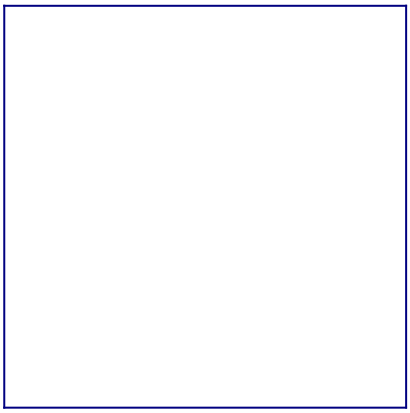
Statistical Thinking News

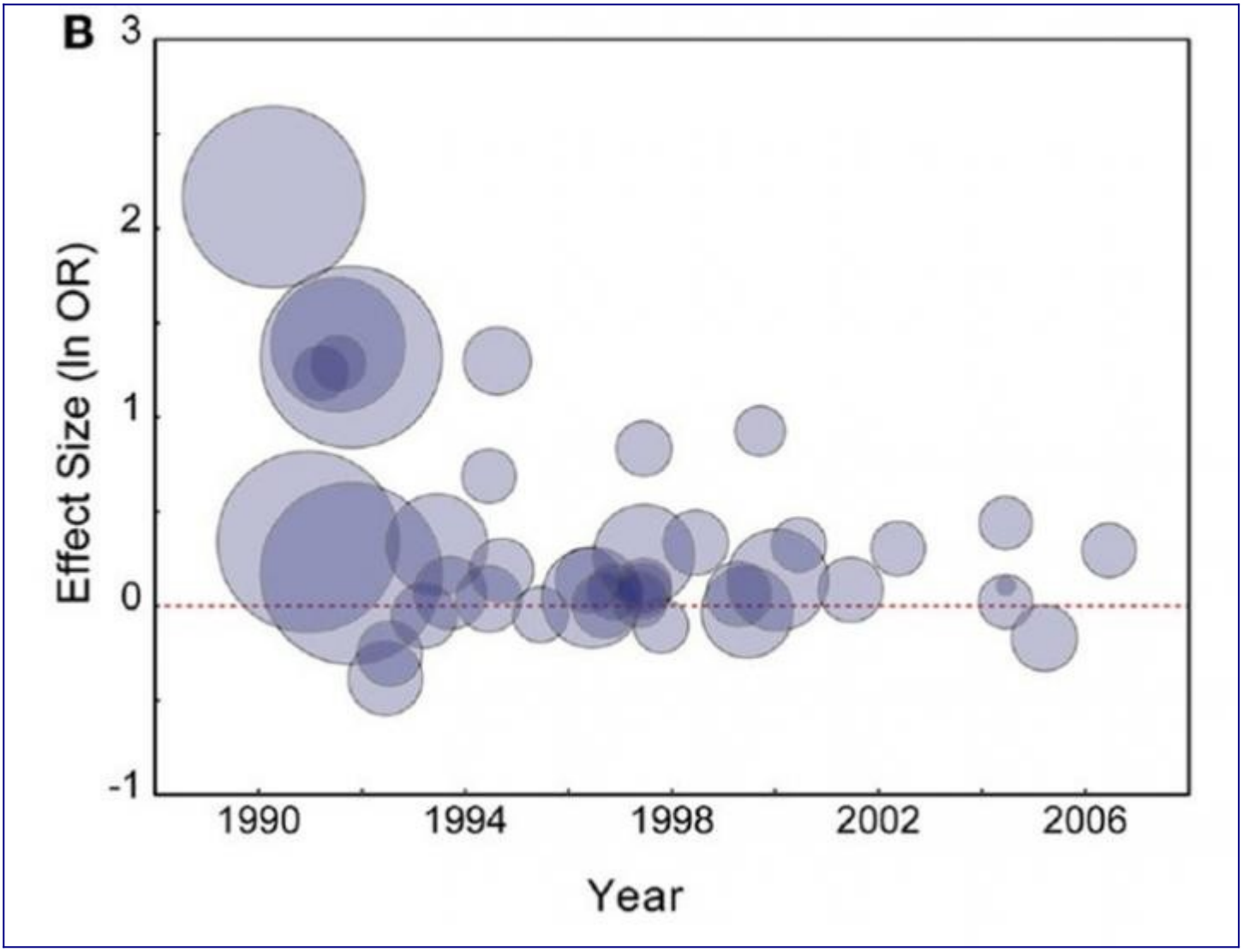
News and Opinions on Data Analysis and Statistical Modeling, Prediction, Statistical Computing, Research Design and Interpretation, Clinical Trials, and Research Integrity

Frank Harrell

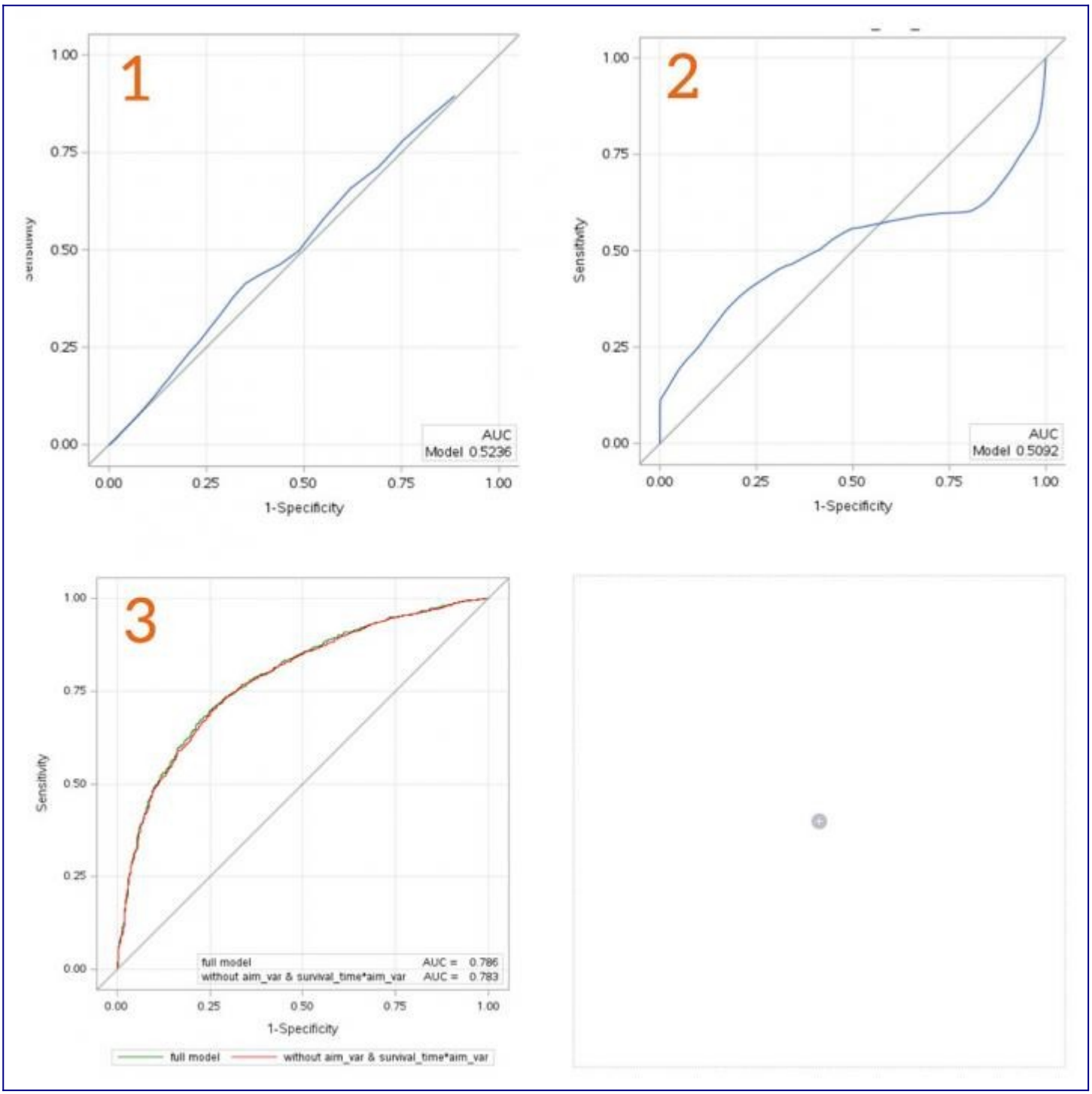
[On improving the efficiency of trials via linear adjustment for a prognostic score](#)
thestatsgeek.com
I've recently had the opportunity to spend a little time looking at an interesting approach for improving the efficiency of estimated treatment effects in clinical trials which exploits histo…
Shared by Jonathan Bartlett



[PNAS is Not a Good Journal](#)
getsyeducated.substack.com
(and Other Hard Truths about Journal Prestige)
Shared by Retraction Watch



ROC has a S shape, why?
discourse.datamethods.org
Background: I am analyzing data from a survival-study. I am a medical PhD, so I tried to get advanced knowledge in biostatistics, but maybe I have some „simple“ mistakes in my understanding. :see_no_ …



[Science and Ethics of “Curing” Misinformation](#)
journalofethics.ama-assn.org
[Misinformation is an urgent new problem, so health professions communities need solutions as much as they need to be wary of ethical pitfalls of rushed interventions.](#)
Shared by [Retraction Watch](#)

Is writing a book chapter still a waste of time?
blogs.lse.ac.uk
Patrick Dunleavy argues thanks to digital open access, book chapters once again rival journal articles in their visibility & usefulness to academic communities.
Shared by Retraction Watch

On the relevance of prognostic information for clinical trials: A theoretical quantification

pubmed.ncbi.nlm.nih.gov

The question of how individual patient data from cohort studies or historical clinical trials can be leveraged for designing more powerful, or smaller yet equally powerful, clinical trials becomes in…

Shared by Stephen John Senn

Here's one article that won't be making any top 50 papers list

retractionwatch.com

Who doesn't love a list? The 500 best rock songs of all time. The 100 tallest buildings on the planet. The 10 smartest dog breeds. The 14 silliest place-names on earth (with Middelfart, Denmark in th…

Shared by Retraction Watch

Long-COVID

icudelirium.org

Long-COVID shares key similarities and differences with PICS, with one key difference being that Long-COVID can impact people who were never severely ill and, in some cases, people who were hardly il…

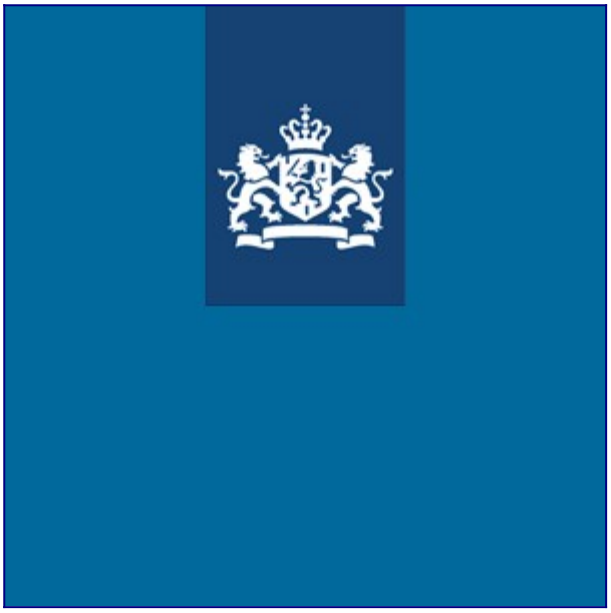
Shared by WesElyMD

Guideline for high-quality diagnostic and prognostic applications of AI in healthcare

datavoorgezondheid.nl

This guideline provides a description of what the work field considers good professional conduct in the development, testing and implementation of an Artificial Intelligence Prediction Algorithm (AIP…

Shared by Maarten van Smeden

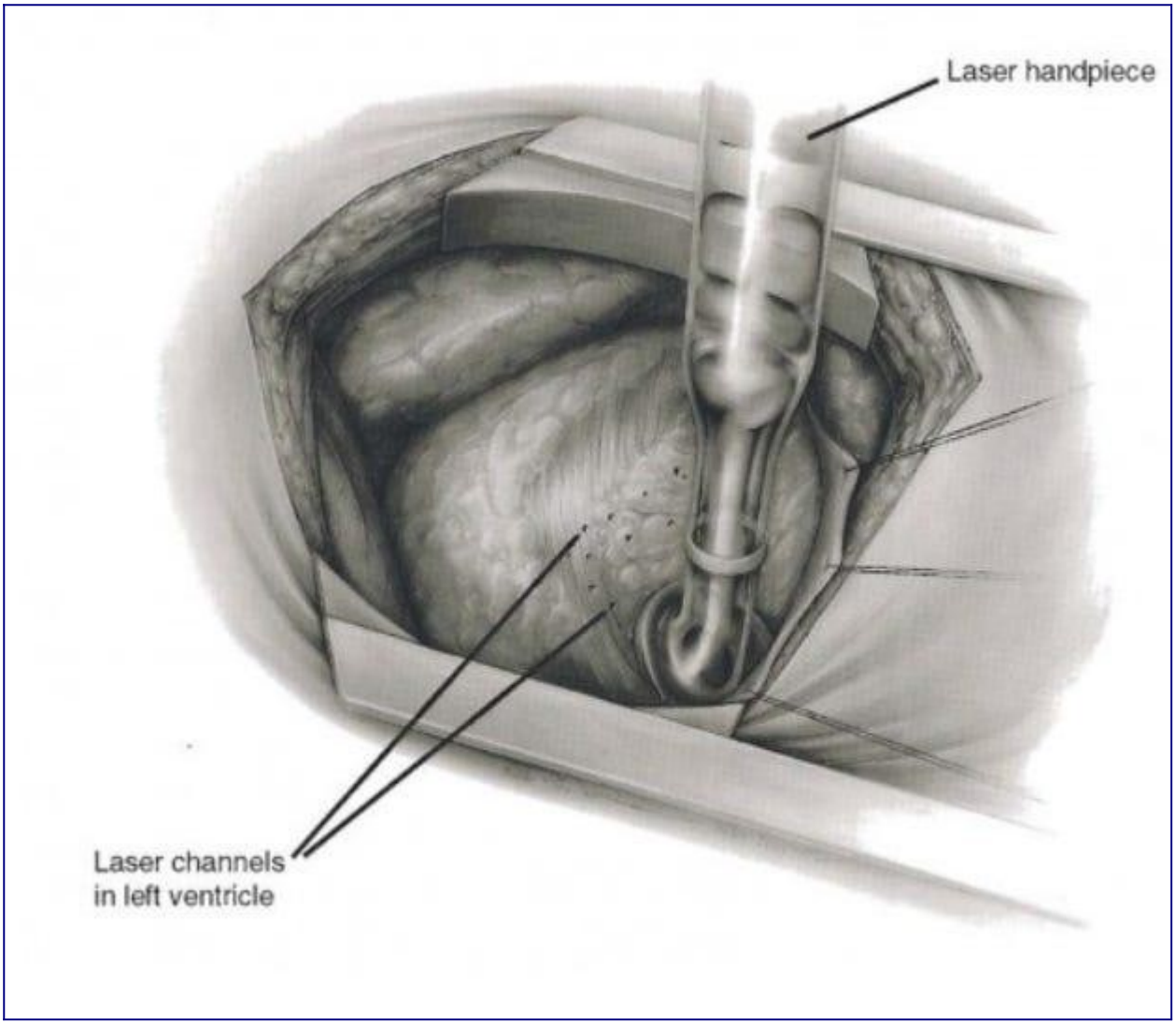


How Doctors are Fooled

sensiblemed.substack.com

For the next few weeks I will explore the need for having a proper control arm in medical studies

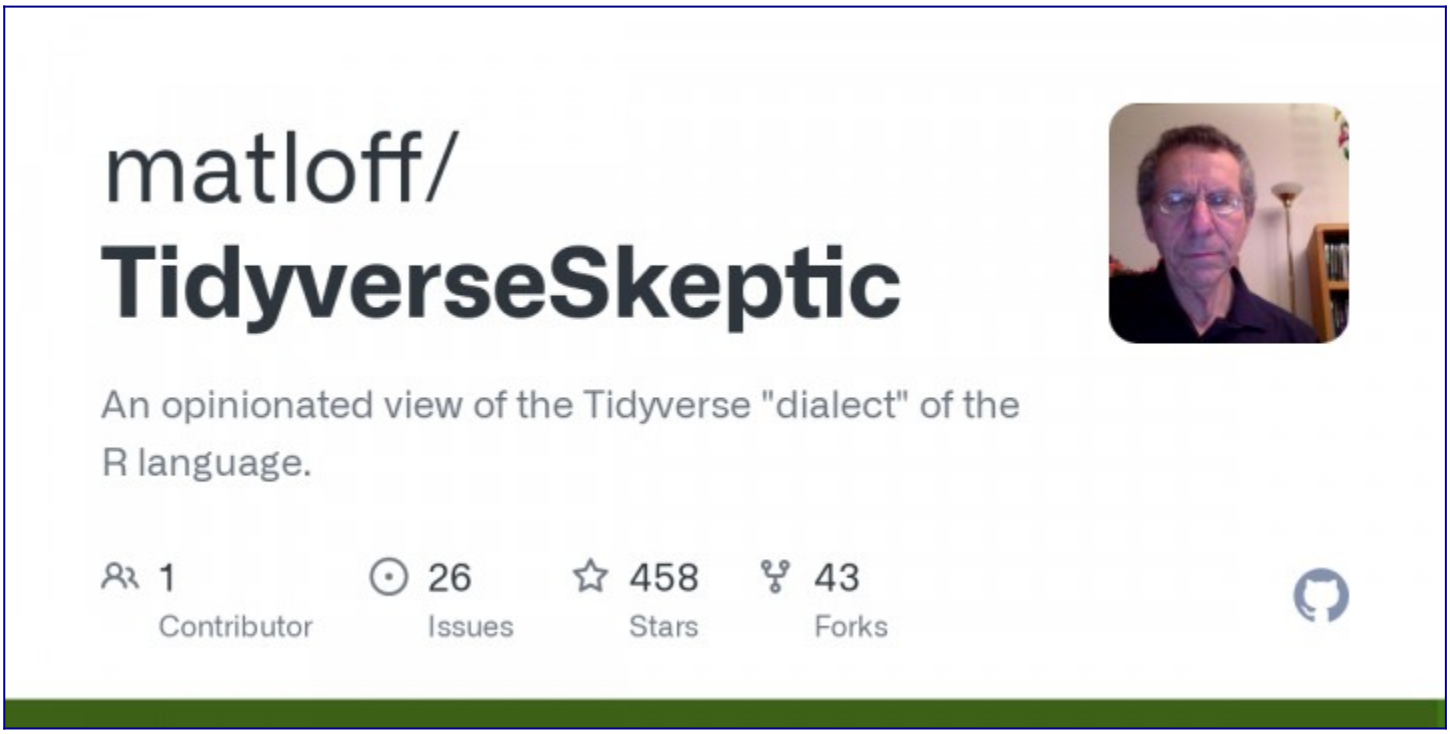
Shared by John Mandrola, MD



Significance testing, the replication crisis, and the market for lemons
statmodeling.stat.columbia.edu
In an article, “Accounting research and the significance test crisis,” David Johnstone writes: There are now hundreds of published papers and statements, echoing what has been said behind closed door…
Shared by Andrew Gelman et al.

Everybody's backyard
statsepi.substack.com
John Tukey said that the best thing about being a statistician is that you get to play in everyone’s backyard. It’s a great quote. While it more obviously reflects the tendency of applied statisticia…
Shared by Darren Dahly https://mastodon.social/@statsepi

GitHub - matloff/TidyverseSkeptic: An opinionated view of the Tidyverse "dialect" of the R language.
github.com
An opinionated view of the Tidyverse "dialect" of the R language. - GitHub - matloff/TidyverseSkeptic: An opinionated view of the Tidyverse "dialect" of the R language.
Dr Matloff's strong arguments against exposing beginners early to the R "tidyverse" are even understated in my opinion.
Shared by Karandeep Singh



The Problem of Old Evidence
bayesianspectacles.org
To my shame and regret, I only recently found the opportunity to read the book “Bayesian philosophy of science” (BPS), by Jan Sprenger and Stephan Hartmann. It turned out to be a wonderful b…
Shared by EJ Wagenmakers

CAPSULE COMPOST
BAYESIAN
PHILOSOPHY
of SCIENCE
JAN SPRENGER
STEPHAN HARTMANN

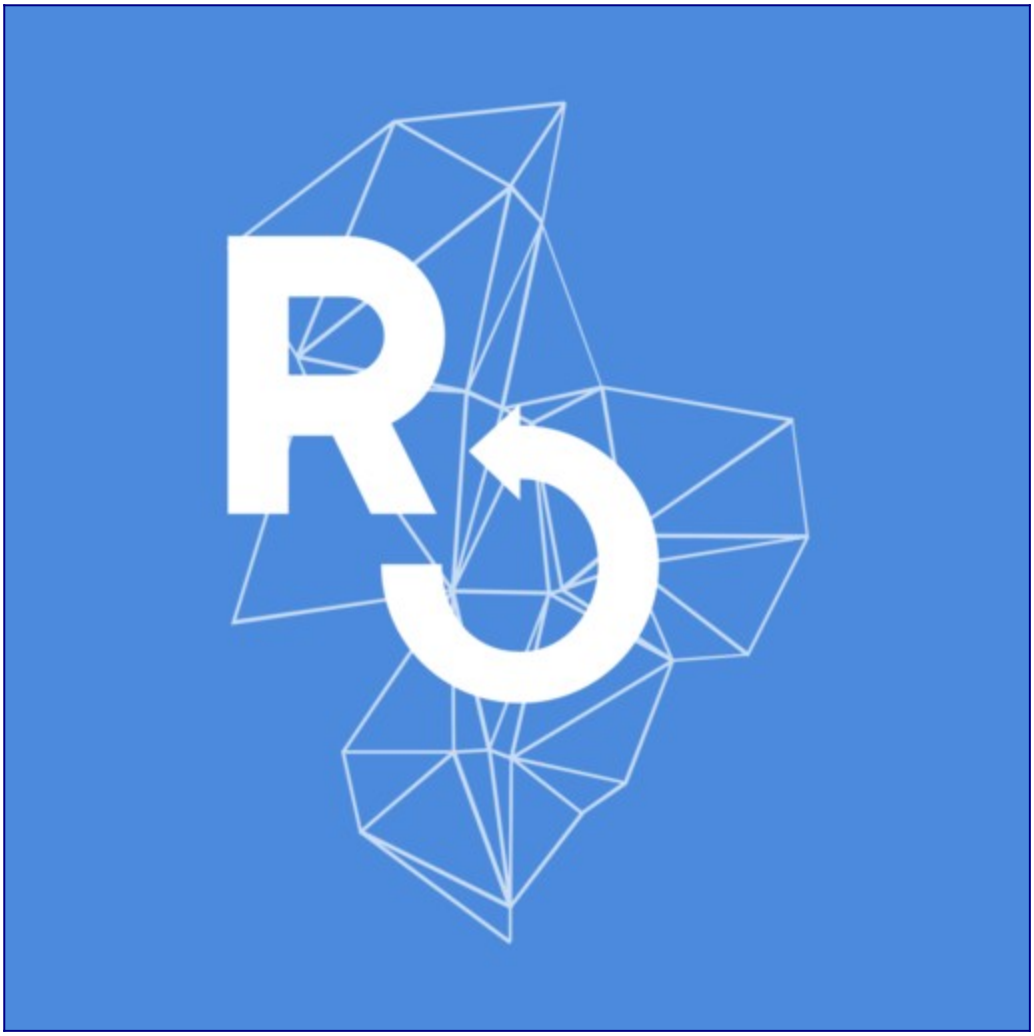
Reference Collection to push back against "Common Statistical Myths"
discourse.datamethods.org
Note: This topic is a wiki, meaning that this main body of the topic can be edited by others. Use the Reply button only to post questions or comments about material contained in the body, or to sugge…
Shared by Atsuhiro Oguri

Receive the newsletter

Why artificial intelligence needs to understand consequences
nature.com
A machine with a grasp of cause and effect could learn more like a human, through imagination and regret.
Shared by Andrew Neff

15 Best Statistics Podcasts
blog.feedspot.com
List of Best Statistics Podcasts. Keep up with statisticians, data scientists community, R and statistics, statistical computing, software used to develop statistical algorithms, visualizations, and …
Shared by Stats + Stories

[Discovering and learning everything there is to know about R packages using r-universe](#)
[ropensci.org](#)
[The goal of r-universe is to provide a central place for browsing through the R ecosystem to discover what is out there, get a sense of the purpose and quality of individual packages, and help you ge…](#)
Shared by [Mara Averick](#)



Rohrer and Arslan's nonet: More ideas regarding interactions in statistical models
statmodeling.stat.columbia.edu
I liked the causal quartet you recently posted and wanted to forward a similar homage (in style if not content) Julia Rohrer and I recently made to accompany this paper. We had to go to a triple trip…
Shared by Andrew Gelman et al.

