Ten simple rules for good research practice

Leonhard Held, Center for Reproducible Science, University of Zurich

University of Zurich[UZH]

QUEST Seminar on Responsible Research

21 June 2022

The UZH Center for Reproducible Science

Network

- 4 **faculties**, 34 **members** and 15 **fellows**
- **Swiss Reproducibility Network**

RN CH Swiss Reproducibility Network

Training

- **Regular Good Research Practice Courses**
- DISK4U: **Digital skills for Open Science** for students & lecturers

Research

- Design and analysis of **Replication Studies**
- **Meta-Research**
- ReproducibiliTea

Outreach

- Scientifica, PRECHECK
- Reproducibility Notes

Ten simple rules for good research practice

PLOS COMPUTATIONAL BIOLOGY

Planning

1. Specify your research question
2. Write and register a study protocol
3. Justify your sample size
4. Write a data management plan
5. Reduce bias

Execution

6. Avoid questionable reseach practices
7. Be cautious with interpretations of statistical significance
8. Make your research open

Reporting

9. Report all findings
10. Follow reporting guidelines

EDITORIAL

Ten simple rules for good research practice

Simon Schwab[1,2]*, Perrine Janiaud[3], Michael Dayan[4], Valentin Amrhein[5], Radoslaw Panczak[6], Patricia M. Palagi[7], Lars G. Hemkens[3,8,9], Meike Ramon[10], Nicolas Rothen[11], Stephen Senn[12], Eva Furrer[1,2], Leonhard Held[1,2]

1 Center for Reproducible Science, University of Zurich, Zurich, Switzerland, **2** Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Zurich, Switzerland, **3** Department of Clinical Research, University Hospital Basel, University of Basel, Basel, Switzerland, **4** Human Neuroscience Platform, Fondation Campus Biotech Geneva, Geneva, Switzerland, **5** Department of Environmental Sciences, Zoology, University of Basel, Basel, Switzerland, **6** Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland, **7** SIB Training Group, SIB Swiss Institute of Bioinformatics, Lausanne, Switzerland, **8** Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, California, United States of America, **9** Meta-Research Innovation Center Berlin (METRIC-B), Berlin Institute of Health, Berlin, Germany, **10** Applied Face Cognition Lab, University of Lausanne, Lausanne, Switzerland, **11** Faculty of Psychology, UniDistance Suisse, Brig, Switzerland, **12** Statistical Consultant, Edinburgh, United Kingdom

* schw4b@gmail.com

osf.io/am5ck/

Rule 1: Specify your research question

- A successful study requires a narrow and clear **research question**.
- **P**opulation, **I**ntervention, **C**omparator, **O**utcome, **T**ime frame: **PICOT** guidelines

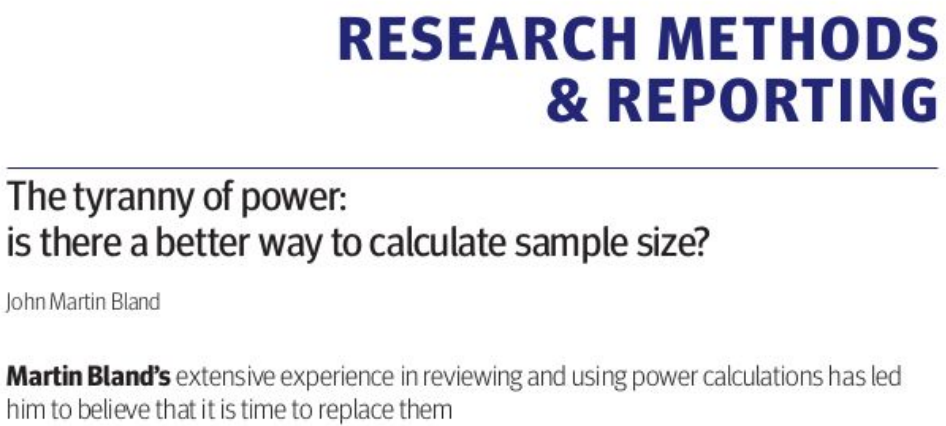
Rule 2: Write and register a study protocol

