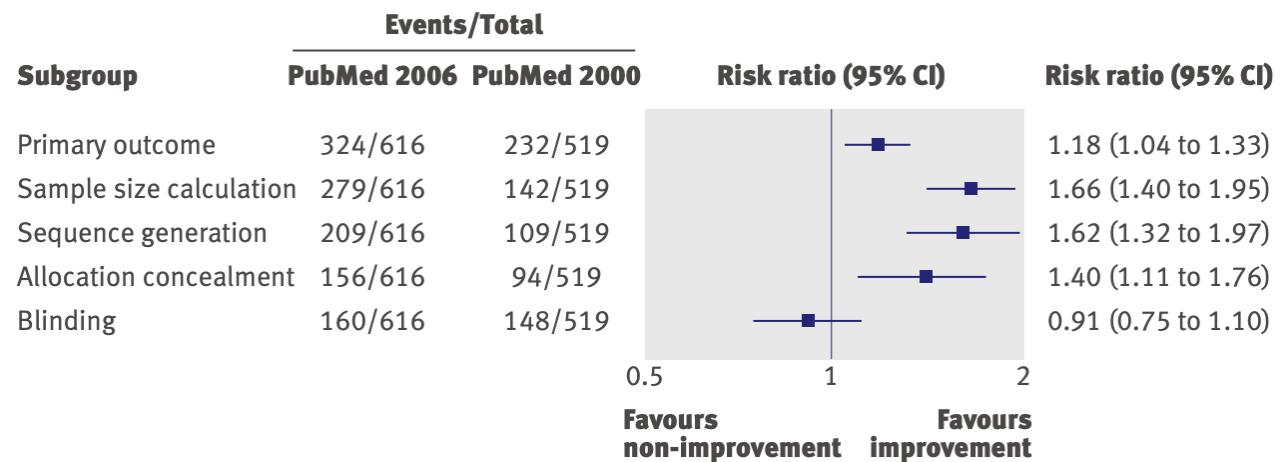


Fig 2 | Differences in reporting of methodological items between 2000 and 2006

Since 2015 funders, journals are embracing *Transparency and Openness* (TOP) guidelines.

8 MODULAR STANDARDS

Citation Standards Describes citation of data	Data Transparency Describes availability and sharing of data
Analytical Methods Transparency Describes analytical code accessibility	Research Materials Transparency Describes research materials accessibility
Design and Analysis Transparency Sets standards for research design disclosures	Preregistration of Studies Specification of study details before data collection
Preregistration of Analysis Plans Specification of analytical details before data collection	Replication Encourages publication of replication studies

ACROSS 3 TIERS

1 DISCLOSURE:
the final research output
must disclose if the work
satisfies the standard

2 REQUIREMENT:
the final research output
must satisfy the standard

3 VERIFICATION:
third party must verify that
the standard is being met

It's still difficult to change norms

Most journals chose *Level 1* (disclosure)

J Am Heart Assoc published 40 original research papers during first half of 2019.

- Posted data: 0
- Posted code: 1
- Data upon "reasonable" request: 30
- Code upon "reasonable" request: 5

MINI-REVIEW

Resource Sharing to Improve Research Quality

Ghassan B. Hamra, PhD; Neal D. Goldstein, PhD; Sam Harper, PhD

Transparency and openness are vital for strengthening the scientific process. However, there is no clear agreement in the scientific community about the elements necessary to qualify scientific research as a transparent and open process. Historically, the description of study methods and results within individual academic publications has been treated as sufficient for establishing transparency; that is, based solely on the written description of study procedures and analytic techniques, a third party can be *assumed* to have all the information needed to reproduce the results of an individual study if the data were available. The core philosophy of *reproducible* research is slightly different and challenges this assumption. Rather than relying on the written report, reproducible research culture demands access to data and analytic code used to produce study results. In this scenario, anyone should be able to exactly reproduce the tables, figures, and evidence presented in a given article. The push for reproducible research and current publication practices do question those findings. While self-correction is natural in science, it is not the norm,¹ and reports have suggested that the extent to which study findings cannot be replicated is alarming, leading to the so-called replication crisis.² Many related reasons have been put forward to explain the replication crisis, including misaligned incentives in academia, the file drawer effect,³ p-hacking,⁴ overreliance on null hypothesis significance testing,⁵ and even outright falsification of data. Some have suggested that our existing assumptions about what qualifies as transparent and open in science may be insufficient and that addressing this can safeguard against further replication crises.

In this commentary, we discuss the importance of transparency and openness, focusing on the 2 major elements necessary for reproducibility: the data and analytic code used to produce the results in a published research report. We highlight how greater openness can support more reliable findings (in the long run) by allowing checks for

Value of reporting guidelines



Improve transparency of reported research

Benefits funders, producers and consumers of research.

May help to improve the quality of research.

More evidence needed, unintended consequences possible.

Better reporting \neq more reliable.

Transparently reported research can still be biased/bad.

Registration, pre-analysis plans, and reporting guides are design strategies to help mitigate bias from underreported research

They do not guarantee reliable or valid research

Break! 

10:00