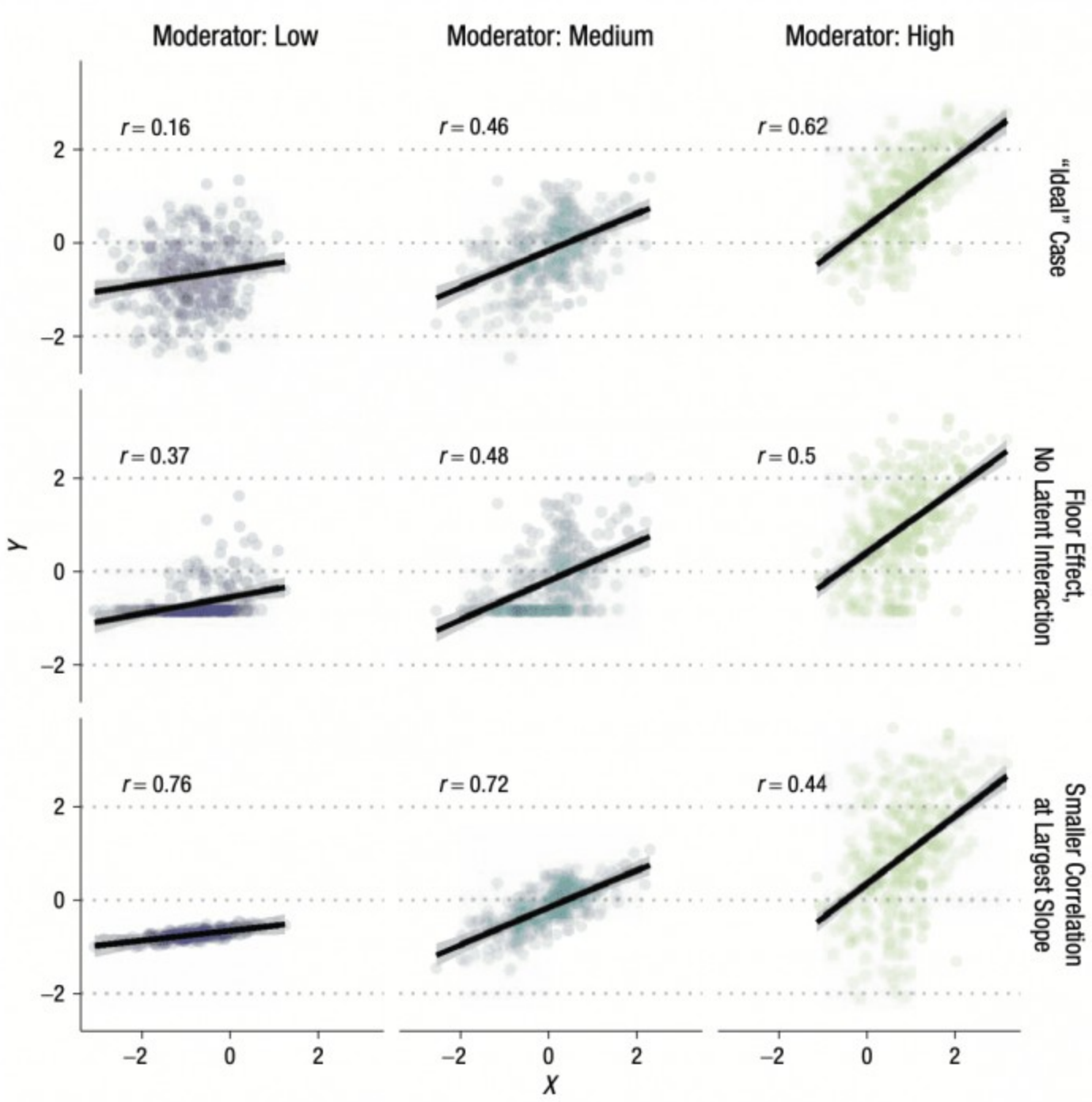
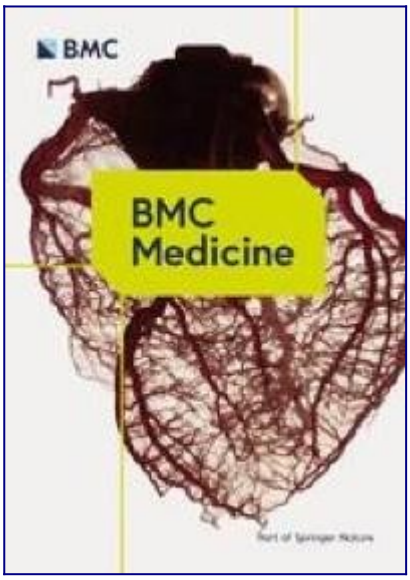


Fig. 6. A triptych in homage to Anscombe. Columns show levels of the moderator variable (*M*). *X* and *M* are the same across rows, whereas *Y* has been generated according to different scenarios but maintains the same mean and variance. The slopes estimated according to a simple linear interaction model (black lines) are the same across rows, but graphing the raw data shows that the data-generating processes were quite different. Shaded regions around the regression lines show 95% credible intervals.

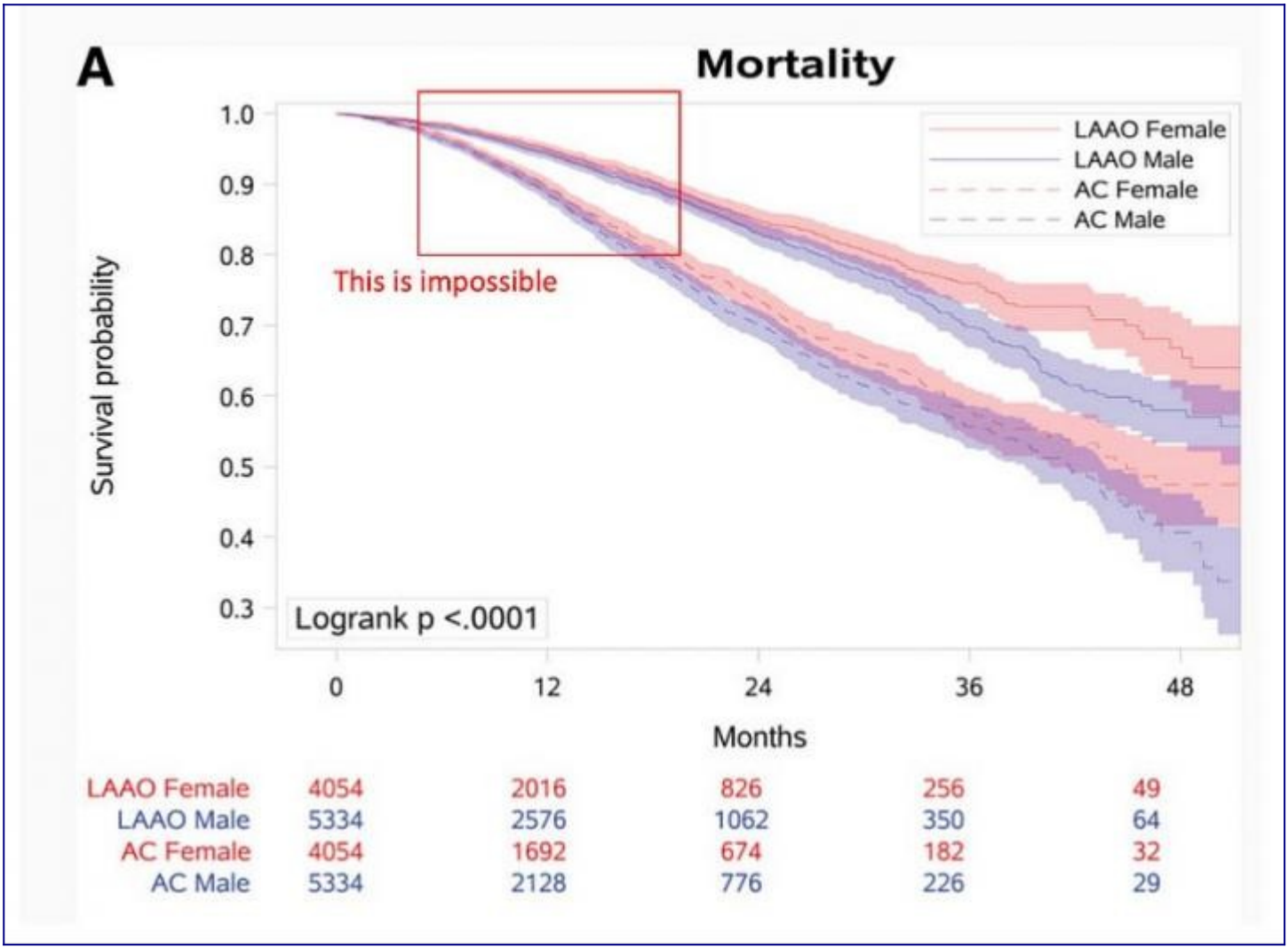
[The new scientific review on masks and Covid isn't what you think](#)
[vox.com](#)
[A meta-analysis seeks to be the last word on the effectiveness of masks, but finding answers in science isn't that easy.](#)

Shared by Dr. Saskia Popescu

There is no such thing as a validated prediction model
bmcmedicine.biomedcentral.com
Clinical prediction models should be validated before implementation in clinical practice. But is favorable performance at internal validation or one external validation sufficient to claim that a pr…
Shared by Darren Dahly https://mastodon.social/@statsepi



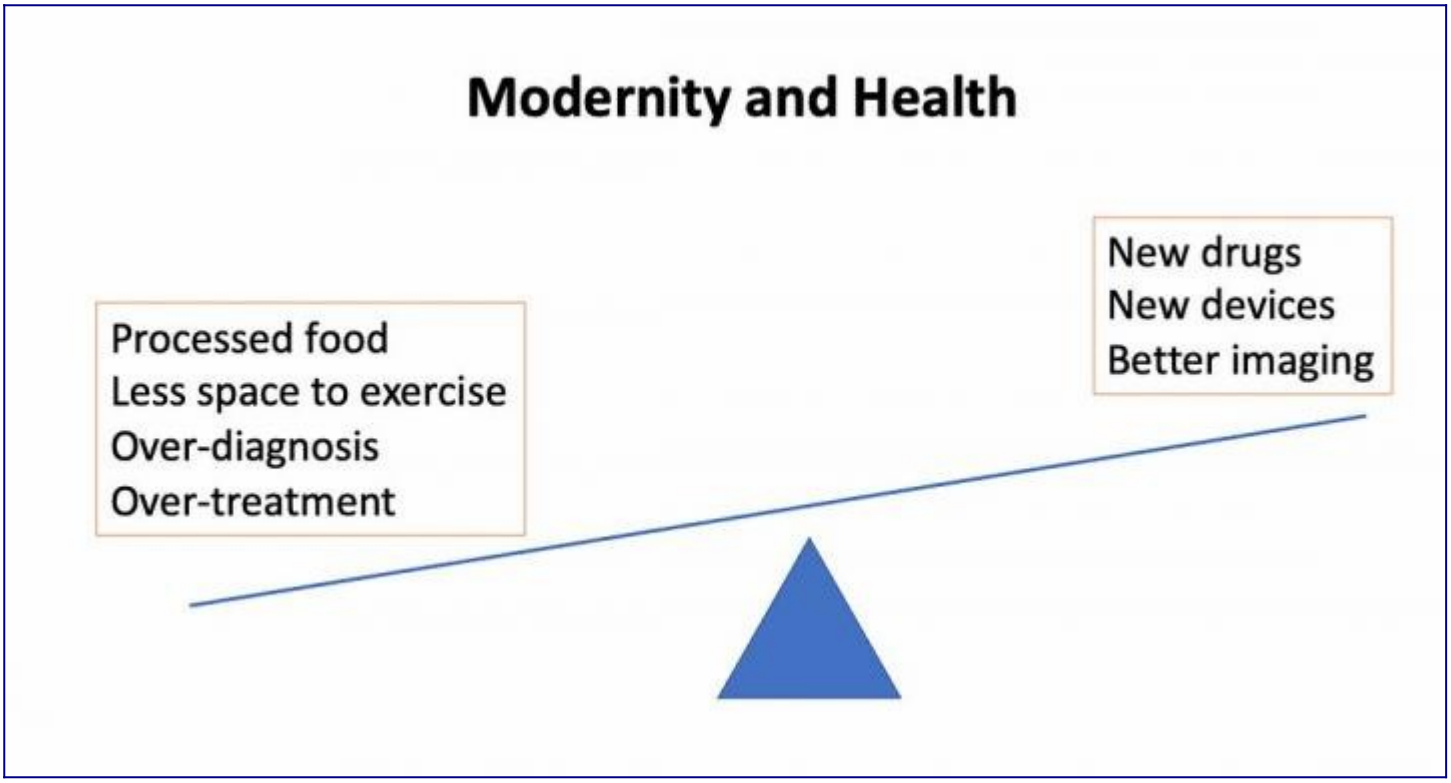
A Flawed Analysis Induces Severe Sadness
sensiblemed.substack.com
A prominent journal in cardiology has published one of the most flawed analyses I have seen in years. And the conclusions are even worse.
Shared by Frank Harrell



[Modernity and Health](johnmandrola.substack.com)

[A recent report on incidental findings from a cardiac monitoring study bolsters my view that modernity is shredding our health. The paper is a re-analysis of one of the most important cardiac trials …]

Shared by [John Mandrola, MD]



FDA "Bad Habits"
discourse.datamethods.org
A February 21, 2023 Wall Street Journal editorial titled “The FDA Returns to Its Bad Habits” discusses a randomized trial of a medication to treat Friedreich’s Ataxia sponsored by Reata Pharmaceutica…
Statistical Thinking - Randomized Clinical Trials Do Not Mimic Clinical Practice, Thank Goodness
fharrell.com
Randomized clinical trials are successful because they do not mimic clinical practice. They remain highly clinically relevant despite this.
Shared by Frank Harrell
Bayesian Logistic Regression with rstanarm
avehtari.github.io
The introduction to Bayesian logistic regression and rstanarm is from a CRAN vignette by Jonah Gabry and Ben Goodrich. CRAN vignette was modified to this notebook by Aki Vehtari. Instead of wells dat…
Shared by 5’UTR
Statistical Thinking - Randomized Clinical Trials Do Not Mimic Clinical Practice, Thank Goodness
fharrell.com
Randomized clinical trials are successful because they do not mimic clinical practice. They remain highly clinically relevant despite this.
Shared by Frank Harrell
Assumptions and analysis planning in studies with missing data in multiple variables: moving beyond the MCAR/MAR/MNAR classification
academic.oup.com

Abstract. Researchers faced with incomplete data are encouraged to consider whether their data are ‘missing completely at random’ (MCAR), ‘missing at random’ (M

Shared by Jonathan Bartlett

How to... use AI to teach some of the hardest skills

oneusefulthing.substack.com

When errors, inaccuracies, and inconsistencies are actually very useful

Shared by Ethan Mollick

Should multiple imputation be stratified by exposure group when estimating causal effects via outcome regression in observational studies?

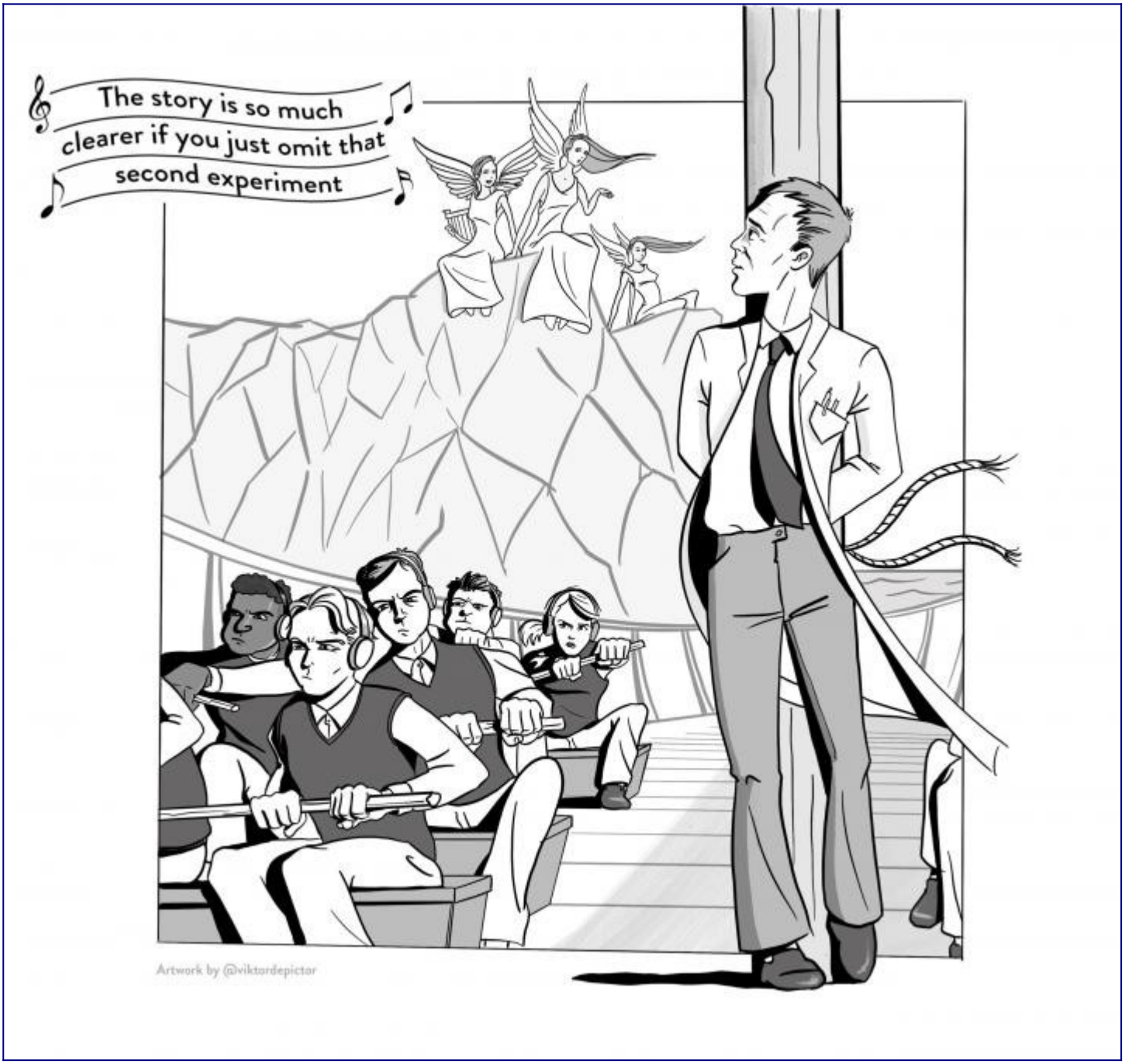
bmcmedresmethodol.biomedcentral.com

Despite recent advances in causal inference methods, outcome regression remains the most widely used approach for estimating causal effects in epidemiological studies with a single-point exposure and…

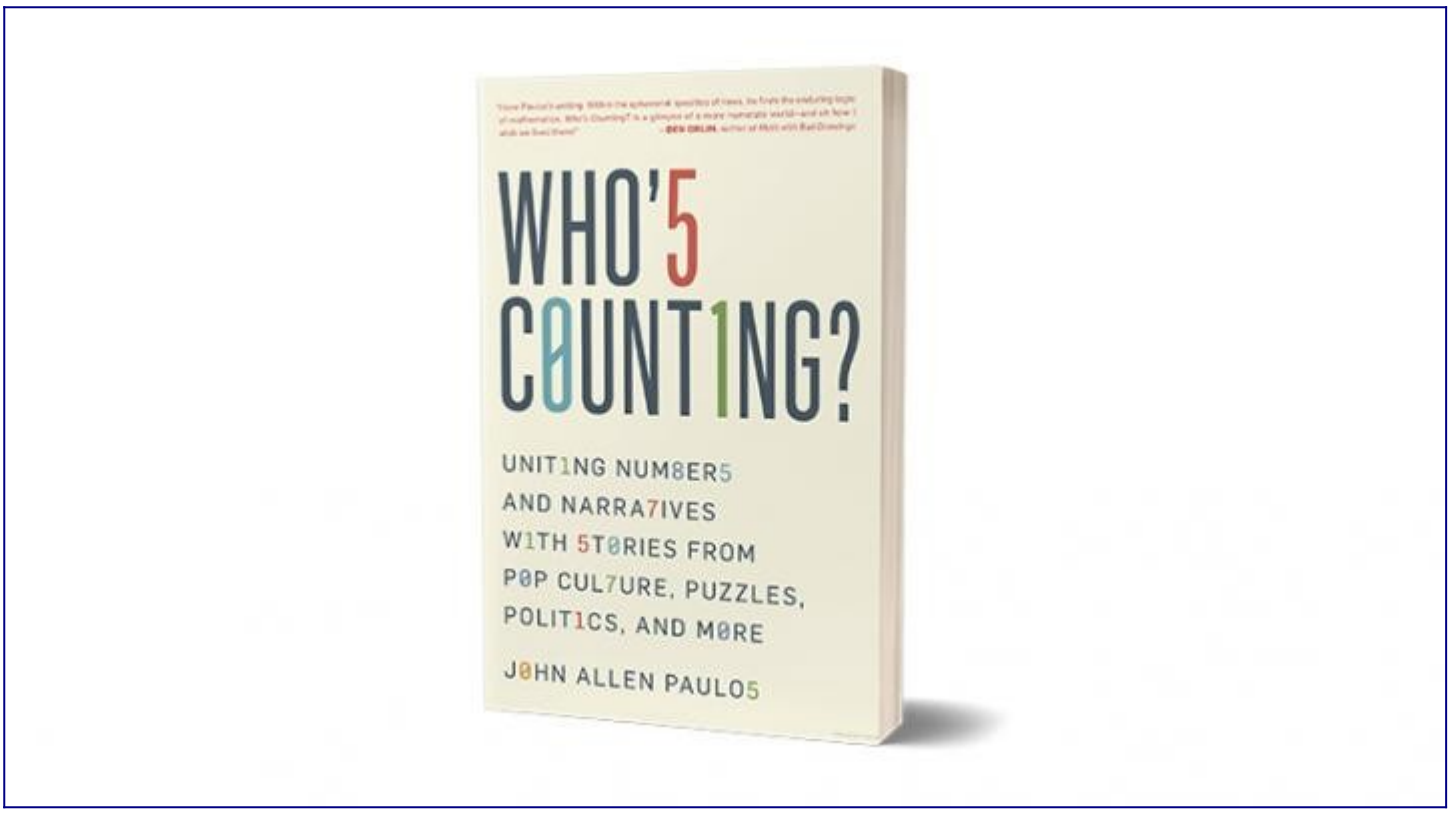
Shared by Jonathan Bartlett

Climate Care is Healthcare with Dr. Reed Omary
healthcarenowradio.com
Now on demand, Health Unabashed episode with guest Reed Omary MD, the founder and director of the Vanderbilt School of Medicine Medical Innovators Development Program.
Shared by Reed Omary, MD