- **Protocol** specifying research question and hypotheses, describing population, sample size, inclusion/exclusion criteria, study design, planned statistical analyses.
- **Registration** reduces bias
- **Registered reports**

Chambers (2019, Nature)

Rule 3: Justify your sample size

- A **sample size** that is too low might
  - increase the risk of finding false negative results
  - overestimate the effect size
- Appropriate **sample size calculation**
  - ensures sufficient statistical power
  - or a small enough width of confidence interval



RESEARCH METHODS & REPORTING

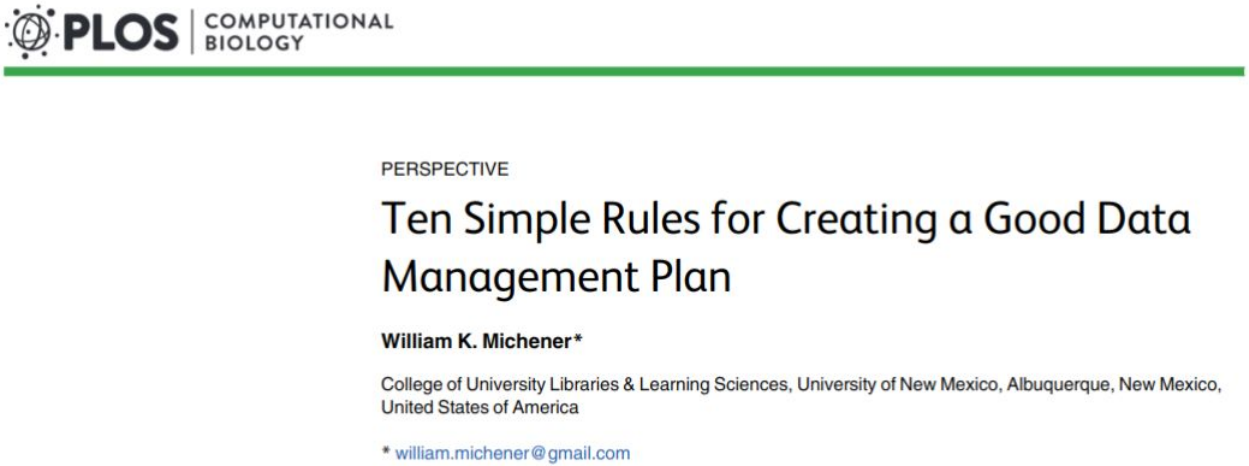
The tyranny of power:
is there a better way to calculate sample size?

John Martin Bland

Martin Bland's extensive experience in reviewing and using power calculations has led him to believe that it is time to replace them

Rule 4: Write a data management plan (DMP)

- **Data** is (recognized as) a key research output.
- DMPs required by funders when applying for grants.
- DMP describes what data is collected, and how it will be organized, stored, protected and shared
- Data should be **F**indable, **A**ccessible, **I**nteroperable and **R**eusable.



Rule 5: Reduce bias

Many different forms of **bias**, that can occur at different stages of research.

