

Computational Reproducibility Seminar

Ten simple rules for good research practice

16. May 2024

Simon Schwab



Photo by Xavi Cabrera on Unsplash



Organ donation and transplantation in Switzerland

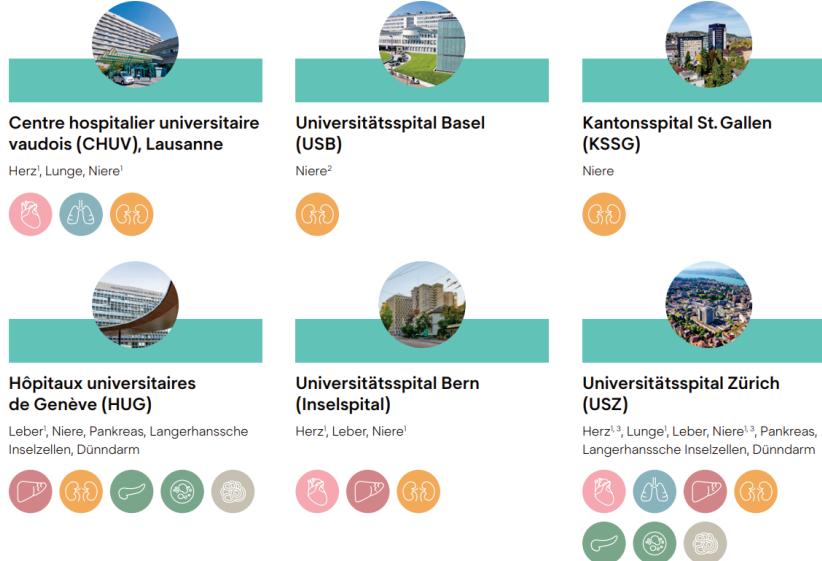


Swisstransplant

- Swiss National Foundation for Organ Donation and Transplantation.
- Manages the waiting list on behalf of the Federal Office of Public Health (FOPH) and is responsible for allocating organs in accordance with the law.



Organ transplantation in Switzerland



¹ Transplantationen auch bei Kindern

² Transplantation bei älteren Kindern von Nieren-Lebendspenderinnen und -spendern

³ In Zusammenarbeit mit dem Kinderspital Zürich



Main sources of data

- SOAS: Swiss organ allocation system
- STCS: Swiss Transplant Cohort Study



Acknowledgments

Many thanks for the invitation

This work was partly supported by the SfwF "Stiftung für wissenschaftliche Forschung" (STWF-19-007)



<https://www.crs.uzh.ch/>



Ten Simple Rules paper

PLOS COMPUTATIONAL BIOLOGY

EDITORIAL

Ten simple rules for good research practice

Simon Schwab^{1,2*}, Perrine Janiaud³, Michael Dayan⁴, Valentin Amrhein⁵, Radoslaw Panczak⁶, Patricia M. Palagi⁷, Lars G. Hemkens^{8,9}, Meike Ramon¹⁰, Nicolas Rothen¹¹, Stephen Senn¹², 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

This is a *PLOS Computational Biology* Methods paper.

Rules are rules!



Ten Simple Rules paper

- Based on our experiences from the Good Research Practice Course (GRP) at UZH
<https://www.crs.uzh.ch/en/training/GoodResearchPractice.html>
- Can serve as a primer for young researcher
- Starting point for others who want to do a similar course
- Focus on clinical research, maybe even biomedical research
- Many things are not covered (misconduct, conflicts of interests, scientific integrity)
- 101 References, among them:

OPEN  ACCESS Freely available online

 PLOS COMPUTATIONAL BIOLOGY

Editorial

Ten Simple Rules for Reproducible Computational Research

Geir Kjetil Sandve^{1,2*}, Anton Nekrutenko³, James Taylor⁴, Eivind Hovig^{1,5,6}

Overview



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 research practices
7. Be cautious with interpretations of statistical significance
8. Make your research open

Reporting

9. Report all findings
10. Follow reporting guidelines

Fig 1. The 10 simple rules for GRP grouped into planning, execution, and reporting of research. GRP, good research practices.



1. Specify your research question

Statistician: "What is your research question?"

Clinician: "The deadline is due tomorrow."