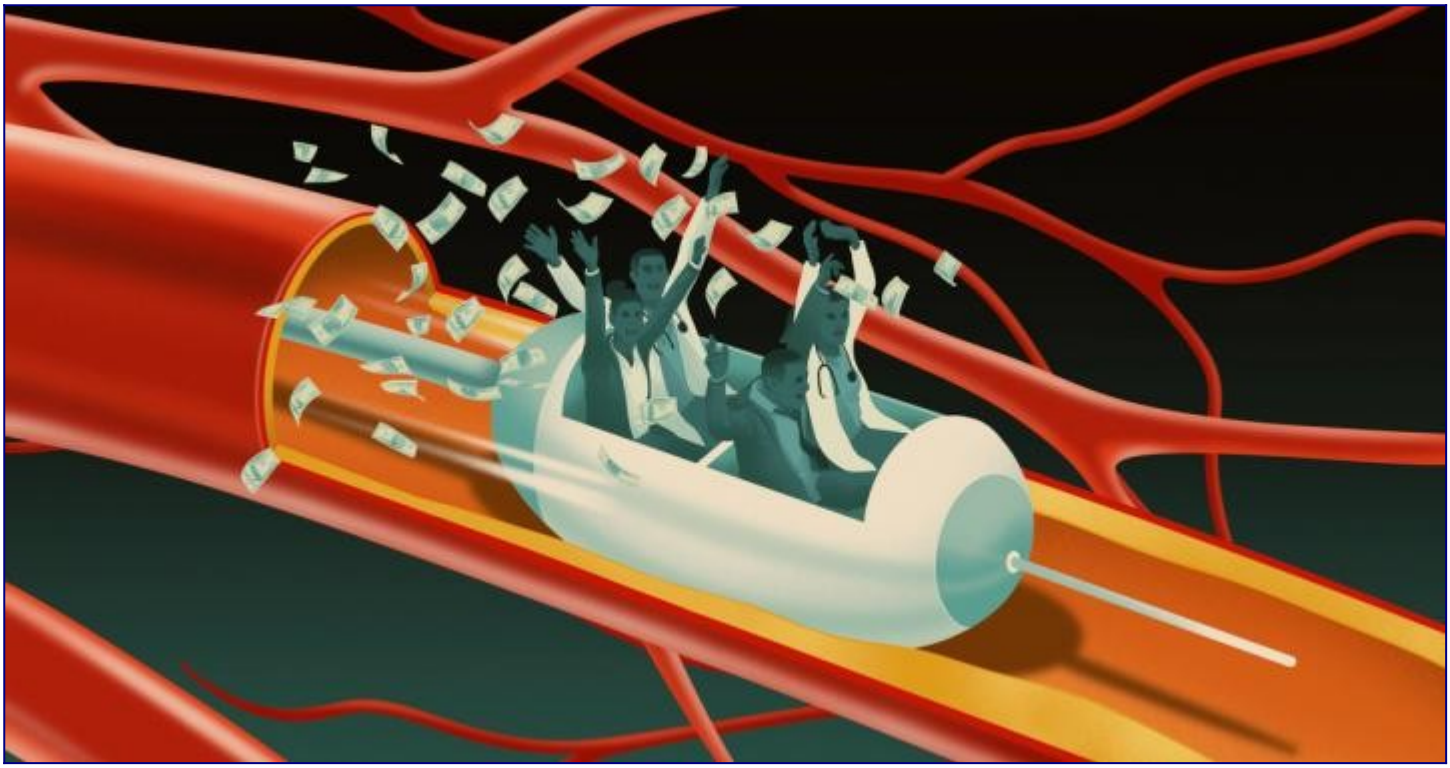
A Presentation on Transparency in Science and Statistics
bayesianspectacles.org
This month I gave a 45-minute presentation “Transparency in Science and Statistics” for the Italian Reproducibility Network. This presentation reflects my recent thinking on the topic. Important th…
Shared by EJ Wagenmakers



[Strength in numbers](#)
[shaastramag.iitm.ac.in](#)
[Paulos links politics, puzzles and probability — and lies, logic and religion — to tell us how we may become better citizens by acting on the connections.](#)
Shared by [John Allen Paulos](#)



[Steak Dinners, Sales Reps and Risky Procedures: Inside the Big Business of Clogged Arteries](propublica.org)
[Text messages, a whistleblower lawsuit and an internal investigation reveal the lengths to which Medtronic, the world's largest medical device company, allegedly "groomed" doctors to overuse its vasc…]
Shared by [Boback Ziaeian]



FDA response to Citizen Petition to include Quality of Life comparisons in RCTs
discourse.datamethods.org
Copying a snippet of the FDA response to advocate's petition (that would be me) along with links to the full text and the original petition. The issue remains. Quality of life is integral to the asse…

rang: make ancient R code run again
blog.schochastics.net
Reproducibility is a big issue in the (computational) world of science. Code that runs today might not run tomorrow because packages are updated, functions deprecated or removed, and whole programmin…
Shared by Mara Averick

Causal and Associational Language in Observational Health Research: A systematic evaluation

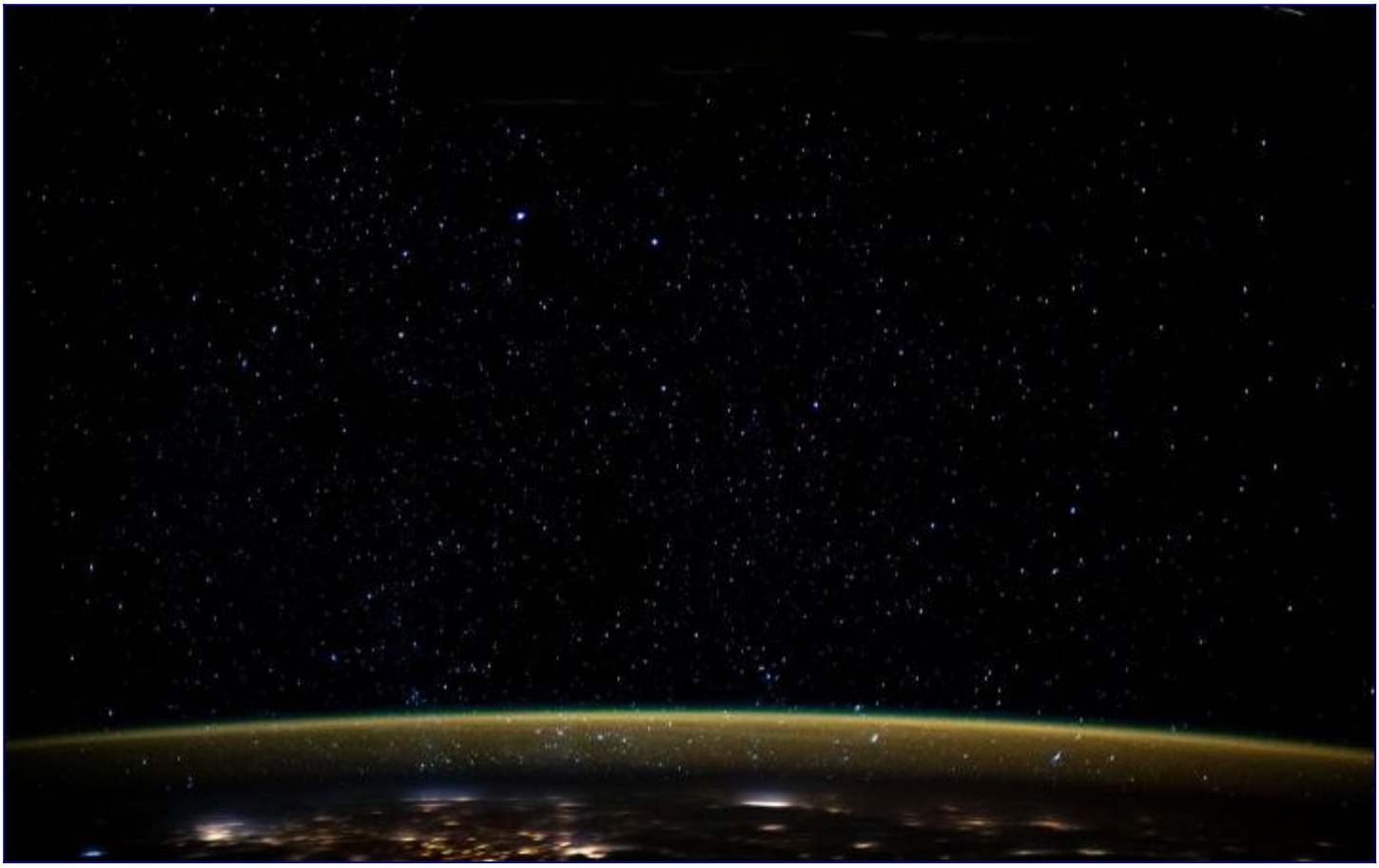
medrxiv.org

medRxiv - The Preprint Server for Health Sciences

Shared by Ethan Mollick

Our cosmic horizon is both unreachable and closer than ever | Aeon Essays
aeon.co
Our planet is a tiny porthole, looking over a cosmic sea. Can we learn what lies beyond our own horizons of perception?

Shared by Aeon+Psyche



Prolific autism researcher has two dozen papers retracted | Spectrum | Autism Research News
spectrumnews.org
Elsevier’s retractions focus on peer review and conflicts of interest.

Shared by Retraction Watch

Toward Guidelines for RCT Measurements.. What are the characteristics of an acceptable measurement for RCT?

discourse.datamethods.org

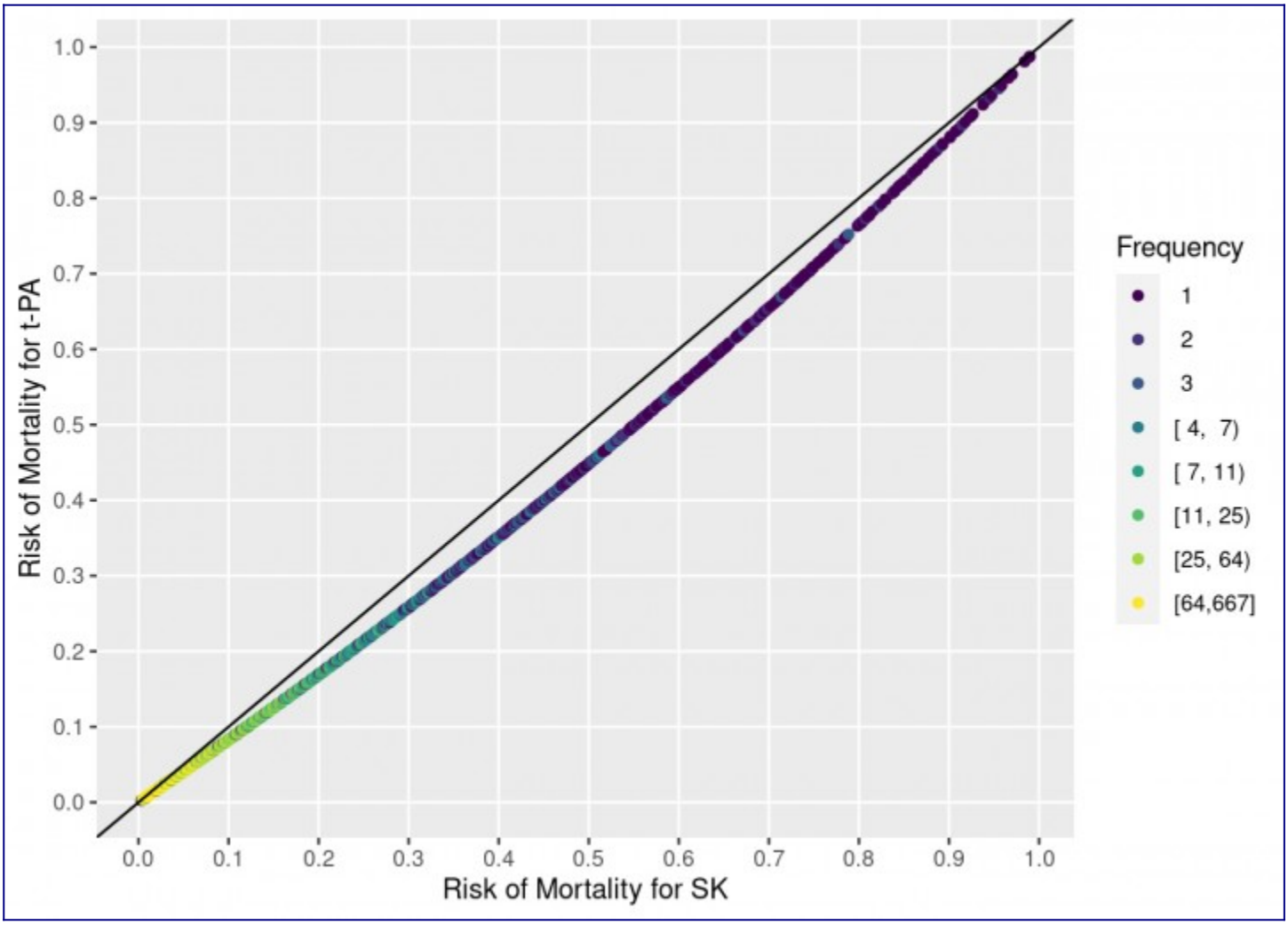
@Stephen Senn provided this reference in a twitter discussion about RCT vs OBS studies. So measurement has long been overlooked as a source of non-reproducibility in RCT. I have been intrigued by the…

Statistical Thinking - Avoiding One-Number Summaries of Treatment Effects for RCTs with Binary Outcomes

fharrell.com

This article presents an argument that for RCTs with a binary outcome the primary result should be a distribution and not any single number summary. The GUSTO-I study is used to exemplify risk differ…

Shared by Frank Harrell

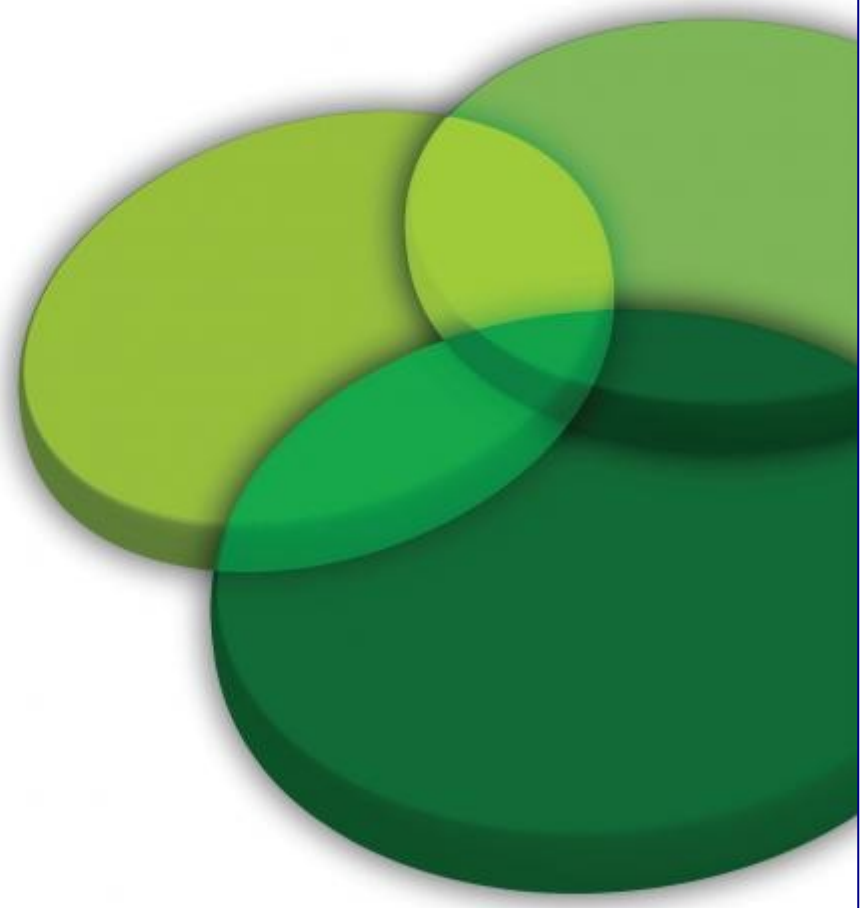


Interpreting the results from the first randomised controlled trial of colonoscopy: does it save lives?
ebm.bmj.com
In screening for colorectal cancer (CRC), colonoscopy is generally referred to as the ‘gold standard’, as it examines both the right and left colon. Despite wide adoption, colonoscopy had not been st…
Shared by Vinay Prasad MD MPH

February 2023 Volume 28 Issue 1

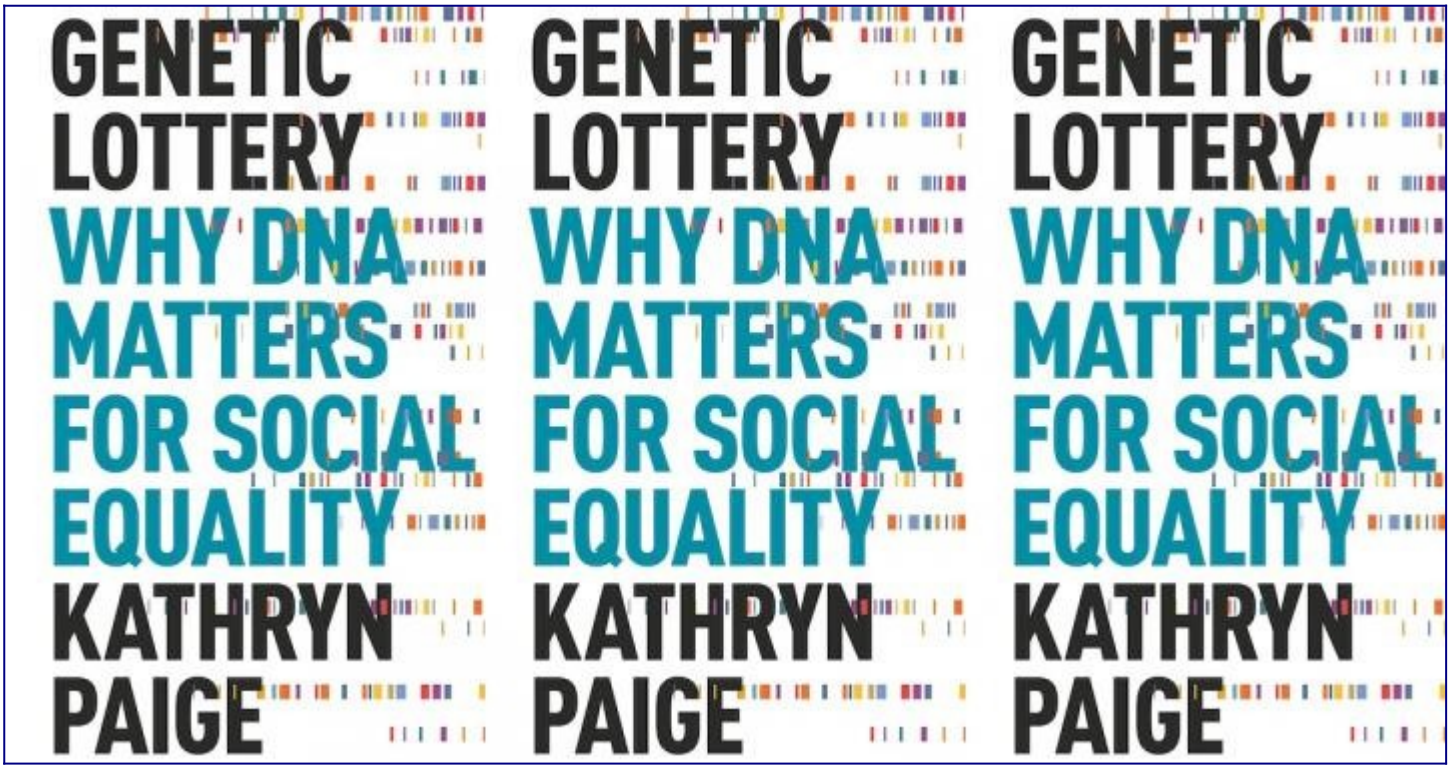
BMJ EVIDENCE-BASED MEDICINE

EXPERTISE VALUES EVIDENCE

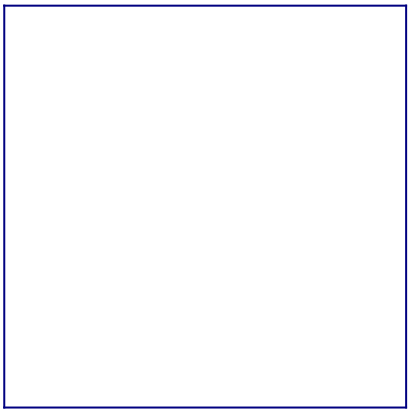


BMJ

The Genetic Lottery is a bust for both genetics and policy
massivesci.com
Kathryn Paige Harden’s book tries to demonstrate how genetics can ameliorate societal ills. She falls well, well short
Shared by pims simp



G-formula for causal inference via multiple imputation
thestatsgeek.com
G-formula (sometimes known as G-computation) is an approach for estimating the causal effects of treatments or exposures which can vary over time and which are subject to time-varying confounding. …
Shared by Jonathan Bartlett



All my classes suddenly became AI classes
oneusefulthing.substack.com
We can't beat AI, but it doesn't need to beat us (or our students)
Shared by Ethan Mollick

The Machines of Mastery
oneusefulthing.substack.com
"Anyone can learn anything they want..." and how technology can help
Shared by Ethan Mollick

Statistical Rethinking with brms, ggplot2, and the tidyverse
bookdown.org
This project is an attempt to re-express the code in McElreath's textbook. His models are re-fit in brms, plots are redone with ggplot2, and the general data wrangling code predominantly follows the …

Editor's note: Other than choosing tiddyverse for data wrangling this looks quite useful. I wish the author had focused on data.table.