https://catalogofbias.org/biases/

Name	Explanation	Prevention
Allocation bias	Systematic difference in the assignment of participants to the treatment and control group in a clinical trial. For example, the investigator knows or can predict which intervention the next eligible patient is supposed to receive due to poorly concealed randomization.	- Randomization with allocation concealment
Attrition bias	Attrition occurs when participants leave during a study that aims to explore the effect of continuous exposure (drop-outs, withdrawal). For example, more drop-outs of patients randomized to an aggressive cancer treatment.	- Good investigator-patient communication - Accessibility of clinics - Incentives to continue
Confounding bias	An artificial association between an exposure and an outcome because another variable is related to both the exposure and outcome. For example, lung cancer risk in coffee drinkers is evaluated, ignoring smoking status (smoking is associated with both, coffee drinking and cancer). A challenge is that many confounders are unknown and/or not measured.	- Randomization (can address unmeasured confounders) - When randomization is not possible: - Restriction to one level of the confounder - Matching on the levels of the confounder - Stratification and analysis within strata - Propensity score matching
Immortal time bias	Survival beyond a certain time point is necessary in order to be exposed (participants are "immortal" in that time period). For example, discharged patients are analyzed but were included in the treatment group only if they filled a prescription for a drug 90 days after discharge from hospital.	- Group assignment at time-zero - Time-dependent analysis may be used
Information bias	Bias that arises from systematic differences in the collection, recall, recording or handling of information. For example, blood pressure in the treatment arm is measured in the morning, and for the control arm in the evening.	- Standardized data collection - Data collection independent from exposure or outcome (e.g. by blinding of intervention status/exposure) - Use of objective measurements
Publication bias	Occurs when only studies with a positive or negative result are published. Affects meta-analyses from systematic reviews and harms evidence-based medicine.	- Writing a study protocol and preregistration - Publishing study protocol or registered report - Report all findings including negative findings

Example: Immortal time bias

30.3 Immortal time

In my opening lecture to a class designed primarily for second year doctoral students in epidemiology, I state the First Rule of survival analysis: *Selection into the study cohort, or into subgroups to be compared in the analysis, must not depend on events that occur after the start of follow-up.* While this point may be obvious to a statistician, certainly one trained to use martingale arguments to justify inferences about how past history influences rates of future events, it was not obvious to many of the epidemiologists. The "immortal time" bias that results from failure to follow the rule has resulted, and continues regularly to result, in grossly fraudulent claims in papers published in the most prestigious medical journals.

Breslow (2014)

VOLUME 26 · NUMBER 24 · AUGUST 20 2008

JOURNAL OF CLINICAL ONCOLOGY | CELEBRATING 25 YEARS OF JCO

Analysis of Survival by Tumor Response and Other Comparisons of Time-to-Event by Outcome Variables

James R. Anderson, *University of Nebraska College of Public Health, Omaha, NE*
Kevin C. Cain, *University of Washington, Seattle, WA*
Richard D. Gelber, *Dana-Farber Cancer Institute and Harvard School of Public Health, Boston, MA*

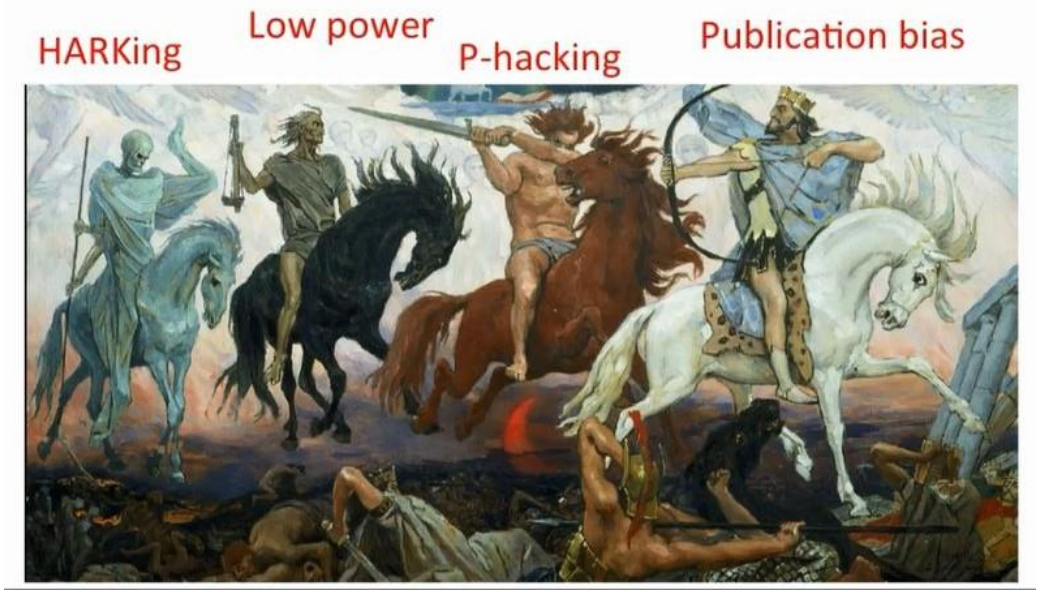
Annals of Internal Medicine | ACADEMIA AND CLINIC