- Ask your PI or the literature for opinion papers (both are just opinions)
- Better: Systematic assessment of research gaps, patient involvement, and public involvement

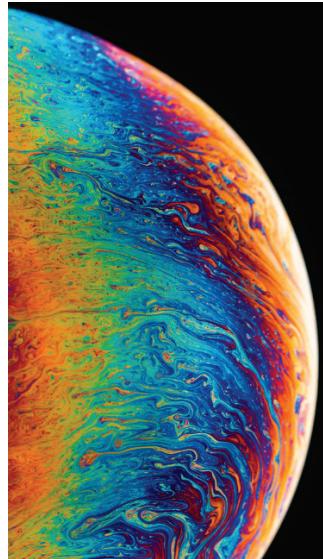
A National Survey Comparing Patients' and Transplant Professionals' Research Priorities in the Swiss Transplant Cohort Study

Sonja Beckmann^{1,2}, Oliver Mauthner^{1,3}, Liz Schick⁴, Jessica Rochat⁵, Christian Loris^{5,6}, Annette Boehler⁷, Isabelle Binet⁸, Uyen Huynh-Do⁹, Sabina De Geest^{1,10*}, the Psychosocial Interest Group[†], and the Swiss Transplant Cohort Study[‡]

¹Department Public Health, Institute of Nursing Science, University of Basel, Basel, Switzerland, ²Center Clinical Nursing Science, University Hospital Zurich, Zurich, Switzerland, ³University Department of Geriatric Medicine Felix Platter, Basel, Switzerland, ⁴Swisstransplant, Bern, Switzerland, ⁵Faculty of Medicine, University of Geneva, Geneva, Switzerland, ⁶Division of Medical Information Sciences, University Hospitals of Geneva, Geneva, Switzerland, ⁷University of Zurich, Zürich, Switzerland, ⁸Service of Nephrology and Transplantation Medicine, Cantonal Hospital, St. Gallen, Switzerland, ⁹Department of Nephrology and Hypertension, Inselspital, University of Bern, Bern, Switzerland, ¹⁰Academic Center for Nursing and Midwifery, Department of Public Health and Primary Care, KU Leuven, Leuven, Belgium



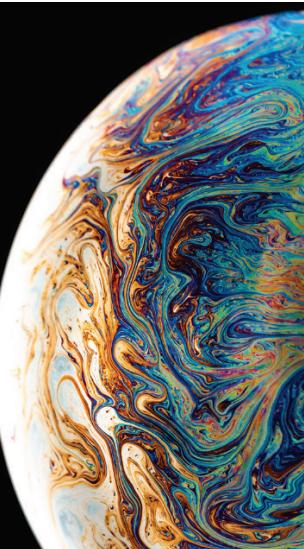
To confirm or to explore?



REPRODUCIBILITY NOTES

Different worlds Confirmatory versus exploratory research

Simon Schwab and Leonhard Held explain the differences between confirmatory and exploratory research and the dangers of confusing the two concepts



Exploratory research	Confirmatory research
No hypothesis required/hypothesis can be vague Generate new hypothesis from data High sensitivity desired, i.e. minimising the risk of false negatives Suitable for making new discoveries and finding the unexpected For example: Testing of new compounds in mice	Clear hypothesis required Test <i>a priori</i> hypothesis with new data High specificity desired, i.e. minimising the risk of false positives Suitable for establishing strong evidence and confirming the expected For example: Assessing the efficacy of a drug in humans

Researchers might present exploratory results as confirmatory to increase the probability of publication.

It is like a game of darts, where the target is drawn around a dart only after it has been thrown.

“A medical statistician will not accept that Columbus discovered America because he said he was looking for India in the trial plan.”

(Stephen Senn)



2. Write and register a study protocol

Schwab *et al.*
Diagnostic and Prognostic Research (2023) 7:6
<https://doi.org/10.1186/s41512-022-00139-5>

Diagnostic and
Prognostic Research

PROTOCOL

Open Access



Clinical prediction model for prognosis in kidney transplant recipients (KIDMO): study protocol

Simon Schwab^{1*} , Daniel Sidler², Fadi Haidar³, Christian Kuhn⁴, Stefan Schaub⁵, Michael Koller⁵, Katell Mellac⁵, Ueli Stürzinger⁶, Bruno Tischhauser⁶, Isabelle Binet⁴, Déla Golshayan⁷, Thomas Müller⁸, Andreas Elmer¹, Nicola Franscini¹, Nathalie Krügel¹, Thomas Fehr⁹, Franz Immer¹, the Swisstransplant Kidney Working Group (STAN) and the Swiss Transplant Cohort Study

- Some researchers are afraid to “get scooped” so they are against (pre)registration