Shared by Richard McElreath

Science journals ban listing of ChatGPT as co-author on papers
theguardian.com
Some publishers also banning use of bot in preparation of submissions but others see its adoption as inevitable

Shared by Retraction Watch

[Publish less, disrupt more]
universityaffairs.ca
[Canadian academics are prolific, but we must reverse a 60-year trend towards progressively less challenging work.]
Shared by [Retraction Watch]

[Don't Wait Until You're Motivated to Write. Take Small Actions! - Edge for Scholars]

edgeforscholars.org
Take some (small) action. Maybe forgetting about that manuscript is the best way to get that manuscript written :)
Shared by Edge for Scholars

Examples of solid causal inferences from purely observational data
discourse.datamethods.org
I would like to catalog here a few great teaching examples where modern principles of causal inference are used to make solid causality statements from purely observational data. Contributions with b…
Shared by Frank Harrell

Guest Post – AI and Scholarly Publishing: A View from Three Experts
scholarlykitchen.sspnet.org
A recap of a recent SSP webinar on artificial intelligence (AI) and scholarly publishing. How can this set of technologies help or harm scholarly publishing, and what are some current trends? What ar…
Shared by Retraction Watch

Former Harvard researchers lose PNAS paper for reusing data
retractionwatch.com
John Blenis A group of cancer researchers once all based at Harvard have earned a retraction after acknowledging data duplication “errors” in an article published more than eight years ago. The paper…
Shared by Ivan Oransky

J&J subsidiary alleges fraud in paper that linked cosmetic talc with mesothelioma
retractionwatch.com

A key paper linking use of talc-based baby powder to cancer contains fraudulent information, according to a new complaint against an author of the article who has testified on behalf of plaintiffs. A…

Shared by Retraction Watch

Changes since the 1970s (ESP edition) | Statistical Modeling, Causal Inference, and Social Science
statmodeling.stat.columbia.edu

In 1979, the computer scientist Douglas Hofstadter published the book Godel Escher Bach. I've always thought this book to be overrated, but maybe that judgment on my part wasn't fair. Godel Escher Ba…

Shared by Andrew Gelman et al.

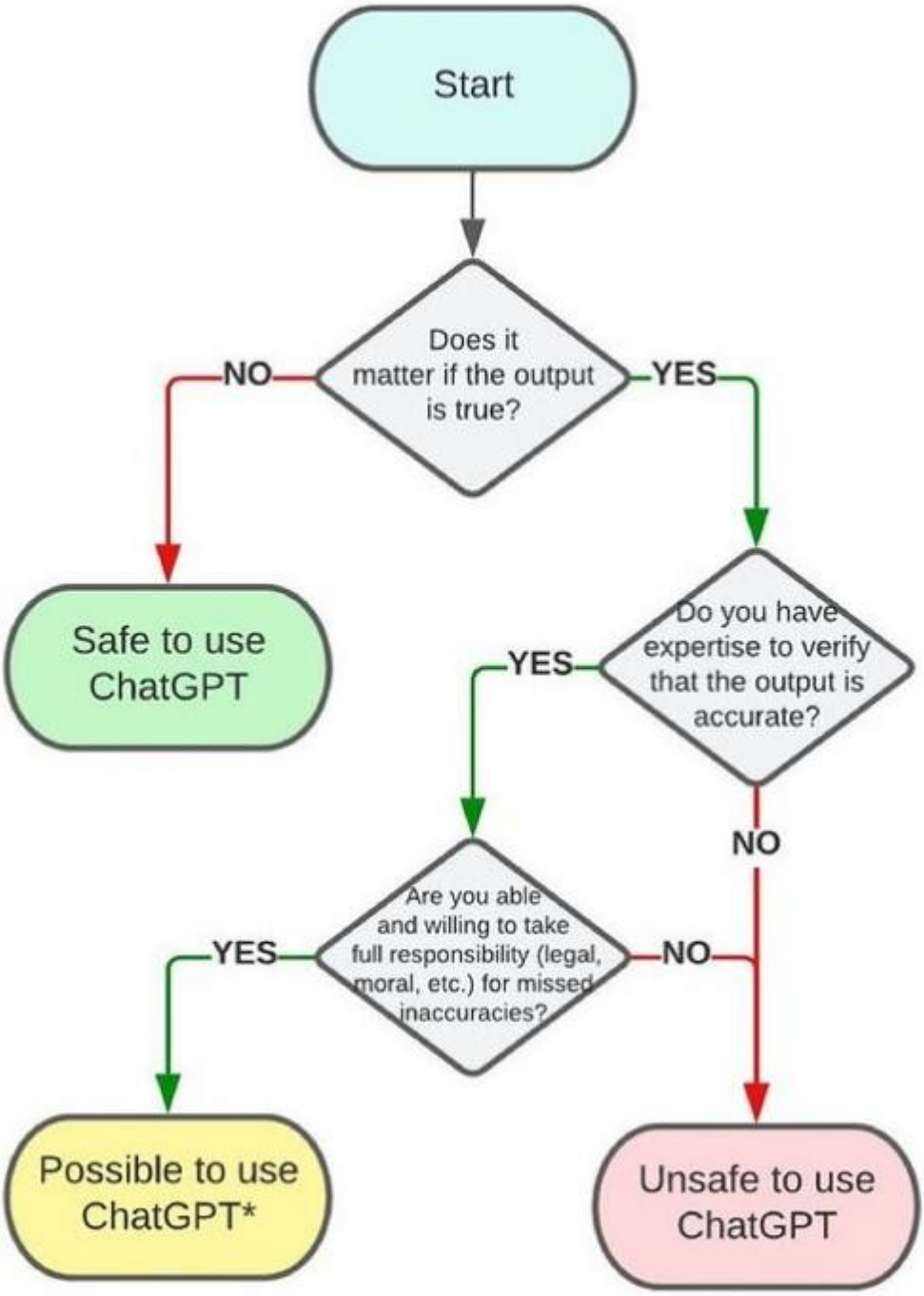
Accuracy of ChatGPT on Statistical Methods
discourse.datamethods.org

OpenAI has created ChatGPT, a "chat generative pre-trained transformer" that has received extensive worldwide attention. ChatGPT has been tested in a variety of situations and has been able to achiev…

Shared by Frank Harrell

Is it safe to use ChatGPT for your task?

Aleksandr Tiulkanov | January 19, 2023



Are universities doing enough to address academic misconduct in research?
universityaffairs.ca
The search for a systemic response to a growing concern.
Editor's note: Universities wrongly think of scientific misconduct as "one-offs" instead of the systemic problem it really is.
Shared by Retraction Watch

Google didn’t break your brain: a history of distraction | Aeon Essays
aeon.co
Worried that technology is ‘breaking your brain’? Fears about attention spans and focus are as old as writing itself
Shared by Aeon+Psyche

[He Was a World-Renowned Cancer Researcher. Now He's Collecting Unemployment.](#)
thefp.com
[Behind the fall of David Sabatini, 'one of the greatest scientists' of his generation.](#)
Shared by [Vinay Prasad MD MPH](#)

[Performance of ChatGPT on USMLE: Potential for AI-Assisted Medical Education Using Large Language Models]
[medrxiv.org]
[medRxiv - The Preprint Server for Health Sciences]
Shared by [Harlan Krumholz]

[Putting ChatGPT's Medical Advice to the (Turing) Test]
[medrxiv.org]

medRxiv - The Preprint Server for Health Sciences
Shared by C. Michael Gibson MD

Why is Java So Weird?
unchartedterritories.tomaspueyo.com
Java has a larger population than Russia. It's smaller than Japan but has a bigger population. It's nearly 8x denser than neighboring Sumatra. Why? Is it volcanoes? Trade? Immigration? colonization?
Editor's note: This is a wonderful graphical geographical history
Shared by Tomas Pueyo

Gestational age at birth and cognitive outcomes in adolescence: population based full sibling cohort study
bmj.com

Objective To investigate the association between gestational age at birth and cognitive outcomes in adolescence. Design Nationwide population based full sibling cohort study. Setting Denmark. Partici…
Editor's note: Categorization of gestational age almost ruined this study
Shared by Sam Finlayson

14 January 2023
380:1-42 No 8366 | ISSN 1759-2151

thebmj

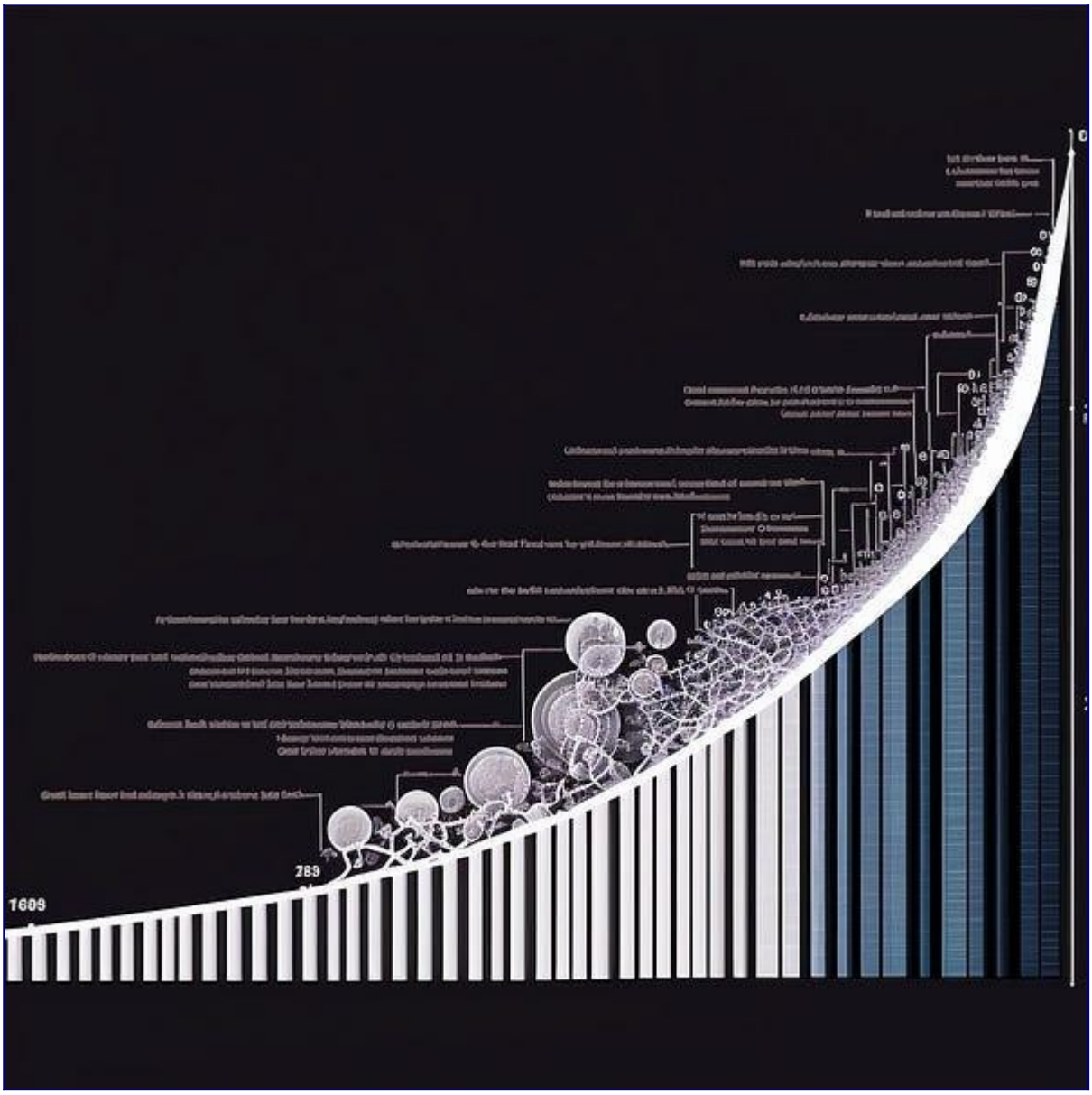
BMJ APPEAL 2022-23

It's not the NHS needing radical reform p 11
Waiting lists: what's really going on p 14
How to manage chronic migraine p 31
Tips on dealing with a racist patient p 40
1 CPD hour in the education pages

The cost of living crisis

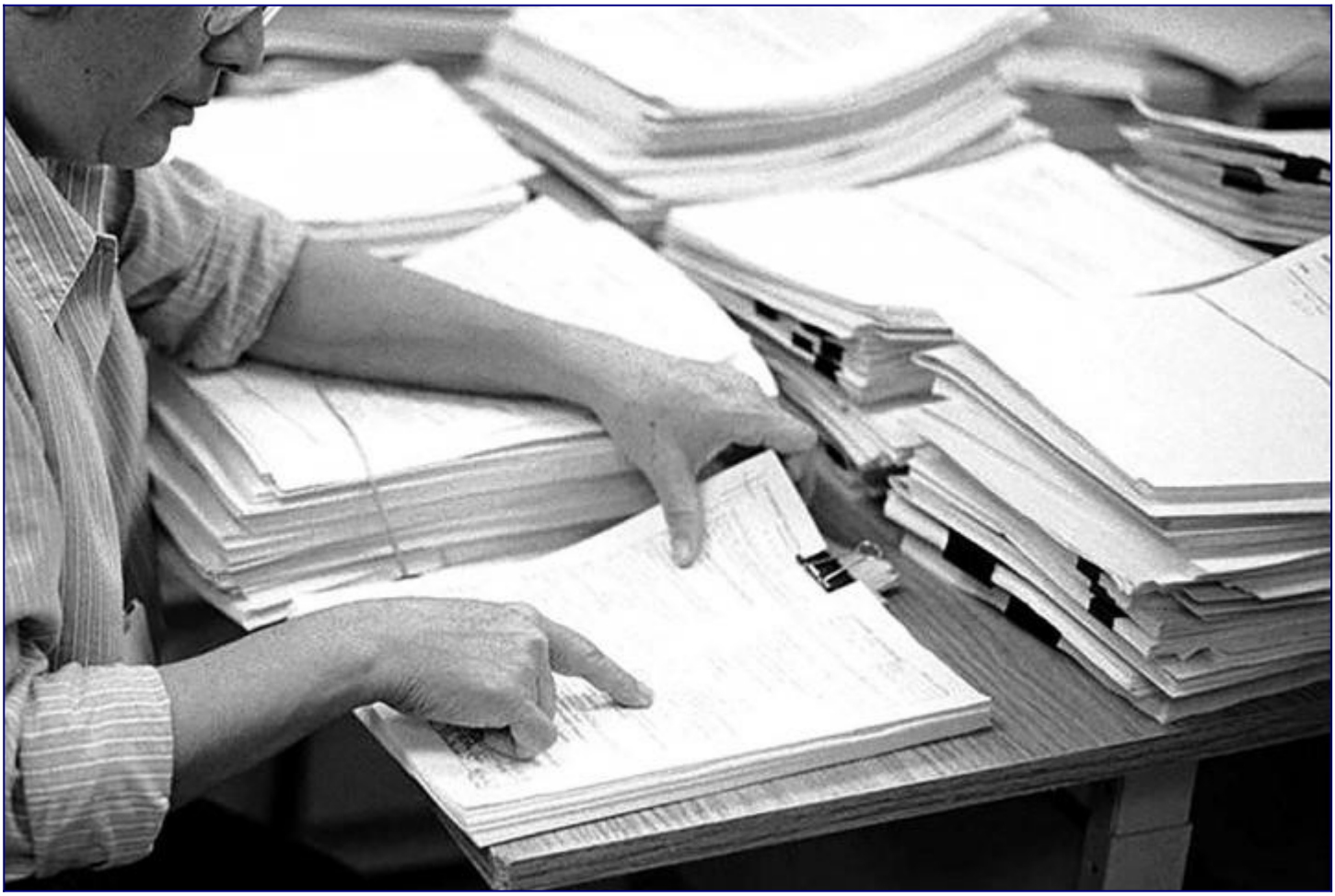
Harming

Becoming strange in the Long Singularity
oneusefulthing.substack.com
A sudden increase in AI capability suggests a weirder world in our near future
Shared by Ethan Mollick



A systematic review of randomisation method use in RCTs and association of trial design characteristics with method selection
bmcmedresmethodol.biomedcentral.com
When conducting a randomised controlled trial, there exist many different methods to allocate participants, and a vast array of evidence-based opinions on which methods are the most effective at doin…
Shared by Marion Campbell

[I Wrote a Viral Screed Against Peer Review. I Got Some Emails.](#)
slate.com
[I said what I said.](#)
Shared by [Retraction Watch](#)



[Yale researchers discover loophole in FDA medical device regulation](#)
yaledailynews.com

Researchers at the Yale School of Medicine and Harvard Medical School found that a loophole in existing regulation has allowed manufacturers to acquire U.S. Food and Drug Administration approval for …

Shared by Harlan Krumholz

Bayesian statistics and machine learning: How do they differ?

statmodeling.stat.columbia.edu

A researcher who would like to remain anonymous writes: My colleagues and I are disagreeing on the differentiation between machine learning and Bayesian statistical approaches. I find them philosophi…

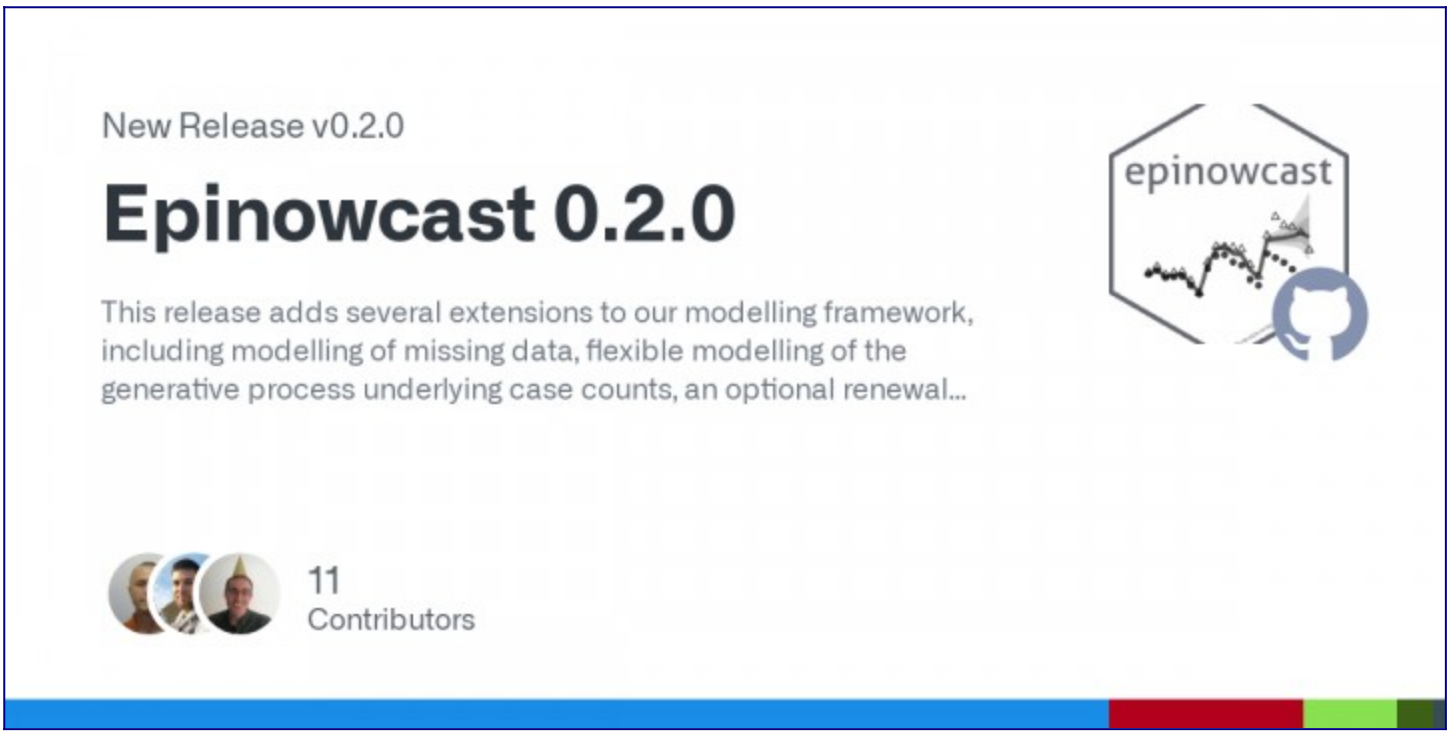
Shared by Andrew Gelman et al.