Do Oscar Winners Live Longer than Less Successful Peers? A Reanalysis of the Evidence

Marie-Pierre Sylvestre, MSc; Ella Huszti, MSc; and James A. Hanley, PhD

Rule 6: Avoid questionable research practices (QRPs)

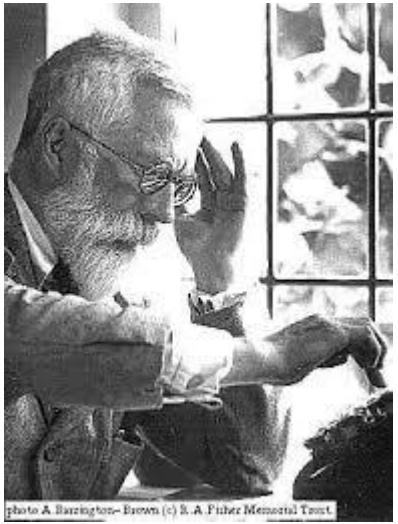
- Many forms of **QRPs**: Low statistical power, p-hacking, selective reporting, HARKing, …
- Can be avoided with **proper planning** of studies or **preregistration**.



Bishop (2019, Nature)

Rule 7: Be cautious with interpretations of statistical significance

- Statistical significance vs. clinical relevance
- Rarely the goal is decision making → report exact p-value



Replication power

Replication power and regression to the mean

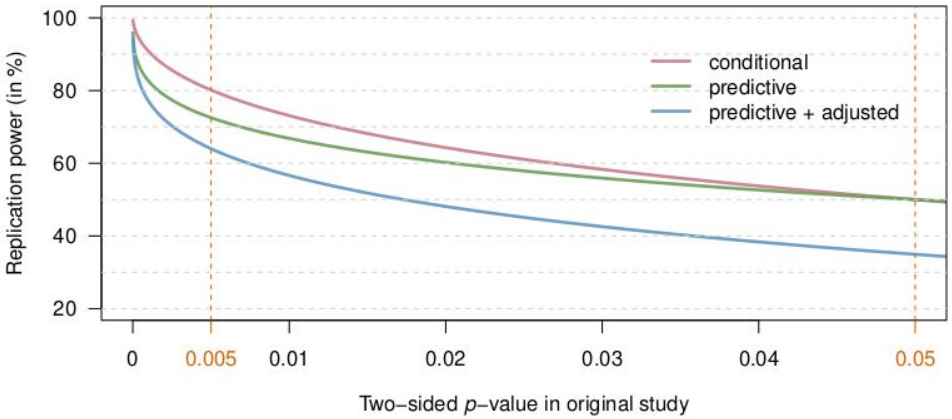
If a scientific study reports a discovery with a *p*-value at or around 0.05, how credible is it? And what are the chances that a replication of this study will produce a similarly "significant" finding? **Leonhard Held**, **Samuel Pawel** and **Simon Schwab**'s answers may surprise you

p-values and significance tests defined

The *p*-value "is defined as the probability, under the assumption of no effect or no difference (the null hypothesis), of obtaining a result equal to or more extreme than what was actually observed".[3] If the *p*-value is smaller than some pre-defined *significance level* (usually 0.05), the result is said to be *statistically significant*. The probability to obtain a statistically significant result is called the *power* of the test, which also depends on the true effect size and the sample size.