3. Justify your sample size

R packages:

- pwr
- pmsampsize
- Monte Carlo simulation



BMJ 2020;368:m441 doi: 10.1136/bmj.m441 (Published 18 March 2020)

Page 1 of 12



RESEARCH METHODS & REPORTING

Calculating the sample size required for developing a clinical prediction model

Clinical prediction models aim to predict outcomes in individuals, to inform diagnosis or prognosis in healthcare. Hundreds of prediction models are published in the medical literature each year, yet many are developed using a dataset that is too small for the total number of participants or outcome events. This leads to inaccurate predictions and consequently incorrect healthcare decisions for some individuals. In this article, the authors provide guidance on how to calculate the sample size required to develop a clinical prediction model.

Richard D Riley *professor of biostatistics*¹, Joie Ensor *lecturer in biostatistics*¹, Kym I E Snell *lecturer in biostatistics*¹, Frank E Harrell Jr *professor of biostatistics*², Glen P Martin *lecturer in health data sciences*³, Johannes B Reitsma *associate professor*⁴, Karel G M Moons *professor of clinical epidemiology*⁴, Gary Collins *professor of medical statistics*⁵, Maarten van Smeden *assistant professor*^{4 5 6}



4. Write a data management plan

The **data life cycle**:

- what data will be collected
- how the data will be organized, stored, handled, and protected
- Sometimes required by funders, e.g. SNF
- There are trainings/courses on that topic
- The Findable, Accessible, Interoperable, and Reusable (FAIR) data principles: promote maximal use of data and enable machines to access and reuse data with minimal human intervention



5. Reduce bias

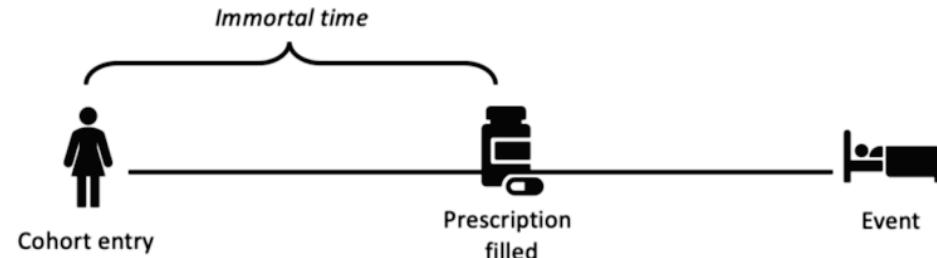
Table 1. Common types of bias that can affect a research study and some measures that may prevent them.

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 (dropouts or withdrawal). For example, more dropouts 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 - Following reporting guidelines

Immortal time bias

Survival beyond a certain time point is necessary to get the treatment

- e.g., COPD patients were assigned to the treatment group if they filled a prescription for a corticosteroid 90 days from discharge.
- they need to survive the time until prescription (immortal)

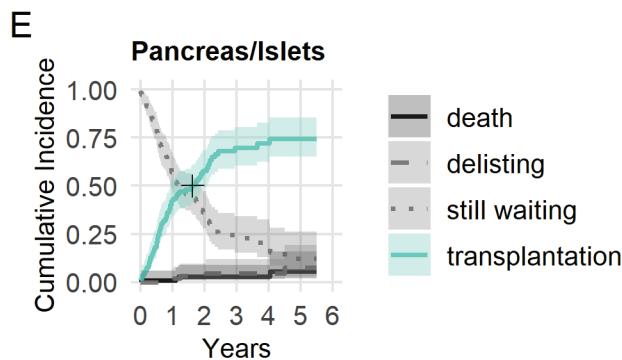
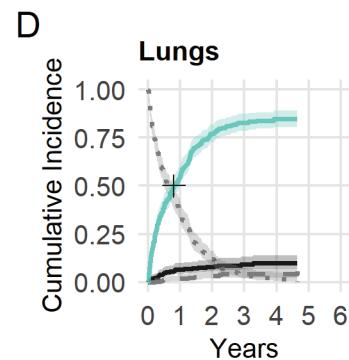
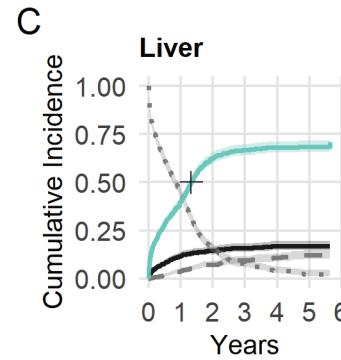
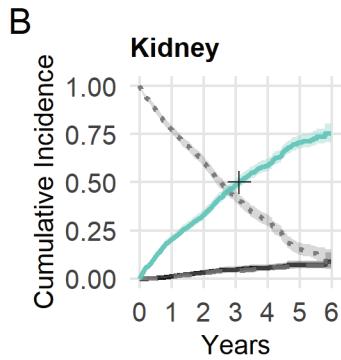
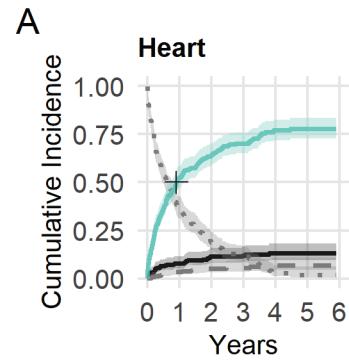