Release Epinowcast 0.2.0 · epinowcast/epinowcast

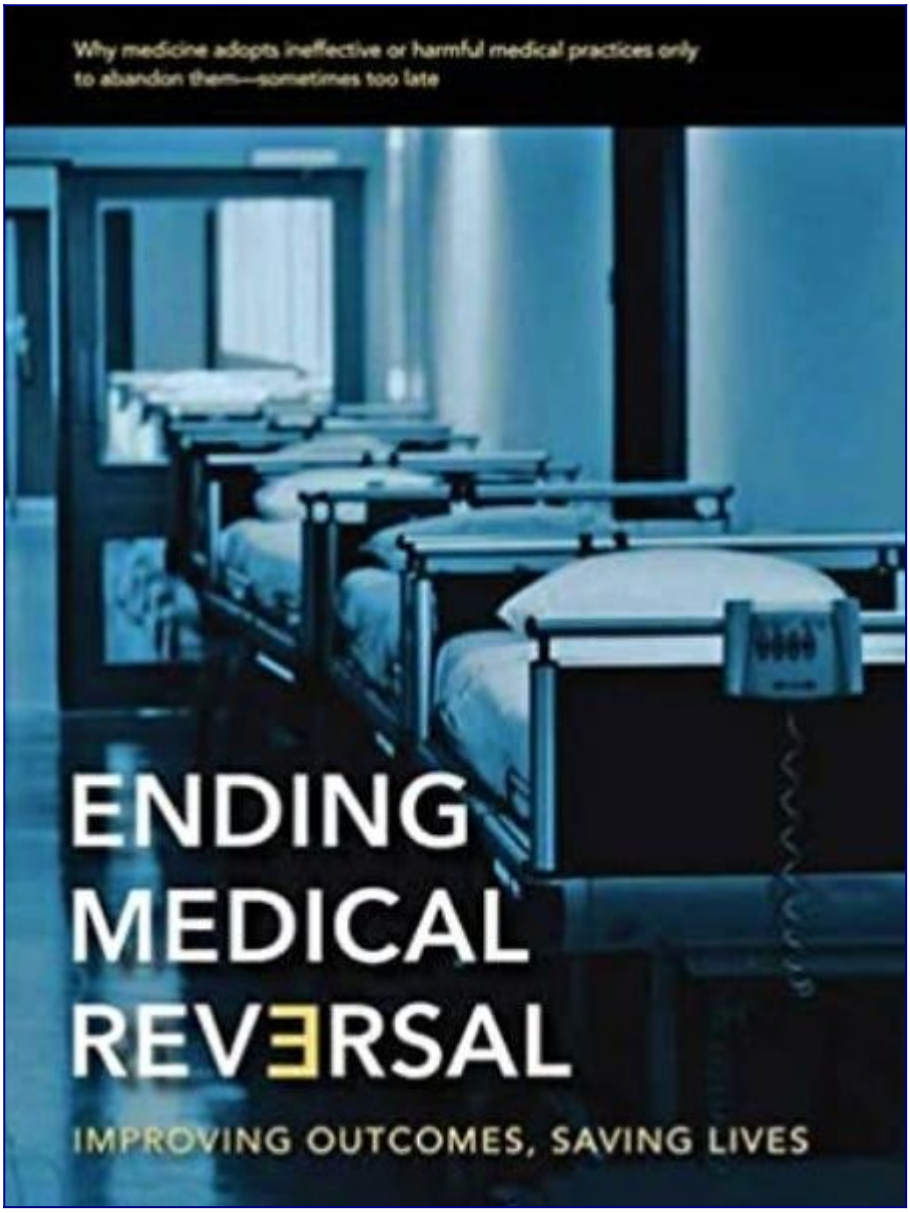
github.com

Tools to enable flexible and efficient hierarchical nowcasting of epidemiological time-series using a semi-mechanistic Bayesian model with support for a range of reporting and generative processes. N…

Shared by Darren Dahly https://mastodon.social/@statsepi



[In Defense of the Randomized Trial]
[sensiblemed.substack.com]
[It took millennia to develop, yet we still can't embrace it]
Shared by [Vinay Prasad MD MPH]



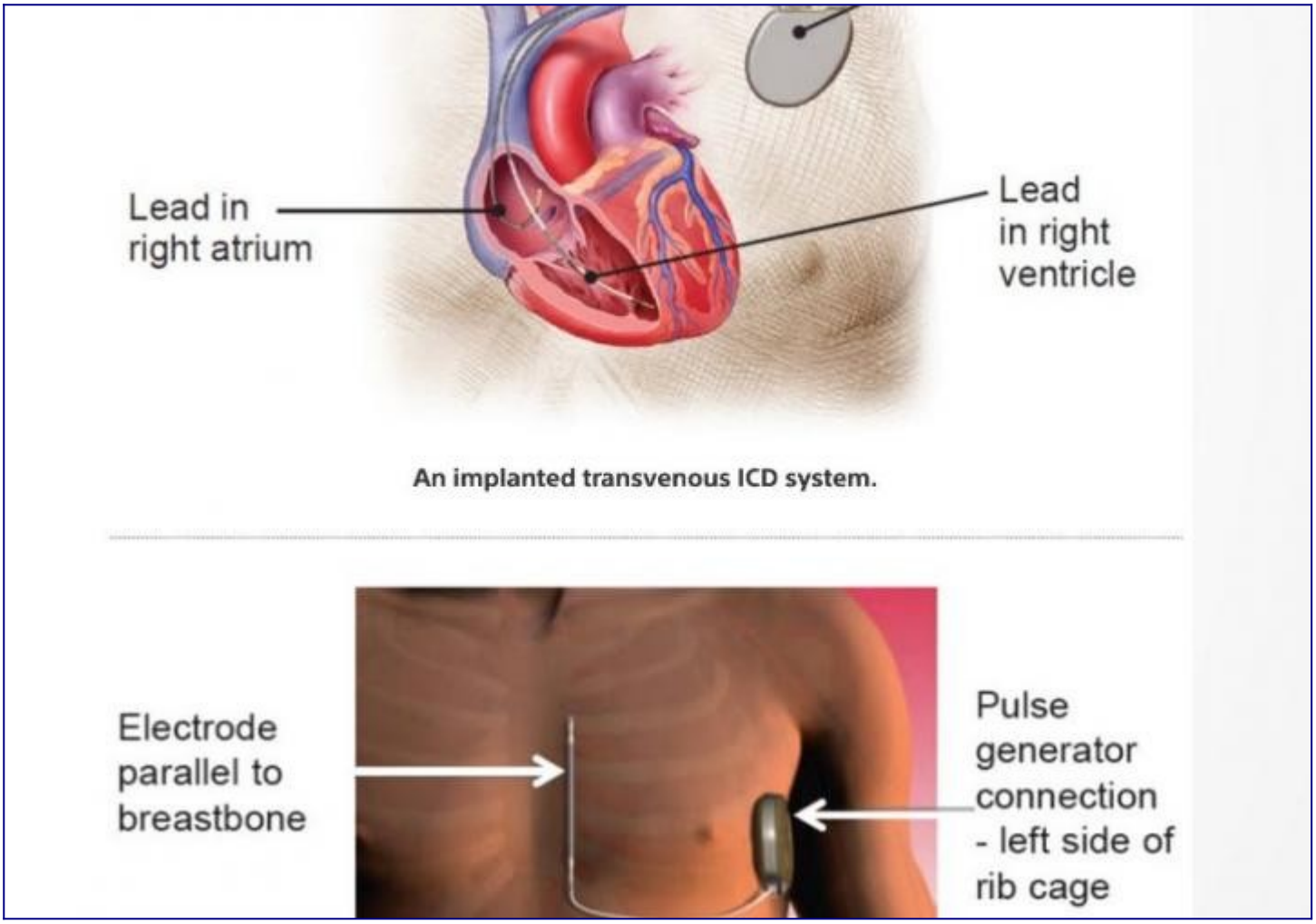
Oops. After ten years and 1,000 studies, epigeneticists uncover trouble in their tool box
bcm.edu
Twenty years ago, after the human genome was first sequenced, geneticists began conducting large genome-wide association studies to identify genomic regions...
Shared by Retraction Watch

Finding Bias Before the First Patient is Enrolled

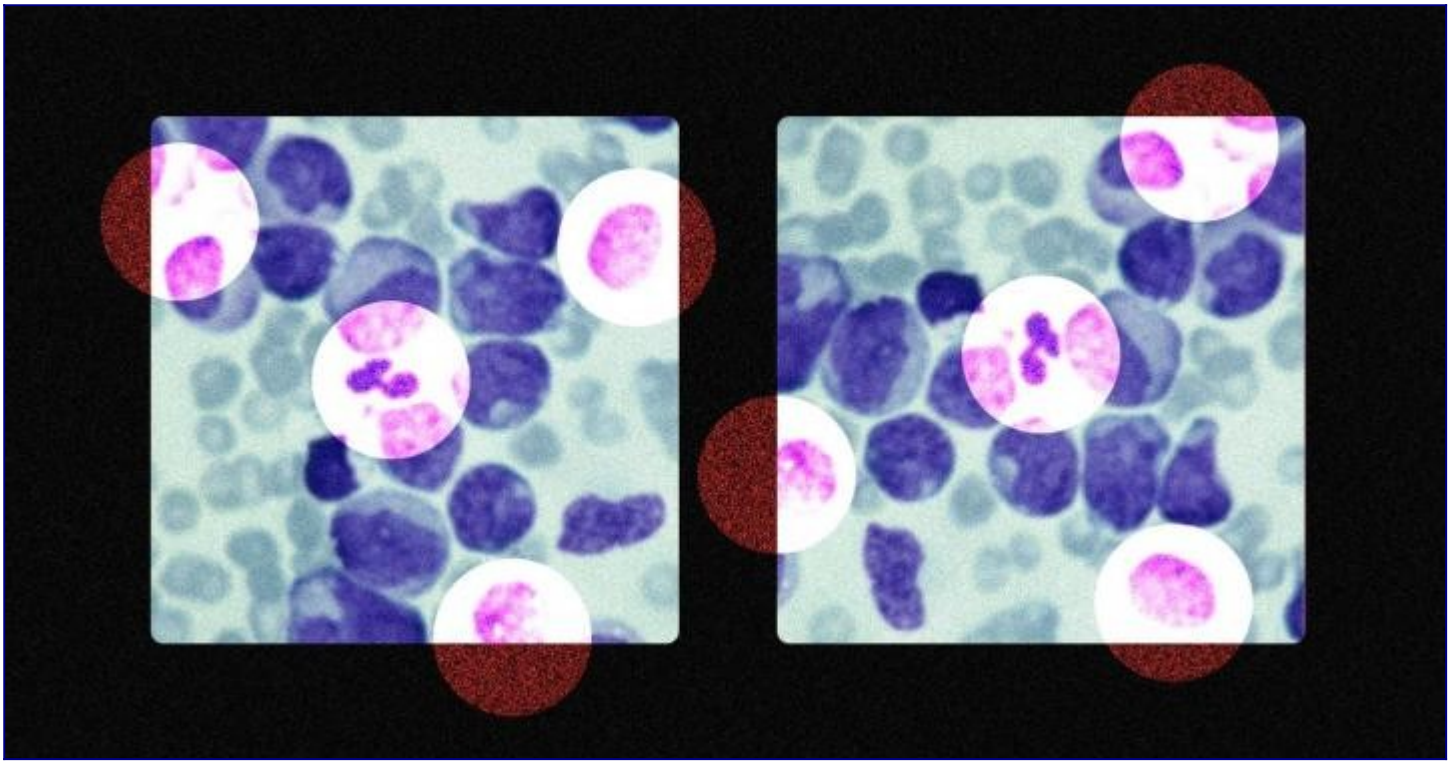
sensiblemed.substack.com

Last Monday, we considered potential bias in data analysis. Today, the Study of the Week considers another source of bias—trial design.

Shared by John Mandrola, MD



[Scientific Fraud Is Slippery to Catch—but Easier to Combat]
[wired.com]
[Fakery spans “beautified” data, photoshopped images, and “paper mills.” Experts and institutions are employing tools to spot deceptive research and mitigate its reach.]
Shared by [Retraction Watch]



[On the 12th Day of Christmas, a Statistician Sent to Me . . .]
bmj.com
The BMJ’s statistical editors relish a quiet Christmas, so make their wish come true and pay attention to the list of common statistical faux pas presented here by Riley and colleagues The weeks lead…
Shared by Gary Collins

14 January 2023
380:1-42 No 8366 | ISSN 1759-2151

thebmj

BMJ APPEAL 2022-23

It's not the NHS needing radical reform p 11
Waiting lists: what's really going on p 14
How to manage chronic migraine p 31
Tips on dealing with a racist patient p 40
1 CPD hour in the education pages

The cost of living crisis

Harming

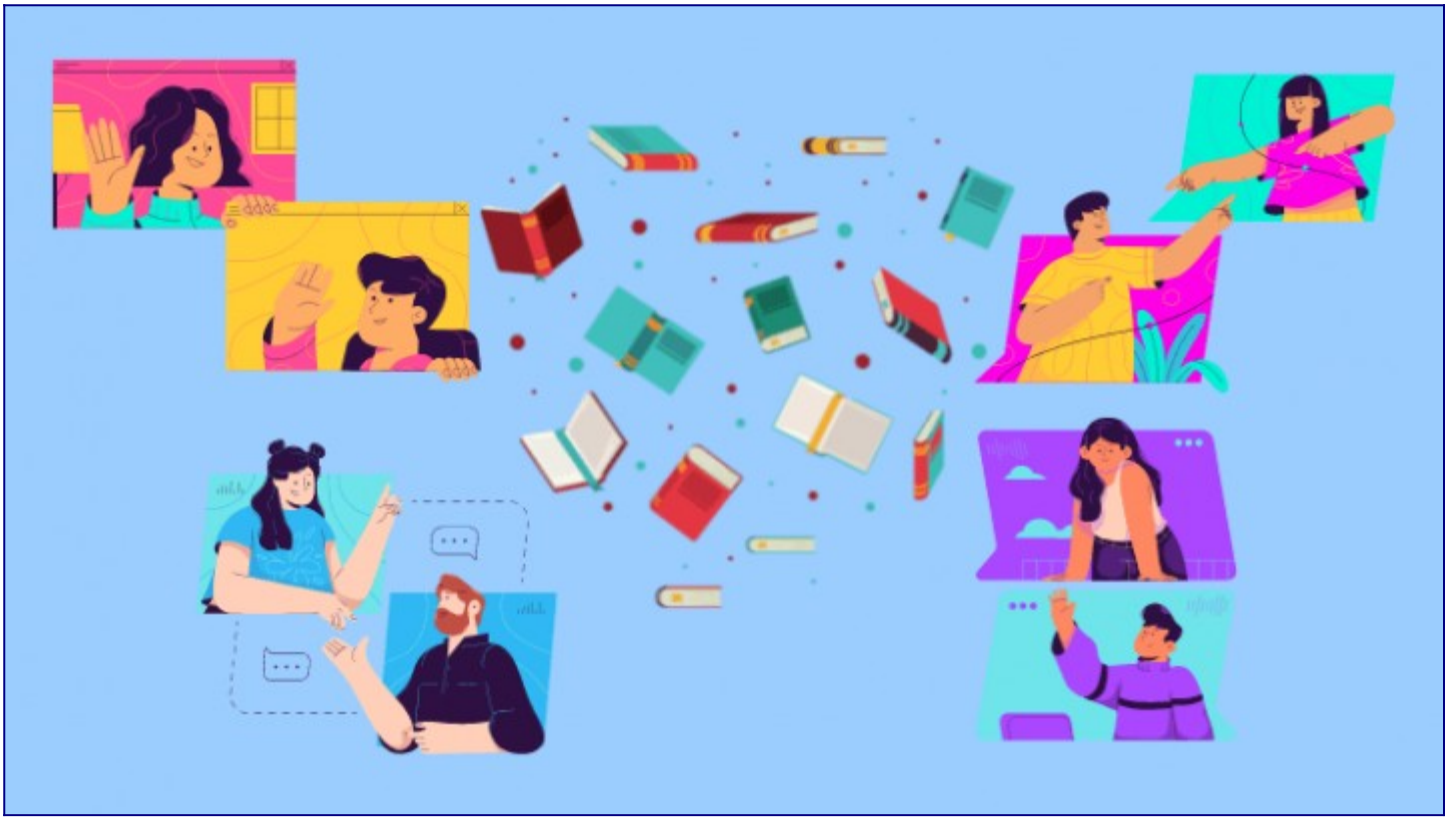
We will never be able to live on another planet. Here’s why | Aeon Essays
aeon.co
The scientific evidence is clear: the only celestial body that can support us is the one we evolved with. Here’s why
Shared by Aeon+Psyche

Can We Please Stop Asking if Crossword Puzzles Prevent Dementia?
sensiblemed.substack.com
In the pantheon of lay press health articles that make me crazy, articles about how brain exercises can ward off dementia lord over all others.
Shared by John Mandrola, MD

Changelog
sebkrantz.github.io
Full data.table support, including reference semantics (set*, :=)!! There is some complex C-level programming behind data.table's operations by reference. Notably, additional (hidden) column pointers…
Shared by collapse

Minimum Sample Size for Developing a Multivariable Prediction Model using Multinomial Logistic Regression
arxiv.org
Multinomial logistic regression models allow one to predict the risk of a categorical outcome with more than 2 categories. When developing such a model, researchers should ensure the number of partic…
Shared by Richard Riley (R^2)

[Do journals need societies, and do societies need journals?]
blogs.lse.ac.uk
Aileen Fyfe asks, do we need to rethink the structure of the learned societies in an increasingly globalised and open access publishing landscape?
Shared by Retraction Watch



[How We Stopped Unintentionally Harming ICU Patients](#)
[ihi.org](#)
[Wesley Ely, MD, remembers some of the patients and professionals who helped change critical care to prevent patients from being irreparably harmed by the care meant to save their lives.](#)
Shared by [WesElyMD](#)

Institute *for* Healthcare Improvement

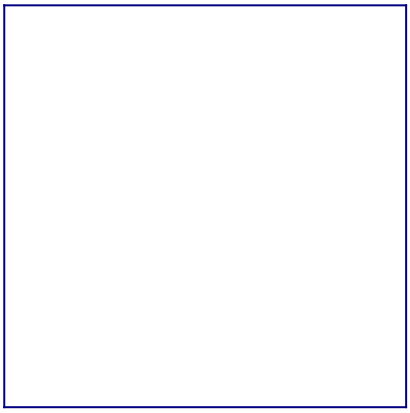
The 8 Most Important Statistical Ideas of the Past 50 Years
medium.com
In this article, we will highlight some of the most important statistical ideas that have emerged over the last half-century and explain how they have helped to shape the field of statistics as we…
Shared by Benedict Neo 梁耀恩

DeepMind CEO Demis Hassabis Urges Caution on AI
time.com
DeepMind brought artificial intelligence mainstream. Now its CEO Demis Hassabis is issuing a warning
Shared by Judea Pearl

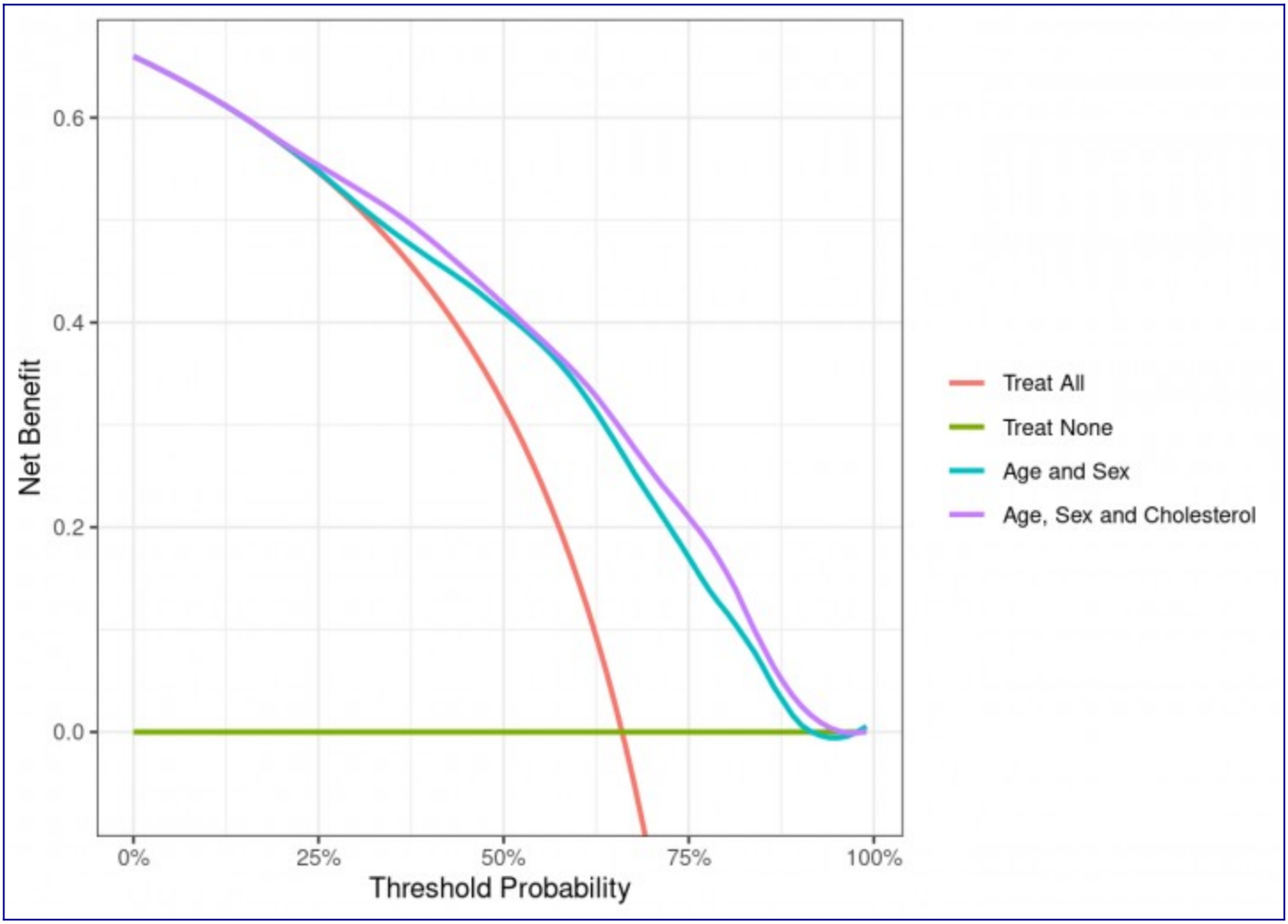
'We Can Save Lives': Damar Hamlin's Resuscitation Is a Teachable Moment
medpagetoday.com
Harlan Krumholz, MD, on the importance of CPR and medical advocacy
Shared by Harlan Krumholz

Multiple imputation for missing baseline covariates in discrete time survival analysis

thestatsgeek.com
A while ago I got involved in a project led by Anna-Carolina Haensch and Bernd Weiß investigating multiple imputation methods for baseline covariates in discrete time survival analysis. The work ha…
Shared by Jonathan Bartlett



Statistical Thinking - Decision curve analysis for quantifying the additional benefit of a new marker
fharrell.com
This article examines the benefits of decision curve analysis for assessing model performance when adding a new marker to an existing model. Decision curve analysis provides a clinically interpretabl…
Shared by 5’UTR



[Modern modelling techniques are data hungry: a simulation study for predicting dichotomous endpoints]
discourse.datamethods.org
I would like to discuss this article by van der Ploeg, which I found reading through some great materials of Professor Harrell yesterday. I have some issues with, and some doubts about, it. I do not …

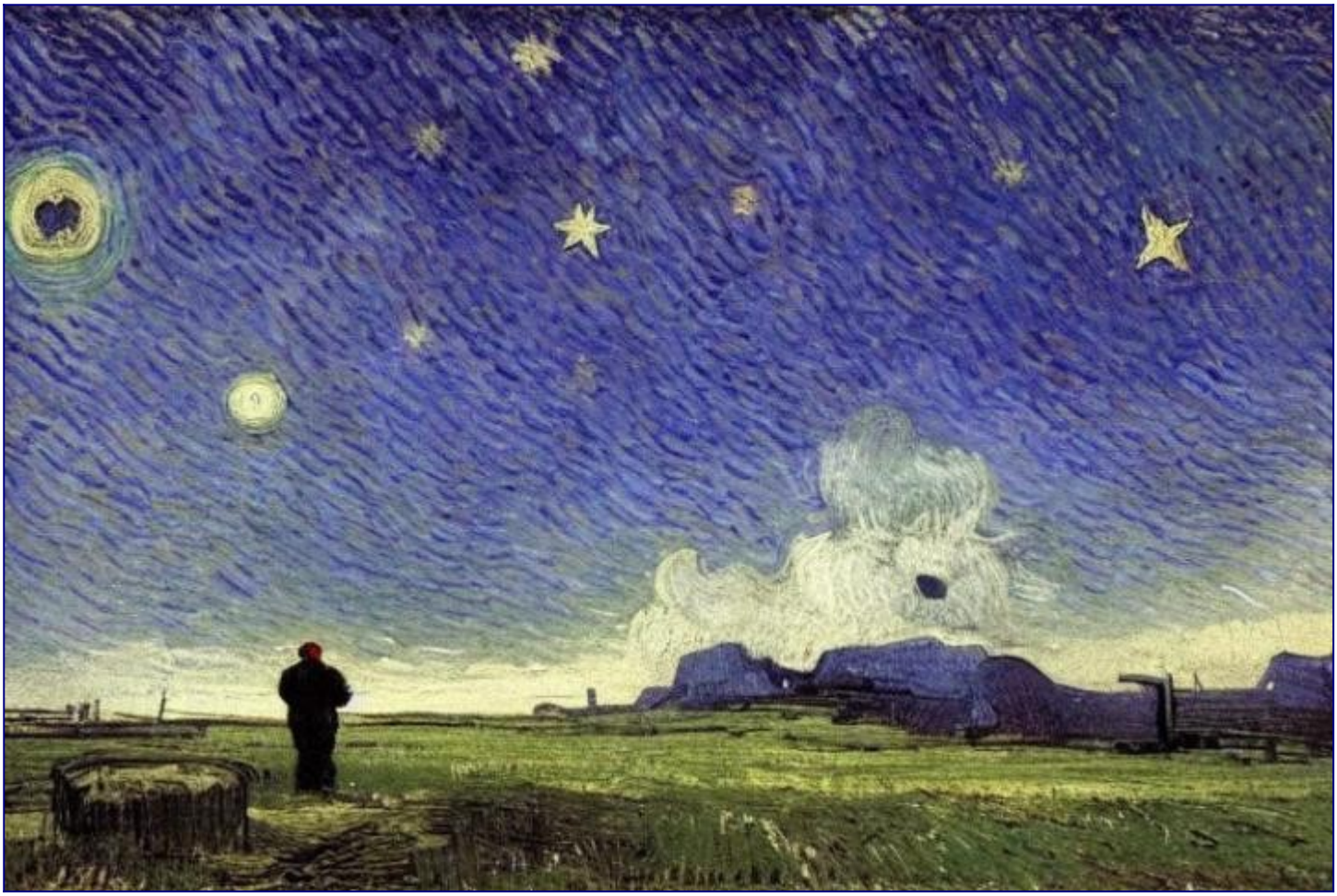
[The Power of Strong Collaborations]
edgeforscholars.org
The right collaborations can accelerate your career - choose wisely!
Shared by Edge for Scholars