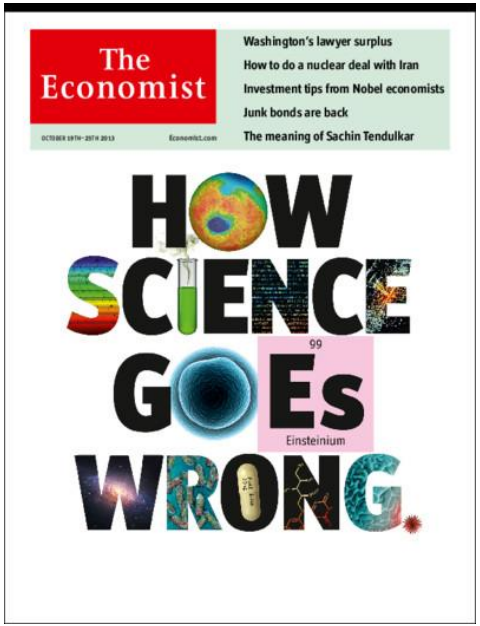
Fisher used "significance" merely to indicate that an observation was **worth following up**, with refutation of the null hypothesis justified only if **further experiments** "rarely failed" to achieve significance

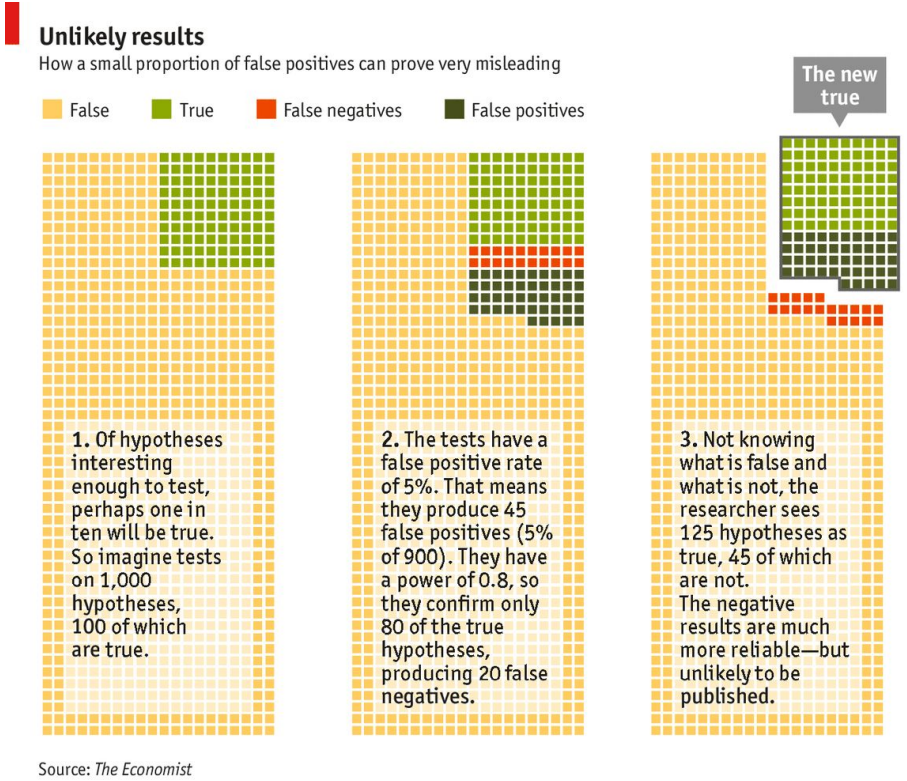
Goodman (2016, Science)