Preventive step:
Patients are assigned to groups at time-zero

Organ waiting list

Primary outcome: Time from listing to transplantation, death, or delisting

We are in a **competing risk** situation



- death
- delisting
- . still waiting
- transplantation

Selection bias when only looking at transplanted individuals when calculating waiting times.



6. Avoid questionable research practices

The seven sins of significance tests

Planning and design

When a study has **low statistical power** it is unlikely to obtain a statistically significant result, even when a true effect is present. Also, low power increases the false discovery rate (the proportion of studies with false positive results among all studies with positive results). Thus, the outcomes of underpowered studies must be interpreted with great caution. Low-powered studies can be avoided by using appropriate sample size calculation and supporting larger studies.

Pseudoreplication is the process of artificially inflating a study's sample size by treating data as independent when they are not. An example might be where repeated measurements are taken from a small sample of animals; however, the data are analysed as if they were independent measures from a larger sample. Doing so can lead to standard errors that are too small, and to results being found to be erroneously statistically significant. Such issues can be avoided by appropriate use of statistical methods that match the experimental design.

Execution and data collection

Repeated inspection of data when new data are added considerably increases the probability that a significant finding is a false positive. This can be avoided by not analysing the data before the study has been completed and the planned sample size has been reached. There are some scenarios in which stopping a study early can be justified. However, such situations need to be predefined, and adequate statistical methods must be used in the analysis.

Data processing and analysis

When researchers analyse data, there are many choices on how to process, transform, and model the data. **p-hacking** occurs when analysts exploit these "researcher degrees of freedom" until a statistically significant result is found. p-hacking can be avoided by writing a statistical analysis plan before data are available.

Presentation and interpretation

Selective reporting occurs when authors do not include particular results for any reason; for example, when the directionality of a particular finding is in conflict with that desired by the researcher. Transparent reporting can be ensured with preregistered studies, publicly available study protocols and adherence to reporting guidelines.

Publication

Confirmatory research requires hypotheses to be specified in advance, before the data are seen and a statistical test performed. Sometimes, however, researchers perform a number of tests until a statistically significant result is obtained. Then, the hypothesis is generated *post hoc*, as if it was in fact an *a priori* hypothesis. **HARKing** (hypothesising after the results are known) can be avoided with a preregistered protocol that includes a primary outcome and a statistical analysis plan.

Publication bias occurs when statistically non-significant results are less likely to be submitted by researchers and published by journals. Publication bias can be avoided with registered reports,⁷ whereby high-quality study protocols are provisionally accepted for publication prior to data analysis and independent of statistical significance.



7. Be cautious with interpretations of statistical significance



Simon Schwab • Sie
Senior Statistiker bei Swisstransplant
1 Monat • Bearbeitet •

Why we should not use the phrasing "statistically significant". I cringe sometimes when I hear this.

1. The 5% significance level is arbitrary. A result with $p=0.047$ is meaningful, but $p=0.051$ is not? It is about the same in terms of how compatible the data is with the null hypothesis.

2. Statistical significance does not mean practical significance or clinical significance. It is not about effects; you can have a very small effect and a very large sample size, and the p value will be very small. Such a result would be statistically significant but clinically meaningless (usually, but there are exceptions).

3. There is way too much emphasis on a statistically significant p value of 0.05 or smaller. Researchers may use questionable research practices to obtain the desired result.

...