INTERVIEW: Geneticist at a Top 5 University

goodscience.substack.com

At the Good Science Project, we've been featuring regular interviews with practicing scientists. One of the chief difficulties in doing so is that many scientists aren't willing to be quoted publicly…

Shared by Harlan Krumholz



Causal inference is a mindset
medium.com
Causal inference from observational data is first and foremost a mindset. Identification is what makes an estimation into causation rather than a simple association.
Shared by PPMI (Progressing Prisoners Maintaining Innocence)
Alcohol: Good or Bad? Why Reductionist Science Can't Answer a Complex Social, Nutritional Question
vinayprasadmdmph.substack.com
If you are reading about NADPH and neutral circuits you aren't even close to thinking about it correctly
Shared by Vinay Prasad MD MPH

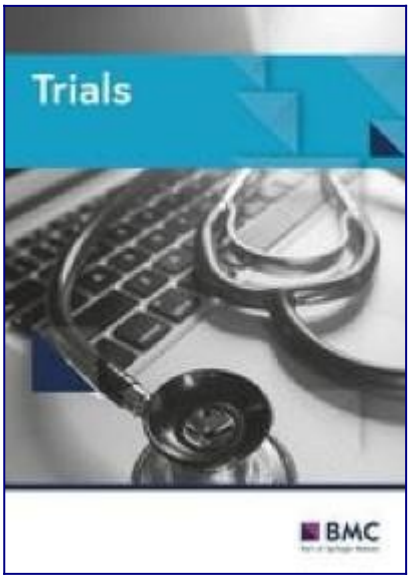
Justification of sample size for the generation of a prediction model
discourse.datamethods.org
Dear all, I plan a retrospective analysis of patients who visited our outpatient clinic. The data will be extracted from our hospital information system. One specific research question is to generate…

Target Trial Emulation to Estimate the Association of Corticosteroids With COVID-19 Mortality
jamanetwork.com
This cohort study compares methods for statistical inference using observational data in replicating findings from a meta-analysis of randomized clinical trials

Shared by Dan Malinsky

Planning a method for covariate adjustment in individually randomised trials: a practical guide
trialsjournal.biomedcentral.com
It has long been advised to account for baseline covariates in the analysis of confirmatory randomised trials, with the main statistical justifications being that this increases power and, when a ran…

Shared by Maarten van Smeden



Validation of prediction models in the presence of competing risks: a guide through modern methods
bmj.com
Thorough validation is pivotal for any prediction model before it can be advocated for use in medical practice. For time-to-event outcomes such as breast cancer recurrence, death from other causes is…

Shared by Maarten van Smeden

24 December 2022 - 7 January 2023
379 463 504 No 8355 ISSN 1759-2151

thebmj

THE BMJ APPEAL 2022-23

In praise of our unsung colleagues p 472
Social media links to cancer myths p 480
She-Hulk: incredible transfusion tale p 482
Alcohol, elite athletes, and health p 487
A crossword to flex diagnosis skills p 504

harm of class imbalance corrections for risk prediction models: illustration and simulation using logistic regression
academic.oup.com
AbstractObjective. Methods to correct class imbalance (imbalance between the frequency of outcome events and nonevents) are receiving increasing interest for de
Shared by Maarten van Smeden
Net Reclassification Index Statistics Do Not Help Assess New Risk Models
pubs.rsna.org
When evaluating a new risk factor for disease (eg, a measurement from imaging studies), many investigators examine its value above and beyond existing biomarkers and risk factors. They compare the ...
Shared by Maarten van Smeden
The rise and fall of peer review
experimentalhistory.substack.com
Why the greatest scientific experiment in history failed, and why that's a great thing
Shared by Retraction Watch

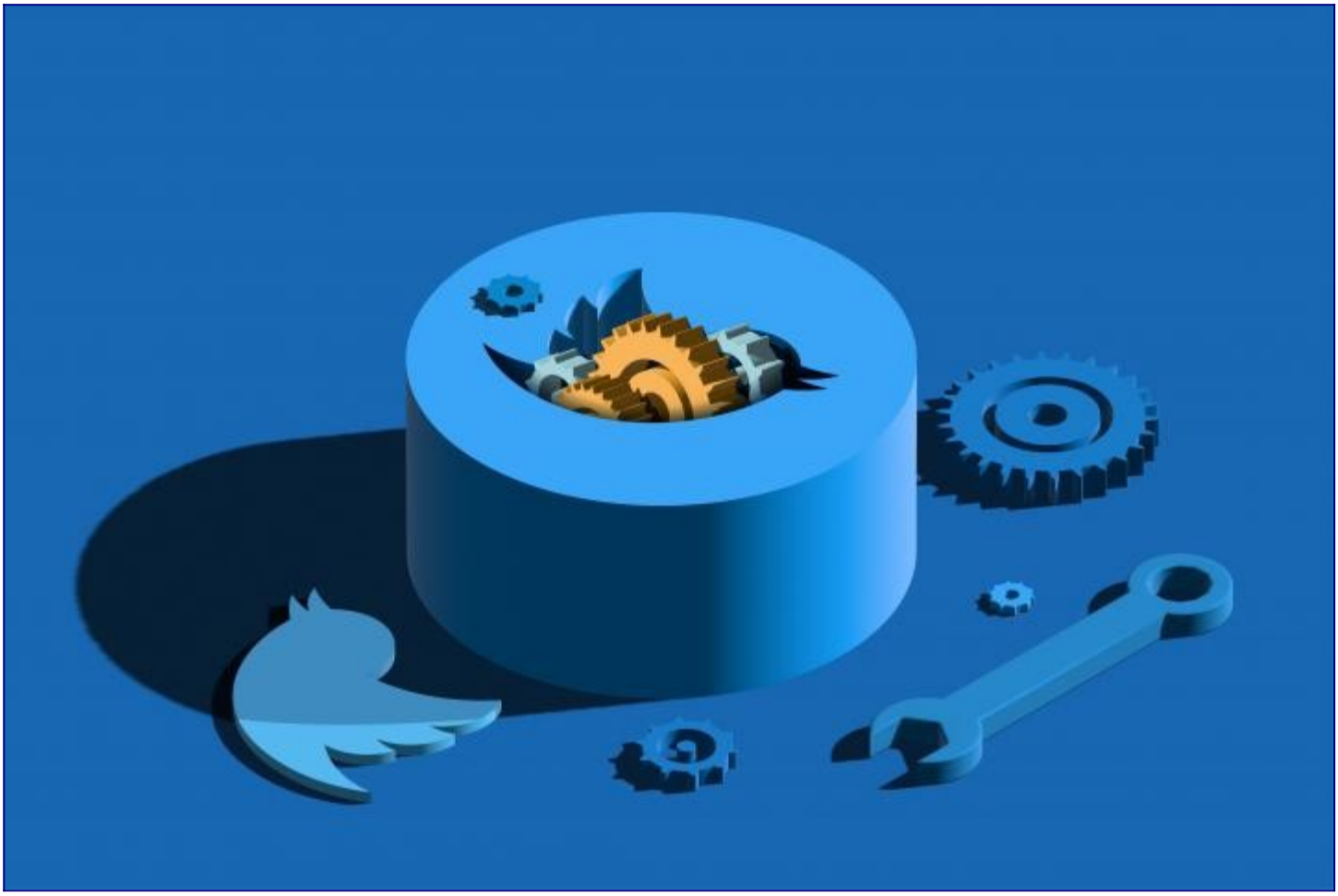
Assessing performance and clinical usefulness in prediction models with survival outcomes: practical guidance for Cox proportional hazards models
medrxiv.org
medRxiv - The Preprint Server for Health Sciences

Shared by Rod Hayward

medRxiv
THE PREPRINT SERVER FOR HEALTH SCIENCES

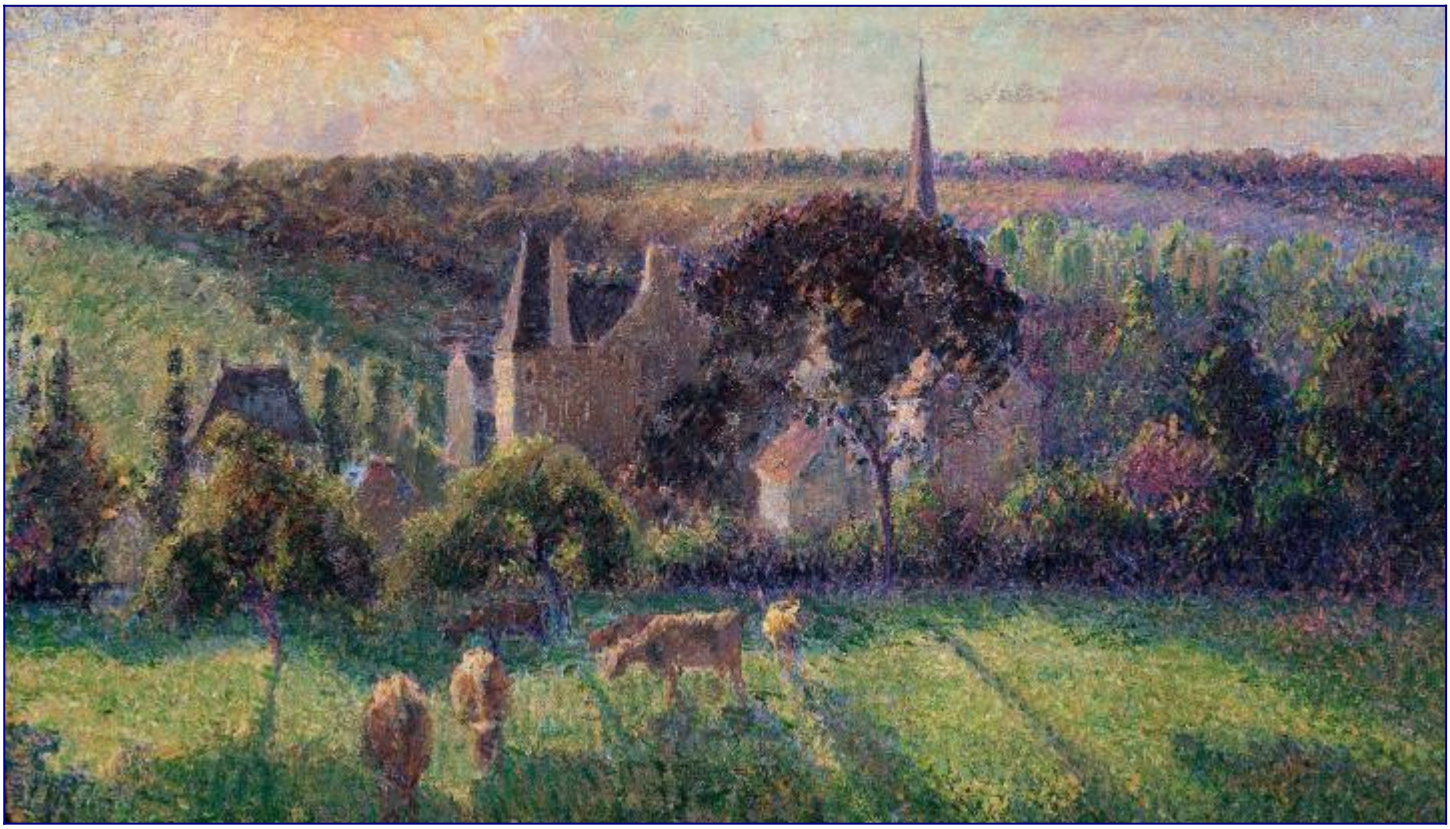
Twitter brings Elon Musk’s genius reputation crashing down to earth
washingtonpost.com
Musk went down conspiracy rabbit holes and sank Tesla's stock with his behavior. And he was confronted with a chorus of boos in the cradle of the tech industry.

Shared by Boback Ziaeian



Glossary of Statistical Terms

hbiostat.org
Vanderbilt University School of Medicine Department of Biostatistics adjusting or controlling for a variable: Assessing the effect of one variable while accounting for the effect of another (confound…
Shared by Frank Harrell
Why it took us thousands of years to see the colour violet
psyche.co
The colour violet was largely missing from art before the Impressionists, and is seen differently by different cultures. Why?
Shared by Aeon+Psyche



On the 12th Day of Christmas, a Statistician Sent to Me . . .
bmj.com
The BMJ's statistical editors relish a quiet Christmas, so make their wish come true and pay attention to the list of common statistical faux pas presented here by Riley and colleagues The weeks lead…
Shared by Richard Riley (R^2)

10-17 December 2022
379 421-462 No 8364 | ISSN 1759-2151

thebmj

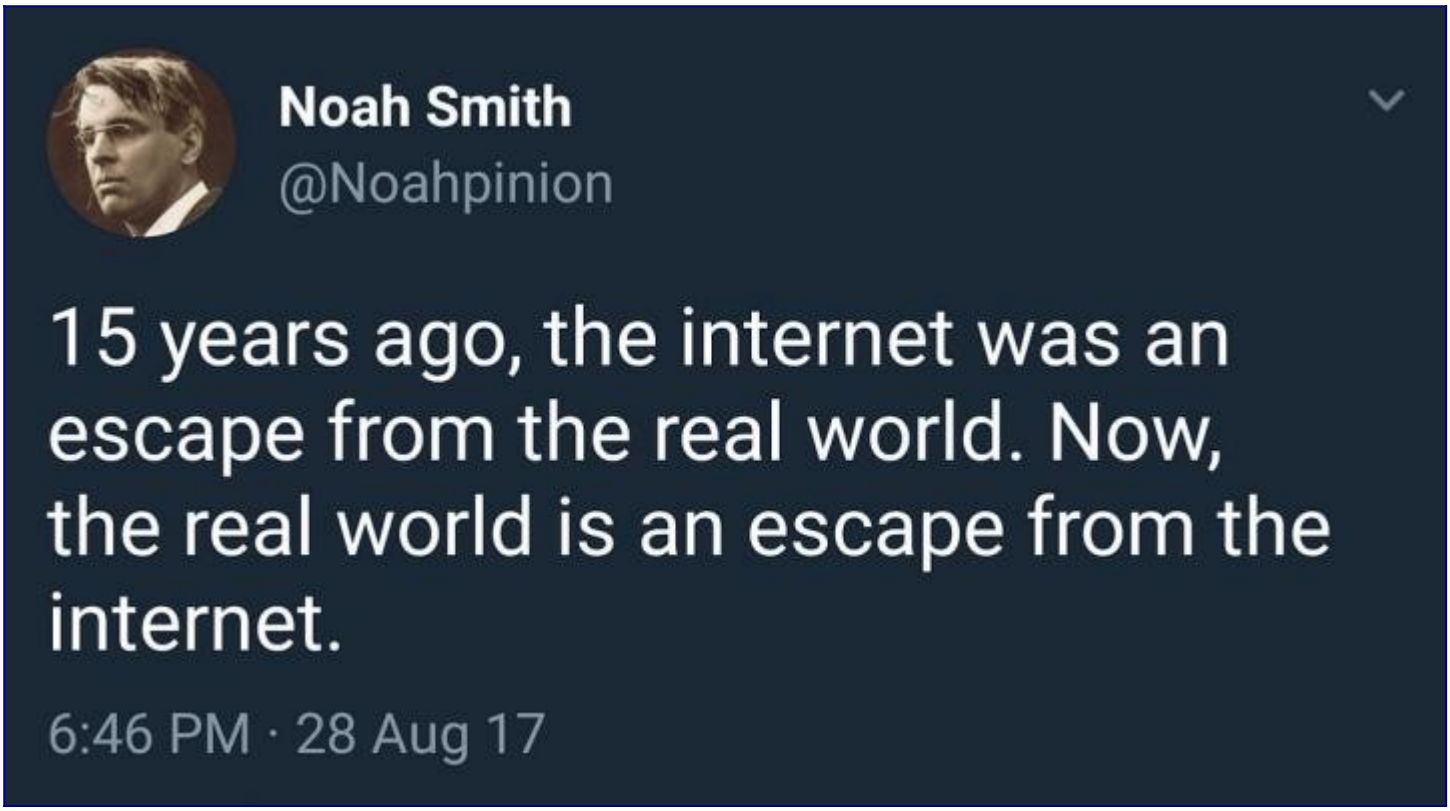
BMJ APPEAL 2022-23

Pension changes "too little too late" p 421
Caring for world's covid orphans p 432
AstraZenca's vaccine equity lesson p 434
Cost of NICE's switch to innovation p 445
1 CPD hour in the education pages

ENDOMETRIOSIS
Presentation and management

Targeted validation: validating clinical prediction models in their intended population and setting
diagnprognres.biomedcentral.com
Clinical prediction models must be appropriately validated before they can be used. While validation studies are sometimes carefully designed to match an intended population/setting of the model, it …
Shared by Gary Collins

The internet wants to be fragmented
noahpinion.substack.com
Throwing the whole world into a single room together doesn't work.
Shared by Ethan Mollick



Boston University School of Public Health Stops Using Twitter

sensiblemed.substack.com
A justified choice or performative theater? And does it advance public health?
Shared by Vinay Prasad MD MPH

"Do Simpler Machine Learning Models Exist and How Can We Find Them?"
statmodeling.stat.columbia.edu
The above is the title of a talk by computer scientist Cynthia Rudin. here's the abstract: While the trend in machine learning has tended towards building more complicated (black box) models, such mo…
Shared by Andrew Gelman et al.
The rise and fall of peer review
experimentalhistory.substack.com
Why the greatest scientific experiment in history failed, and why that's a great thing
Shared by Retraction Watch

[Methodological guidance for the evaluation and updating of clinical prediction models: a systematic review]
bmcmedresmethodol.biomedcentral.com
Clinical prediction models are often not evaluated properly in specific settings or updated, for instance, with information from new markers. These key steps are needed such that models are fit for p…
Shared by Giuseppe Biondi-Zoccai

A probability isn’t just a number; it’s part of a network of conditional statements
statmodeling.stat.columbia.edu
This came up awhile ago in comments, when Justin asked: Is p(aliens exist on Neptune that can rap battle) = .137 valid “probability” just because it satisfies mathematical axioms? Is p(aliens exist o…
Shared by Andrew Gelman et al.

The cost of dichotomising continuous variables
bmj.com
Measurements of continuous variables are made in all branches of medicine, aiding in the diagnosis and treatment of patients. In clinical practice it is helpful to label individuals as having or not …
Shared by Maarten van Smeden

10-17 December 2022
379 421-462 No 8364 | ISSN 1759-2151

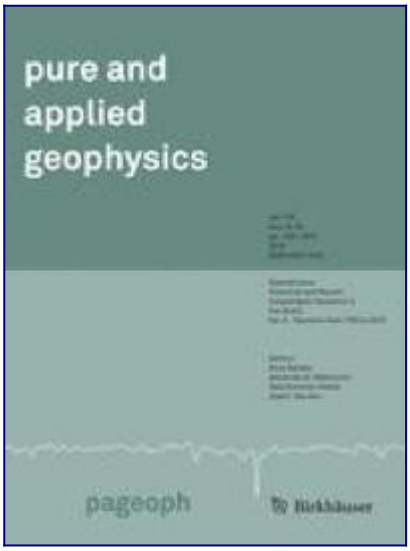
thebmj

BMJ APPEAL 2022-23

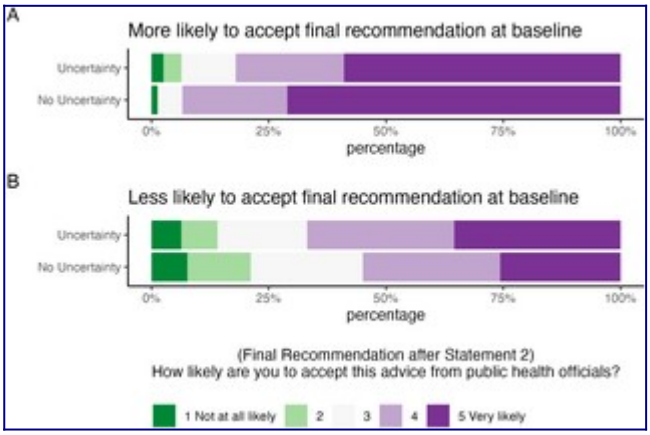
Pension changes "too little too late" p 421
Caring for world's covid orphans p 432
AstraZenca's vaccine equity lesson p 434
Cost of NICE's switch to innovation p 445
1 CPD hour in the education pages

ENDOMETRIOSIS
Presentation and management

Pay No Attention to the Model Behind the Curtain
link.springer.com
Many widely used models amount to an elaborate means of making up numbers—but once a number has been produced, it tends to be taken seriously and its
Shared by Sander Greenland



Randomized controlled trial: Quantifying the impact of disclosing uncertainty on adherence to hypothetical health recommendations
journals.plos.org
We conducted a randomized controlled trial to assess whether disclosing elements of uncertainty in an initial public health statement will change the likelihood that participants will accept new, dif…
Shared by Lucy D'Agostino McGowan



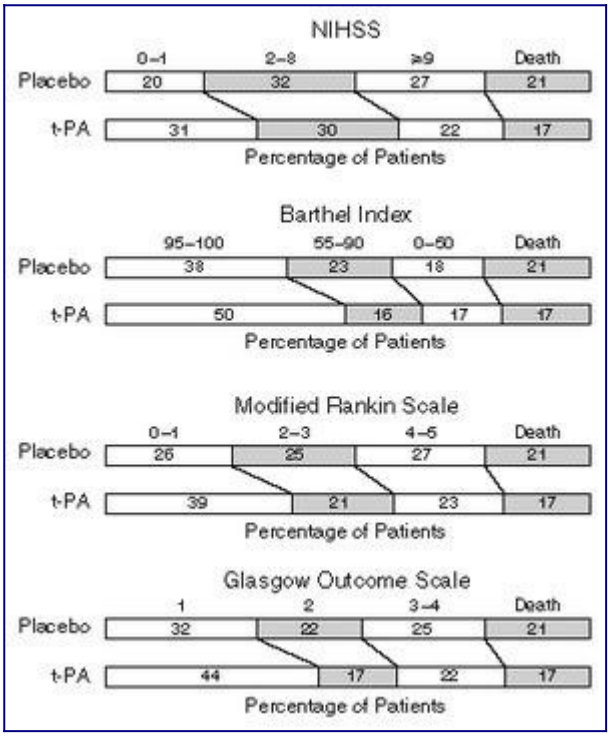
Data Science Pedagogy | Stats + Stories Episode 152 — Stats + Stories
statsandstories.net
In the past, Introduction to Statistics classes spent a lot of time covering distribution tables, teaching students to run stats by hand and focusing on statistical procedures. However, educators…
Shared by Stats + Stories