Held et al, Significance 2020

False positives

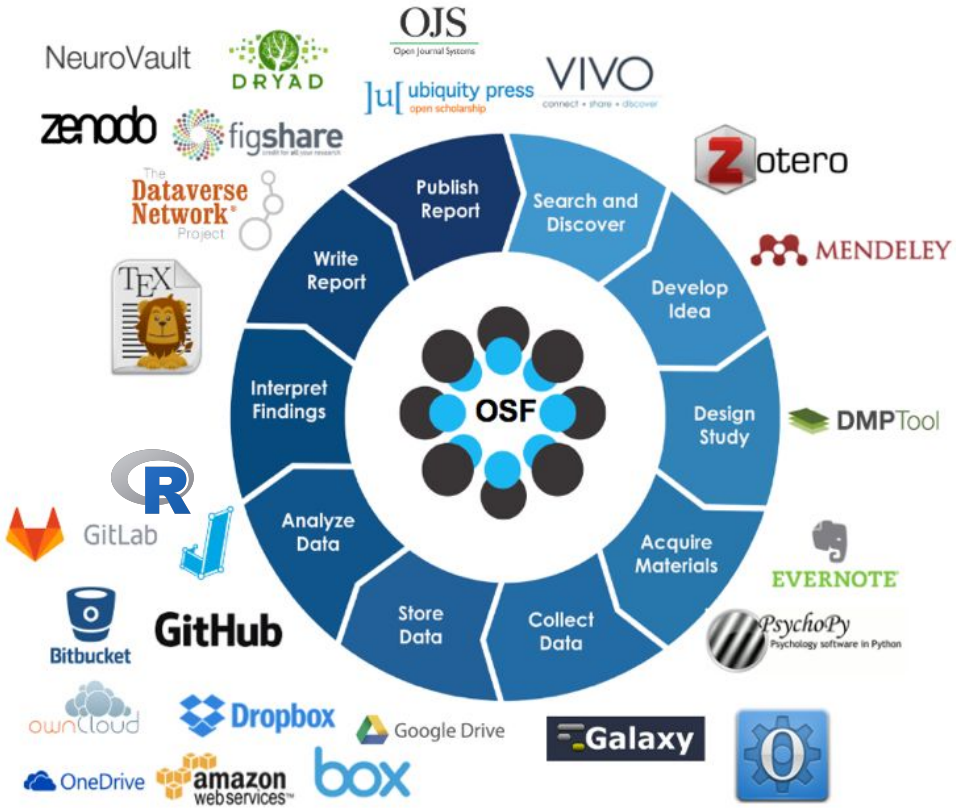




Source: *The Economist*

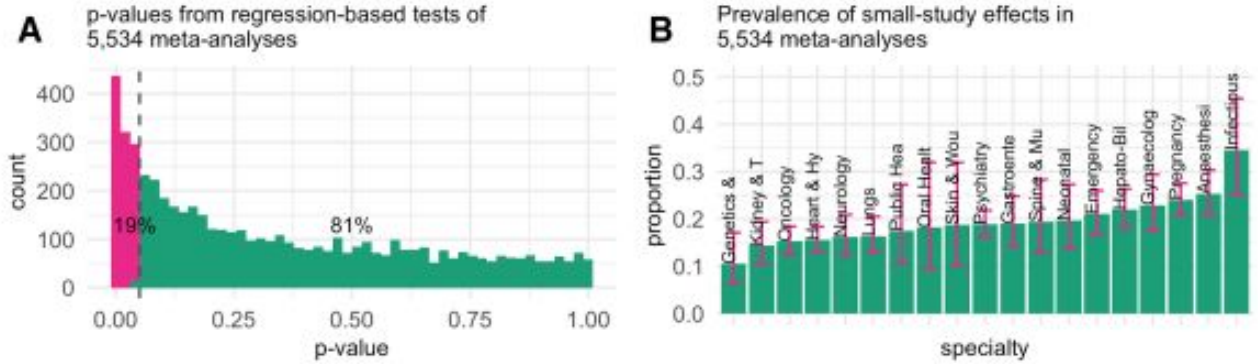
Rule 8: Make your research open

- To foster **transparency** and **accessibility**.
- Research paper can link to data and analysis code to facilitate (and encourage) reproducibility.
- Increased visibility of datasets can help career building.