4. A p value around 0.05 is pretty weak anyway. Claims of new discoveries should probably use a lower threshold.

5. We believe that journals or reviewers may prefer a result $p \leq 0.05$. However, it is often us, the researchers, who believe 0.05 or lower is noteworthy, and larger than 0.05 is not worth reporting. This leads to publication bias and harms evidence based medicine such as meta-analyses and systematic reviews.

6. Multiple studies can, in fact, each show a statistically non-significant result, all p values > 0.05 , but a meta analysis that combines all the studies can demonstrate evidence for a treatment benefit.

So, what to do about it? We should quit categorizing the result. Report exact p-values and interpret them in a graded way.

Moreover, we should put less emphasis on statistical significance and p values, anyway. It is just the tip of the iceberg. What is more important is how the result was obtained. Study protocols, reproducible analyses, following reporting guidelines etc. That all is a measure of the credibility of a study.

The p value or statistical significance is not.

<https://www.linkedin.com/feed/update/urn:li:activity:7177054485049942017/>



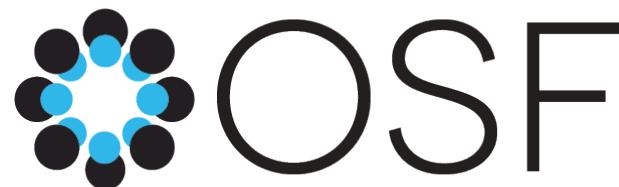
8. Make your research open

Open Science

- Open Study Protocols
- Open Code (GitHub)
- Open Data
- Open Access

2 – Gold Open Access – same publishing process as above. The difference is that when an article is accepted for publication, the author/s or funder/s pay an Article Processing Charge (APC). The final version of the published article is then free to read for everyone. The APC to publish Gold Open Access in *Nature Neuroscience* is £8890.00/\$12290.00/€10290.00.

Open Science Framework <https://osf.io/>





Hello, Quarto

```
---
```

```
title: "ggplot2 demo"
author: "Norah Jones"
date: "5/22/2021"
format:
  html:
    fig-width: 8
    fig-height: 4
    code-fold: true
---
```

```
## Air Quality
```

```
@fig-airquality further explores the impact of temperature on ozone level.
```

```
```{r}
#| label: fig-airquality
#| fig-cap: "Temperature and ozone level."
#| warning: false
```