Social Resilience in Online Communities: The Autopsy of Friendster
arxiv.org
We empirically analyze five online communities: Friendster, Livejournal, Facebook, Orkut, Myspace, to identify causes for the decline of social networks. We define social resilience as the ability of…
Shared by Ethan Mollick

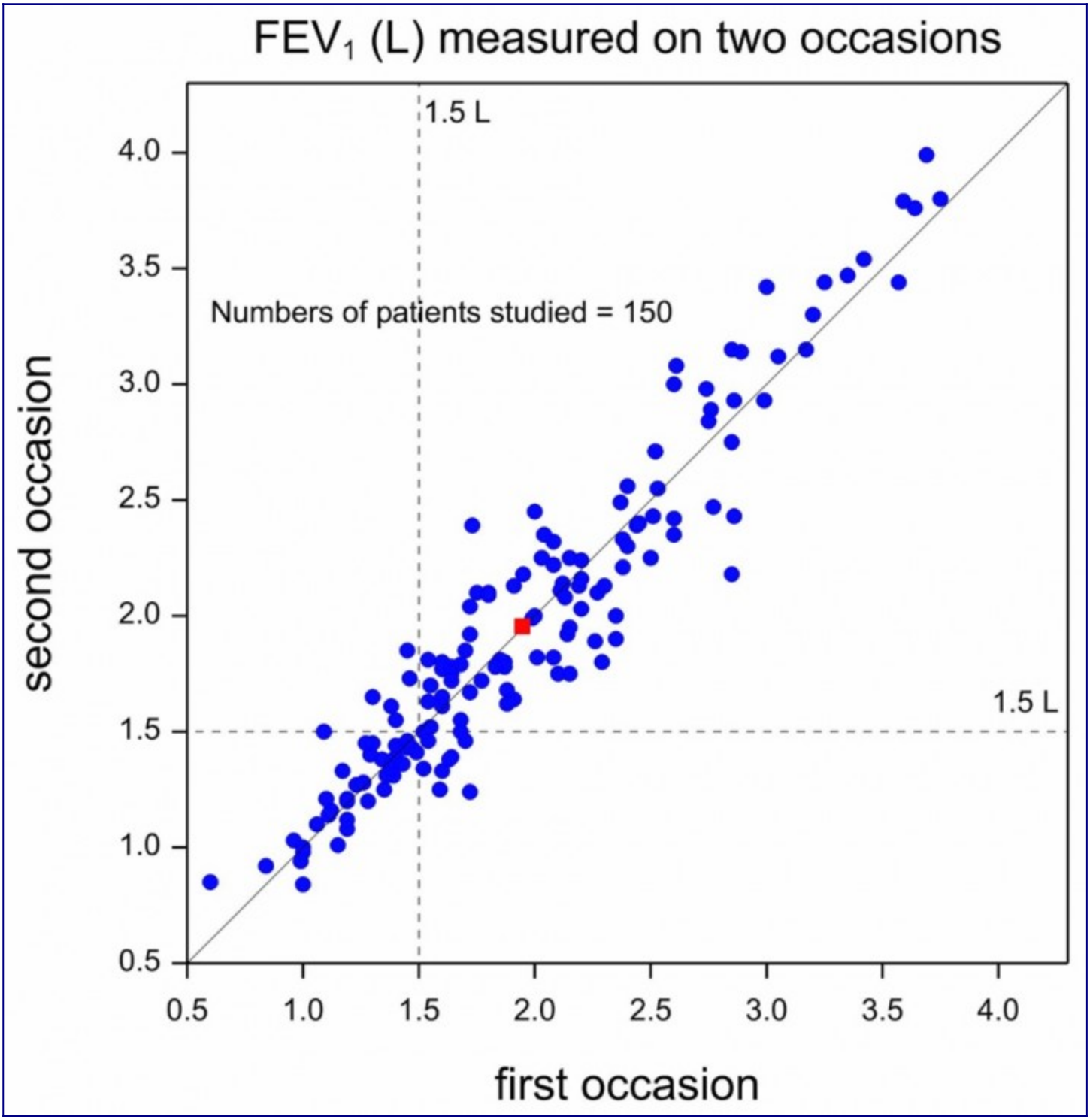
[The Clinical and Statistical Saga of Thrombolysis in Acute Ischemic Stroke](#)

[discourse.datamethods.org](#)

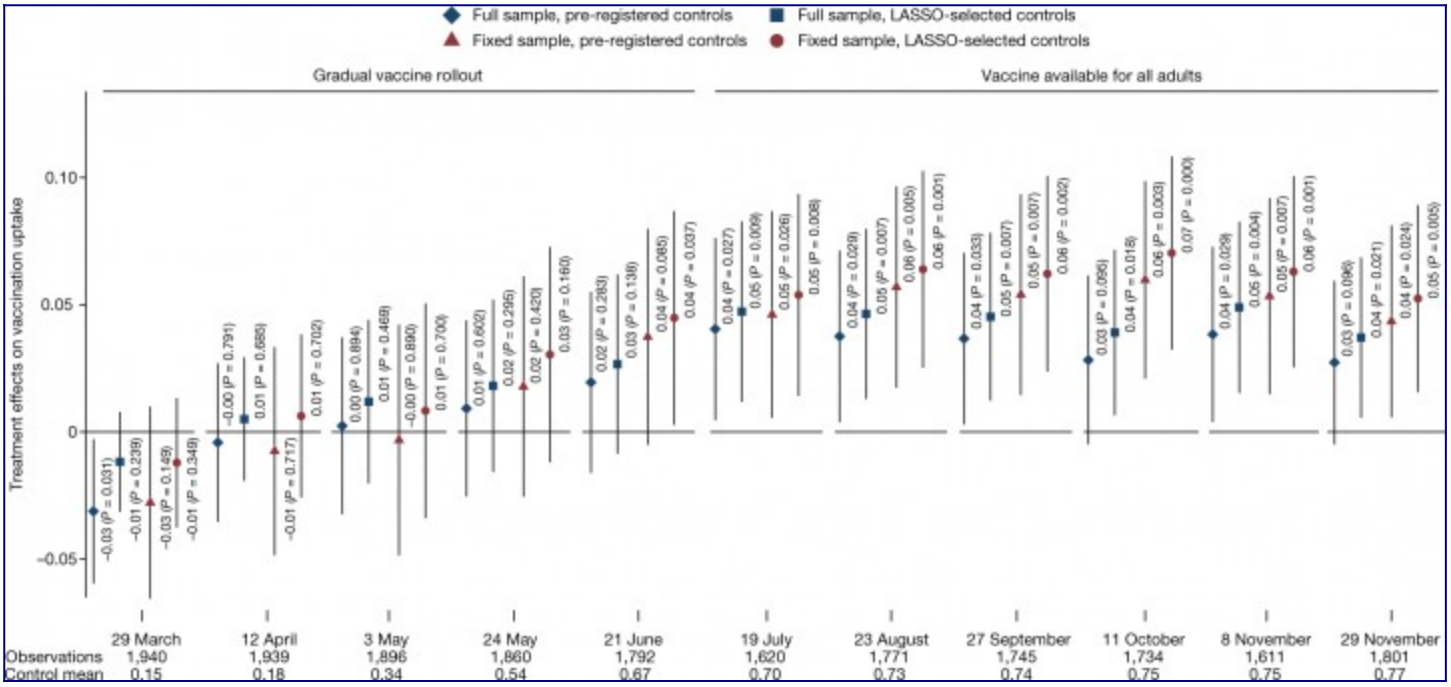
[There is a longstanding dispute between stroke neurologists and some emergency physicians over over the risk/benefit calculus for intravenous thrombolytics in treating AIS. The controversy dates back…](#)



[The mean side of the force : How regression to the mean can fool us - FifteenEightyFour | Cambridge University Press]
cambridgeblog.org
[Regression to the mean is a powerful and common source of bias in interpreting data. Once understood, its potential to mislead is obvious. Yet many scientists are regularly fooled by it. In this blog…]
Shared by [Stephen John Senn]



Successful randomization and covariate “imbalance” in a survey experiment in Nature
statmodeling.stat.columbia.edu
Last year I wrote about the value of testing observable consequences of a randomized experiment having occurred as planned. For example, if the randomization was supposedly Bernoulli(1/2), you can ch…
Shared by Andrew Gelman et al.



[Board members decry their own journal's retraction of paper on predatory publishers](retractionwatch.com)

retractionwatch.com

[Academics affiliated with a journal that retracted a paper on predatory publishing last year — after one of the publishers mentioned in the analysis complained — have put out a letter critiquing the …]

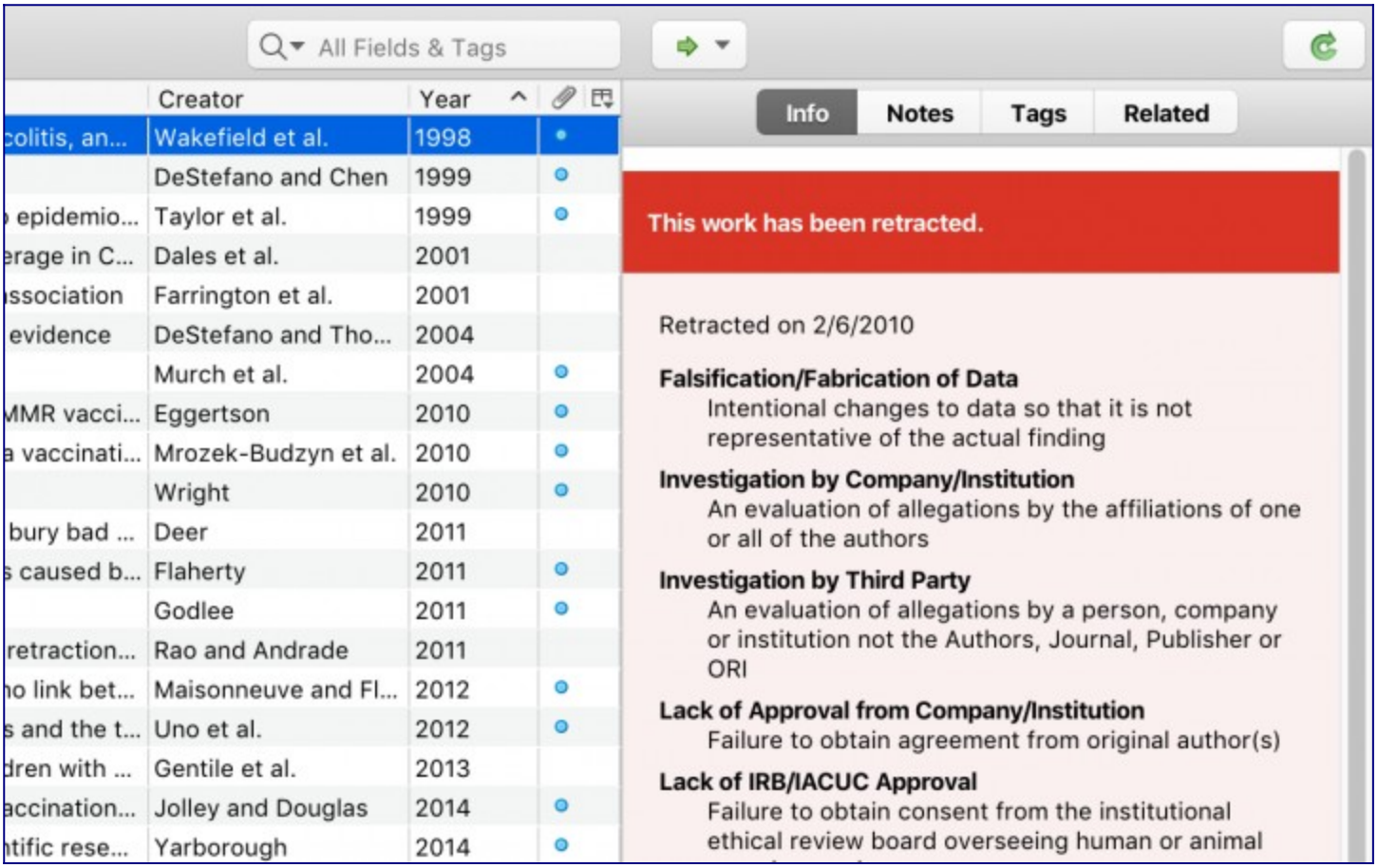
Shared by Ivan Oransky

[Zotero Blog » Blog Archive » Retracted item notifications with Retraction Watch integration](zotero.org)

zotero.org

[Zotero is a free, easy-to-use tool to help you collect, organize, cite, and share research.]

Shared by Retraction Watch

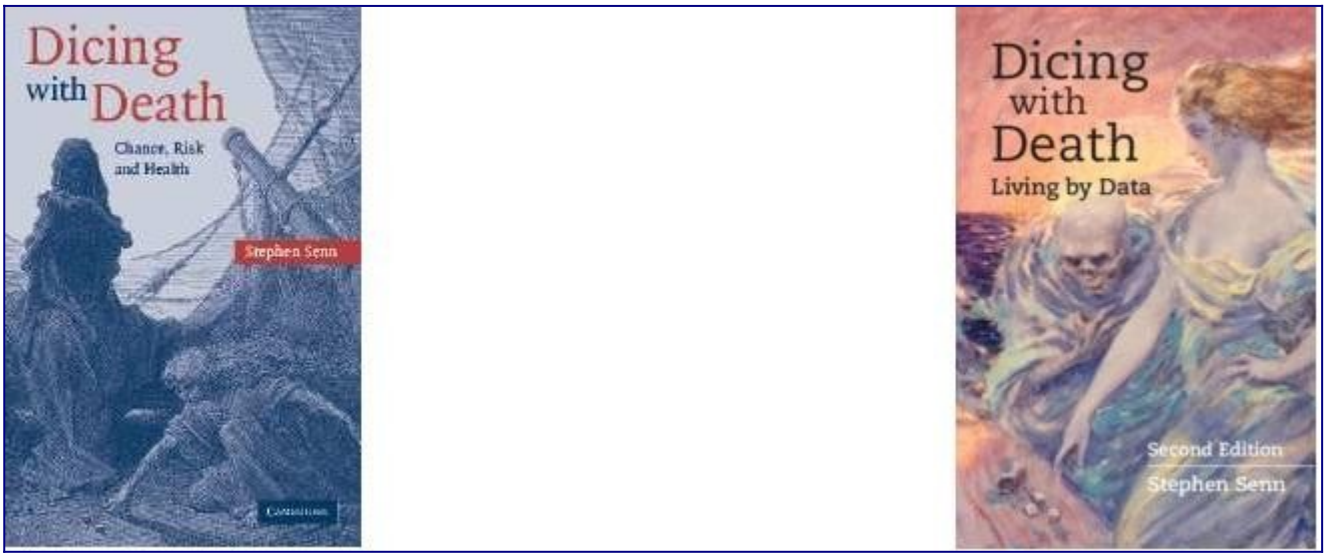


DICE

senns.uk

Statistics is a wonderful discipline. It has it all: mathematics and philosophy, analysis and empiricism, as well as applicability, relevance and the fascination of data. It demands clear thinking, g…

Shared by Stephen John Senn

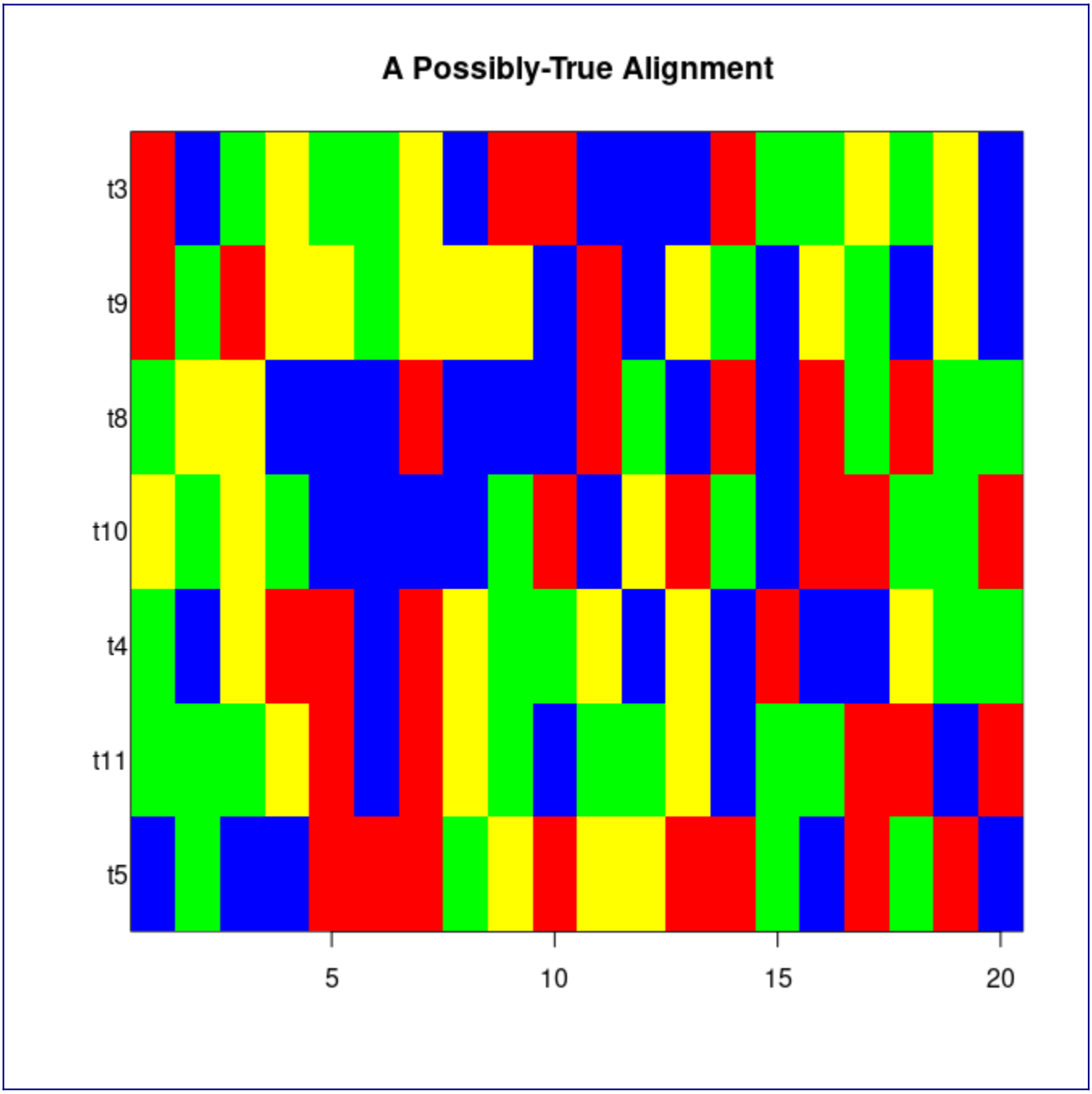


October 2022: "Top 40" New CRAN Packages

rviews.rstudio.com

An R community blog edited by RStudio

Shared by Curtis K



Area Man Accepts Burden Of Being Only Person On Earth Who Understands How World Actually Works

theonion.com

INDIANAPOLIS—Solemnly declaring that he alone could enlighten a human race crippled by ignorance, area man Aaron Krause said Thursday that he had accepted the burden of being the only person on earth…

Shared by Boback Ziaeian

Meta Trained an AI on 48M Science Papers. It Was Shut Down After 2 Days

cnet.com

Galactica was supposed to help "organize science." Instead, it spewed misinformation.

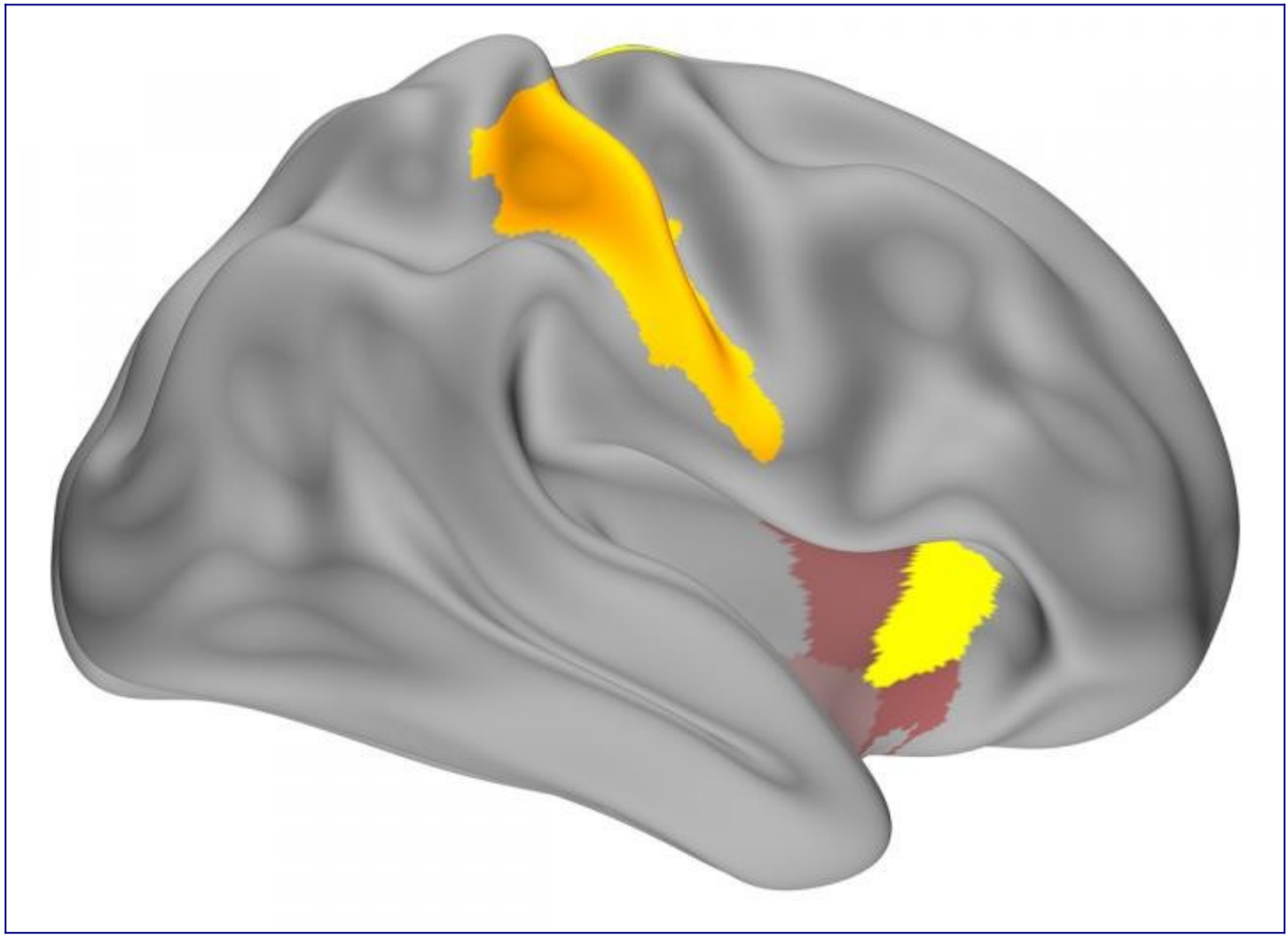
Shared by Retraction Watch

Study shows how taking short breaks may help our brains learn new skills

nih.gov

NIH scientists discover that the resting brain repeatedly replays compressed memories of what was just practiced.