Rule 9: Report all findings

- Avoid **publication** and **outcome reporting bias**
- Nonsignificant “negative” findings are worth to be reported!



Open access Original research

BMJ Open **Assessing treatment effects and publication bias across different specialties in medicine: a meta-epidemiological study**

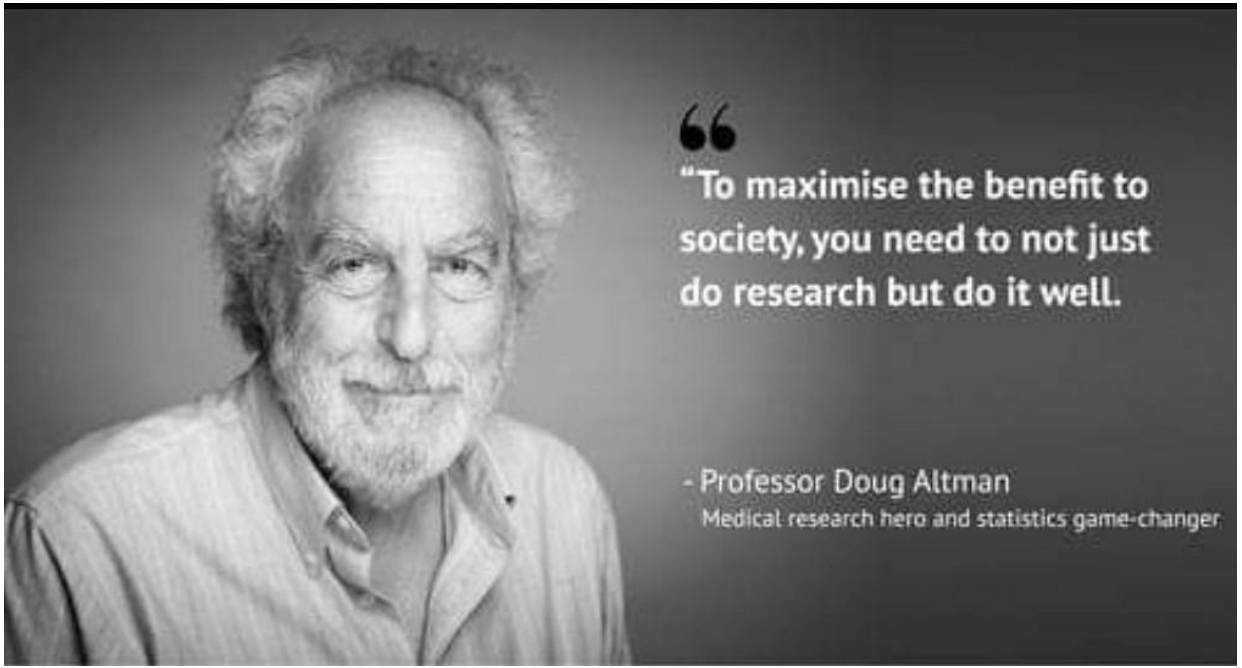
Simon Schwab ,[1] Giuachin Kreiliger,[2] Leonhard Held[1]

Rule 10: Follow reporting guidelines

Reporting guidelines provide the minimum information needed to ensure that scientific findings can be used and studies replicated.

Guideline name	Study type
ARRIVE	Animal experiments
CONSORT	Randomized trials
STROBE	Observational studies
PRISMA	Systematic reviews
SPIRIT	Study protocols
STARD/TRIPOID	Diagnostic/prognostic studies

Summary


"To maximise the benefit to society, you need to not just do research but do it well.
- Professor Doug Altman
Medical research hero and statistics game-changer