```
library(ggplot2)
```

```
ggplot(airquality, aes(Temp, Ozone)) +
 geom_point() +
 geom_smooth(method = "loess")
)```

```

## ggplot2 demo

Norah Jones

May 22nd, 2021

### Air Quality

[Figure 1](#) further explores the impact of temperature on ozone level.

#### ► Code

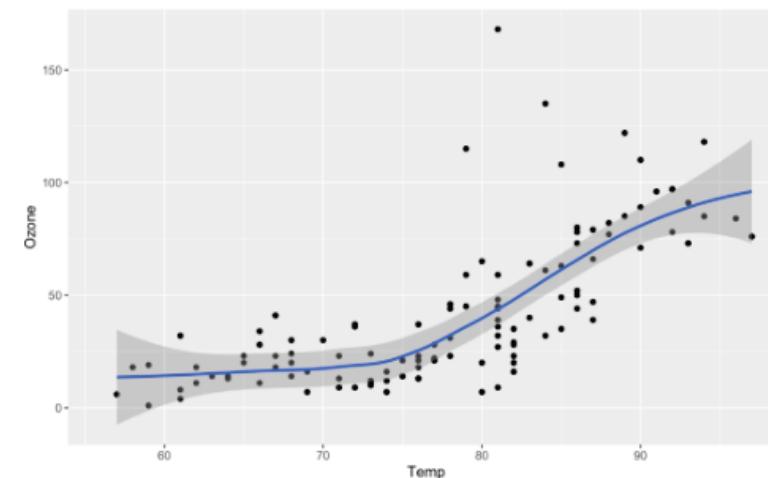


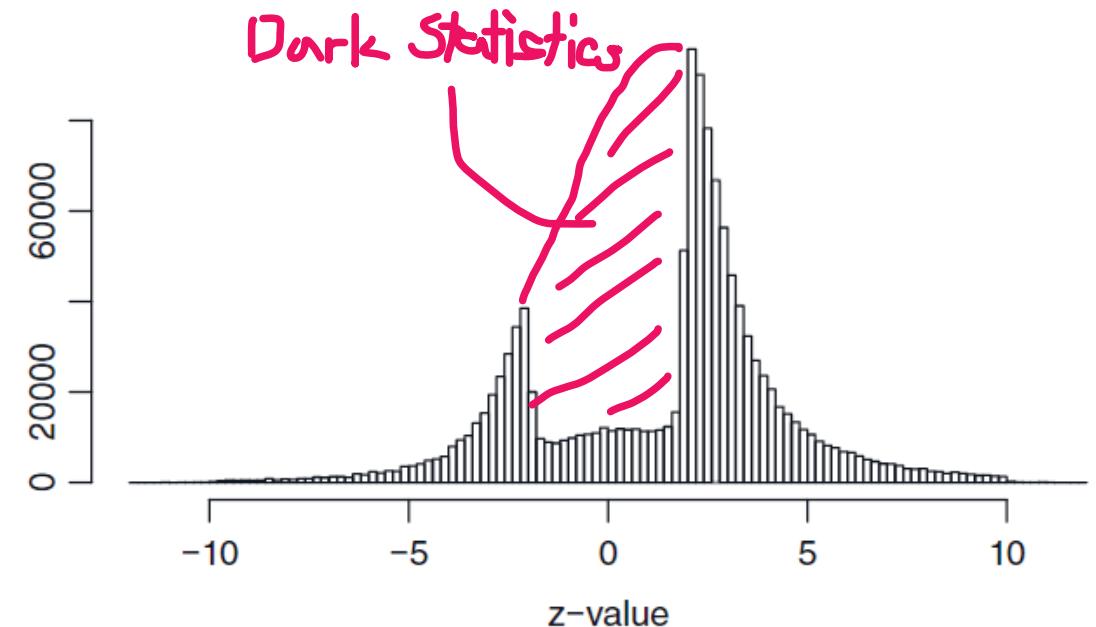
Figure 1: Temperature and ozone level.

## 9. Report all findings

- 1 million z-values from Medline (1976–2019)
- Under-representation of z-values between –2 and 2 is striking

van Zwet & Cator (2020)

In clinical research, publication bias can **mislead decision-making** and harm patients.





## 10. Follow reporting guidelines

<b>Reporting guideline</b>	<b>Study type</b>
CONSORT	Randomized trials
STROBE	Observational studies
PRISMA	Systematic reviews
SPIRIT	Study protocols (clinical trials)
TRIPOD	Prognostic studies
START	Diagnostic studies

<https://www.equator-network.org/>



## RESEARCH METHODS AND REPORTING

OPEN ACCESS



### TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods

Gary S Collins,<sup>1</sup> Karel G M Moons,<sup>2</sup> Paula Dhiman,<sup>1</sup> Richard D Riley,<sup>3,4</sup> Andrew L Beam,<sup>5</sup> Ben Van Calster,<sup>6,7</sup> Marzyeh Ghassemi,<sup>8</sup> Xiaoxuan Liu,<sup>9,10</sup> Johannes B Reitsma,<sup>2</sup> Maarten van Smeden,<sup>2</sup> Anne-Laure Boulesteix,<sup>11</sup> Jennifer Catherine Camaradou,<sup>12,13</sup> Leo Anthony Celi,<sup>14,15,16</sup> Spiros Denaxas,<sup>17,18</sup> Alastair K Denniston,<sup>4,9</sup> Ben Glocker,<sup>19</sup> Robert M Golub,<sup>20</sup> Hugh Harvey,<sup>21</sup> Georg Heinze,<sup>22</sup> Michael M Hoffman,<sup>23,24,25,26</sup> André Pascal Kengne,<sup>27</sup> Emily Lam,<sup>12</sup> Naomi Lee,<sup>28</sup> Elizabeth W Loder,<sup>29,30</sup> Lena Maier-Hein,<sup>31</sup> Bilal A Mateen,<sup>17,32,33</sup> Melissa D McCradden,<sup>34,35</sup> Lauren Oakden-Rayner,<sup>36</sup> Johan Ordish,<sup>37</sup> Richard Parnell,<sup>12</sup> Sherri Rose,<sup>36</sup> Karandeep Singh,<sup>38</sup> Laure Wynants,<sup>40</sup> Patricia Logullo<sup>1</sup>

## Contact

Simon Schwab  
[simon.schwab@swisstransplant.org](mailto:simon.schwab@swisstransplant.org)

**Swisstransplant**  
Effingerstrasse 1  
Postfach  
CH-3011 Bern  
Telefon +41 58 123 80 00  
[www.swisstransplant.org](http://www.swisstransplant.org)  
[info@swisstransplant.org](mailto:info@swisstransplant.org)