Shared by Edge for Scholars



How heaven became a place among the stars | Aeon Essays

aeon.co

For hundreds of years, Christians knew exactly where heaven was: above us and above the stars. Then came the new cosmologists

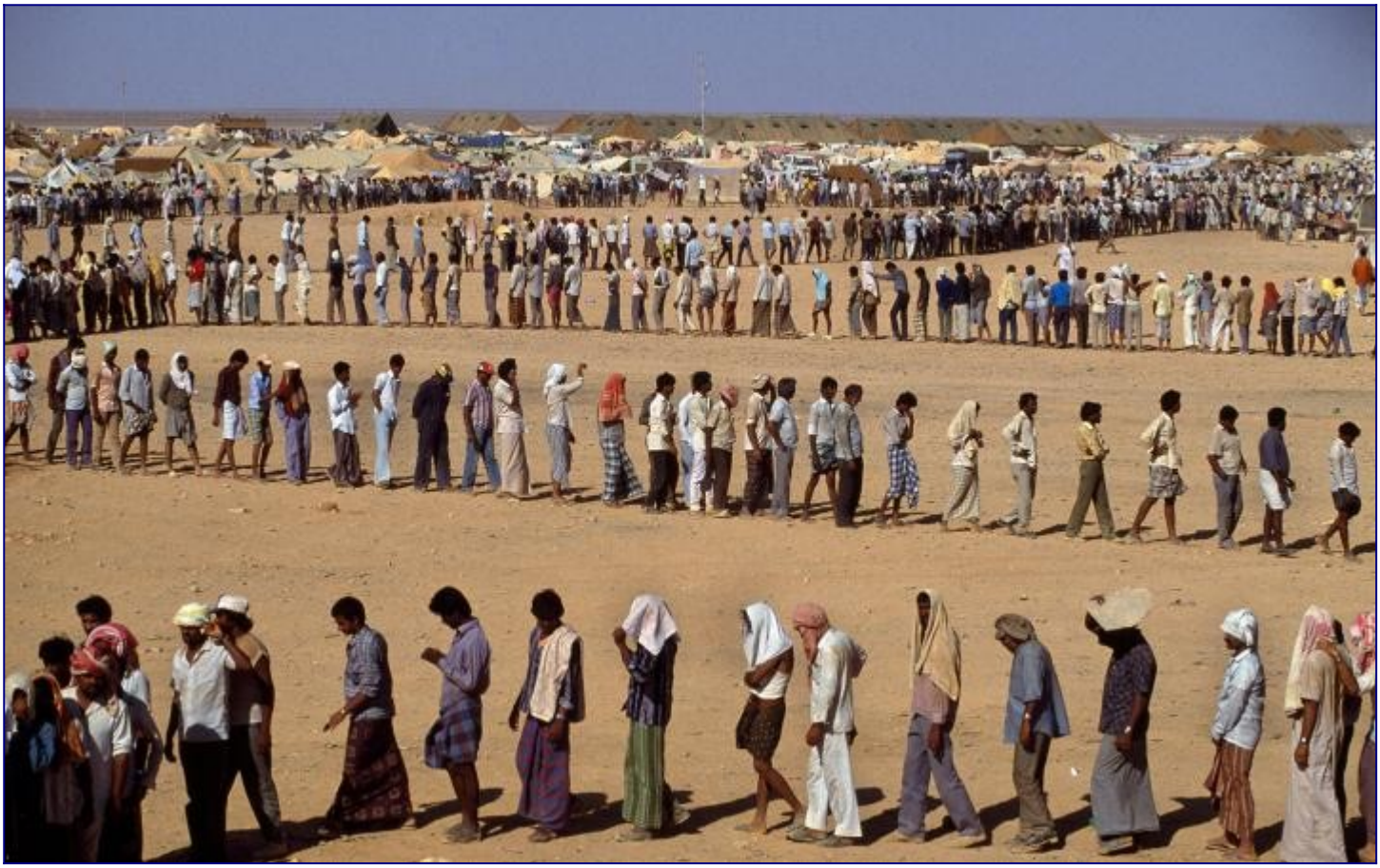
Shared by Aeon+Psyche

How to solve moral problems with formal logic and probability | Aeon Essays

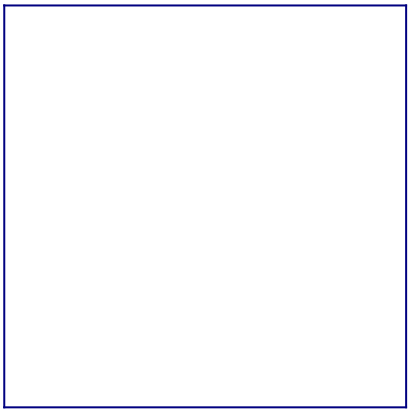
aeon.co

Subjecting the problems of ethics to the cool quantifications of logic and probability can help us to be better people

Shared by Aeon+Psyche



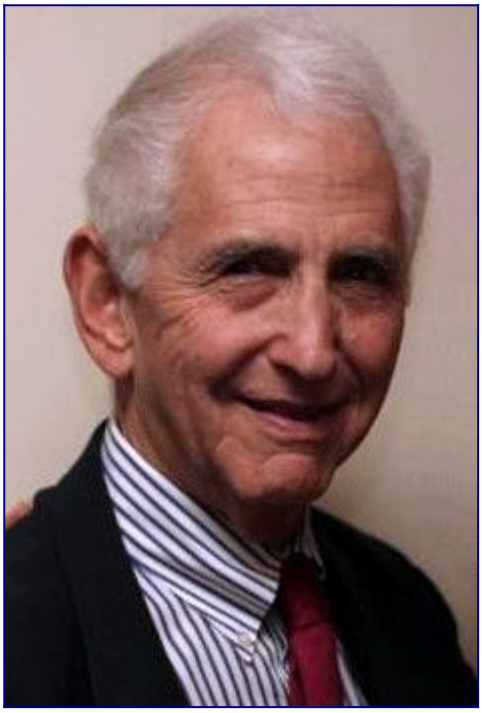
Release of CmdStan 2.31
blog.mc-stan.org
We are very happy to announce that the 2.30.0 release of CmdStan is now available on Github! As usual, the release of CmdStan is accompanied by new releases of Stan Math, core Stan, and Stanc3. A PyS…
Shared by Stan (bayes.club/@mcmc_stan)



Understanding exchangeability in statistical modeling: a Thanksgiving-themed post
statmodeling.stat.columbia.edu
Several years ago, on the day after Thanksgiving, I was on the phone with my sister telling her how I'd used turkey leftovers to make a delicious chili. The process started by pulling out the pieces …
Shared by Andrew Gelman et al.

R Workflow
hbiostat.org
R Workflow for Reproducible Data Analysis and Reporting
Shared by Frank Harrell
Information, Evidence, and Statistics for critical research appraisal
discourse.datamethods.org
In another post about understanding p-values, ESMD made an important observation: ESMD wrote: I suspect hat much of the confusion among physicians learning critical appraisal stems from the fact that…

Ellsberg paradox
en.wikipedia.org
In decision theory, the Ellsberg paradox (or Ellsberg's paradox) is a paradox in which people's decisions are inconsistent with subjective expected utility theory. Daniel Ellsberg popularized the par…
Shared by John Allen Paulos



Questionable science and reproducibility in electrical brain stimulation research
journals.plos.org
Electrical brain stimulation (EBS) is a trendy new technique used to change brain function and treat neurological, psychiatric and psychological disorders. We were curious whether the published liter…
Shared by Retraction Watch

Questionable research practices	Others (%)	Self (%)	Audit (%)
Adjust statistical analyses in order to optimise the results	43	25	0
Not report all experimental conditions	36	13	0
Screen whether subjects are responders and not report it	21	4	0
Exclude data based on a gut feeling	21	8	0
Exclude data after looking at impact on results	20	9	0
Exclude trials or subjects without support of statistical analysis	22	8	2
Selectively report outcomes	41	14	0
Selectively report time points	18	3	0
Selectively report types of EBS used in study	12	4	0
Selectively report sub-groups of subjects	24	14	0

See S2 File for the exact wording used in the online survey.

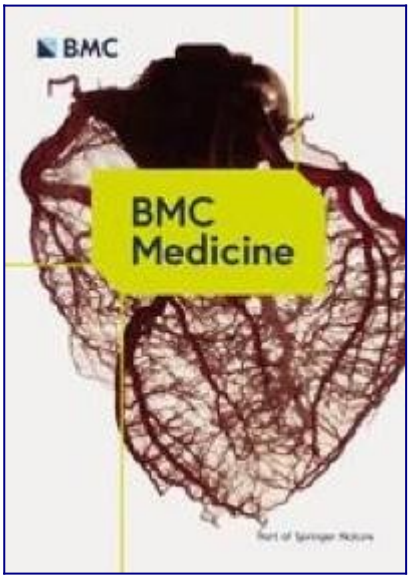
https://doi.org/10.1371/journal.pone.0175635.t003

Desperate covid long-haulers turn to costly, unproven treatments
washingtonpost.com
Unproven treatments put patients at risk of potentially harmful health effects as well as having their hopes dashed and their wallets emptied.
Shared by Harlan Krumholz

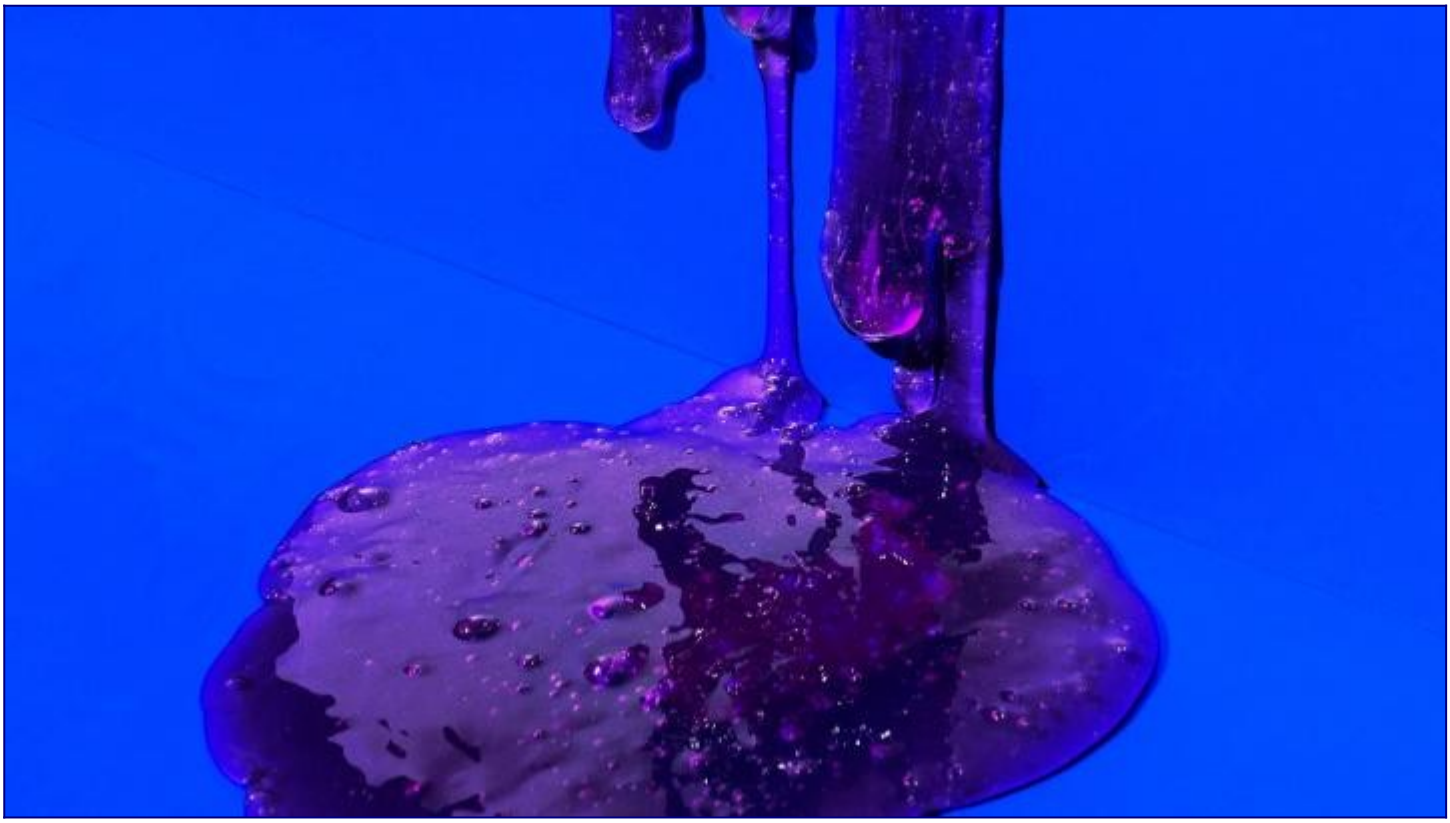
[You’ve decided to quit Twitter. Here’s what you can use to replace it.]
washingtonpost.com
The best alternatives to Twitter for people who don't want to stay on the Elon Musk-owned site.
Shared by Edward Tufte

[Why doesn’t physics help us to understand the flow of time? | Aeon Essays]
aeon.co
From past to present, into the future: the flow of time is central to human experience. Why isn’t it central to physics?
Shared by Aeon+Psyche

How often do cancer researchers make their data and code available and what factors are associated with sharing?
bmcmedicine.biomedcentral.com
Various stakeholders are calling for increased availability of data and code from cancer research. However, it is unclear how commonly these products are shared, and what factors are associated with …
Shared by Retraction Watch



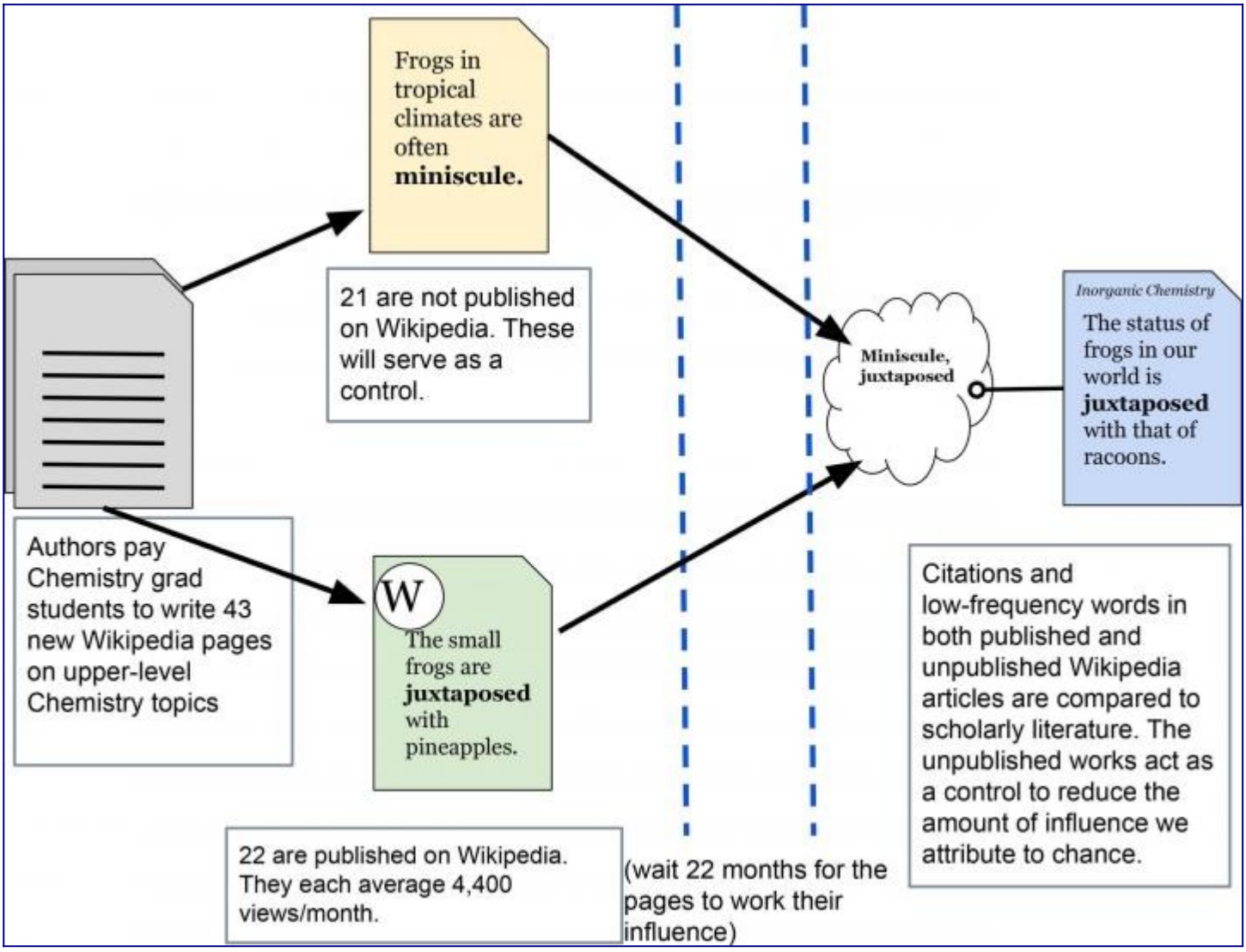
Twitter Users Have Caused a Mastodon Meltdown
wired.co.uk
The network’s servers are overrun by almost half a million new sign-ups, as Elon Musk’s Twitter takeover has thrust it into the spotlight.
Shared by Boback Ziaeian



How Schopenhauer’s thought can illuminate a midlife crisis
aeon.co
How to solve a midlife crisis, the Schopenhauer way: go with the flow and don’t look for projects that require completion
Shared by Aeon+Psyche

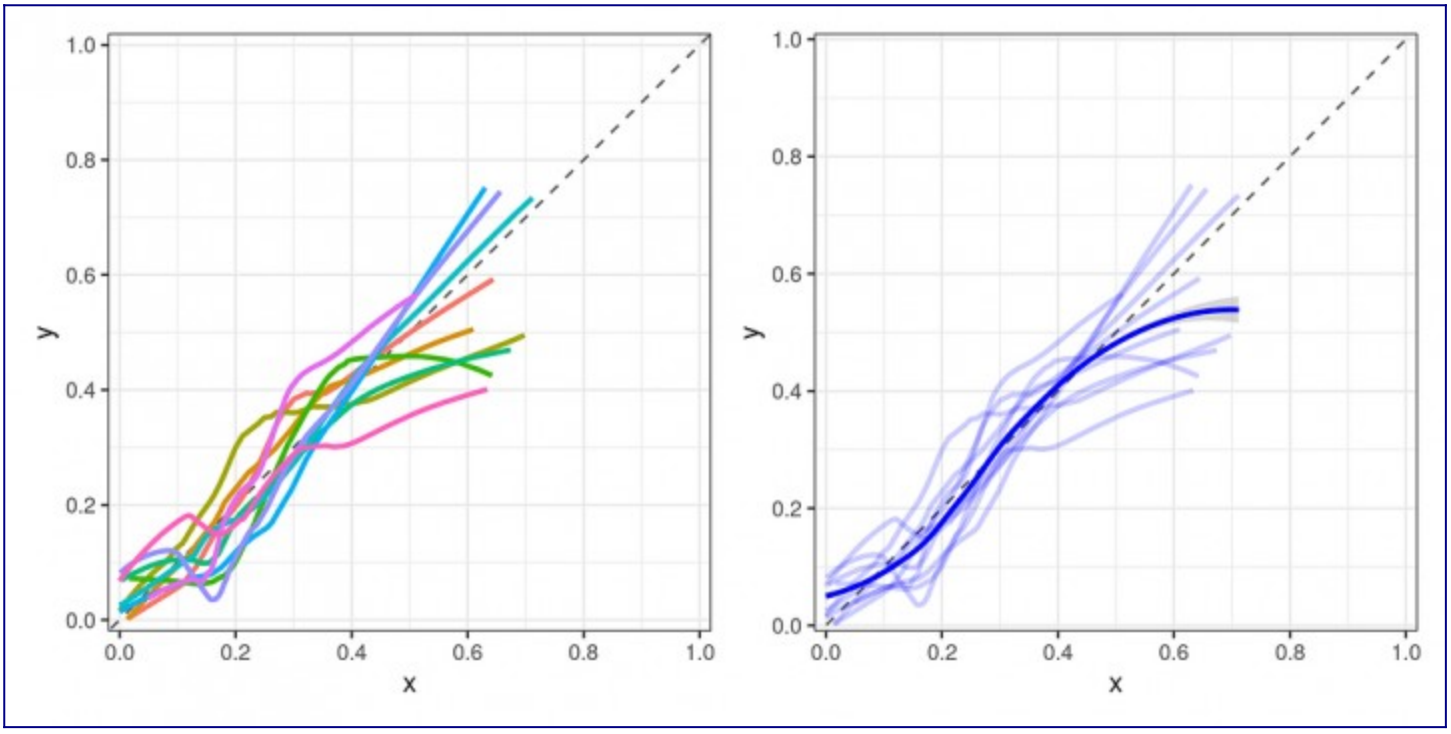
[On the use of non-concurrent controls in platform trials: A scoping review]
arxiv.org
Platform trials gained popularity during the last few years as they increase flexibility compared to multi-arm trials by allowing new experimental arms entering when the trial already started. Using …
Shared by Martin Posch

Guest Post – Wikipedia's Citations Are Influencing Scholars and Publishers
scholarlykitchen.sspnet.org
Rachel Helps, the Wikipedian-in-residence at the BYU libraries discusses the intersection of scholarly journals and Wikipedia.
Shared by Retraction Watch



[Proper way to get an averaged calibration curve from nested resampling](discourse.datamethods.org)

[Dear all, I am working on some projects where the overarching goal is to assess whether a selection of machine learning algorithms perform better than “traditional” statistical methods such as logist…]



DICE

senns.uk

Statistics is a wonderful discipline. It has it all: mathematics and philosophy, analysis and empiricism, as well as applicability, relevance and the fascination of data. It demands clear thinking, g…

Shared by Stephen John Senn

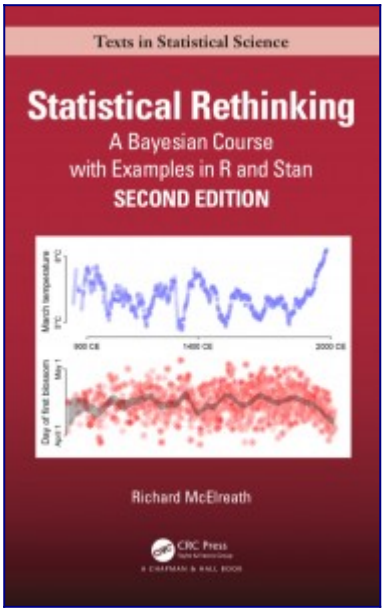
Journal editor in chief who published controversial Covid papers resigns

retractionwatch.com

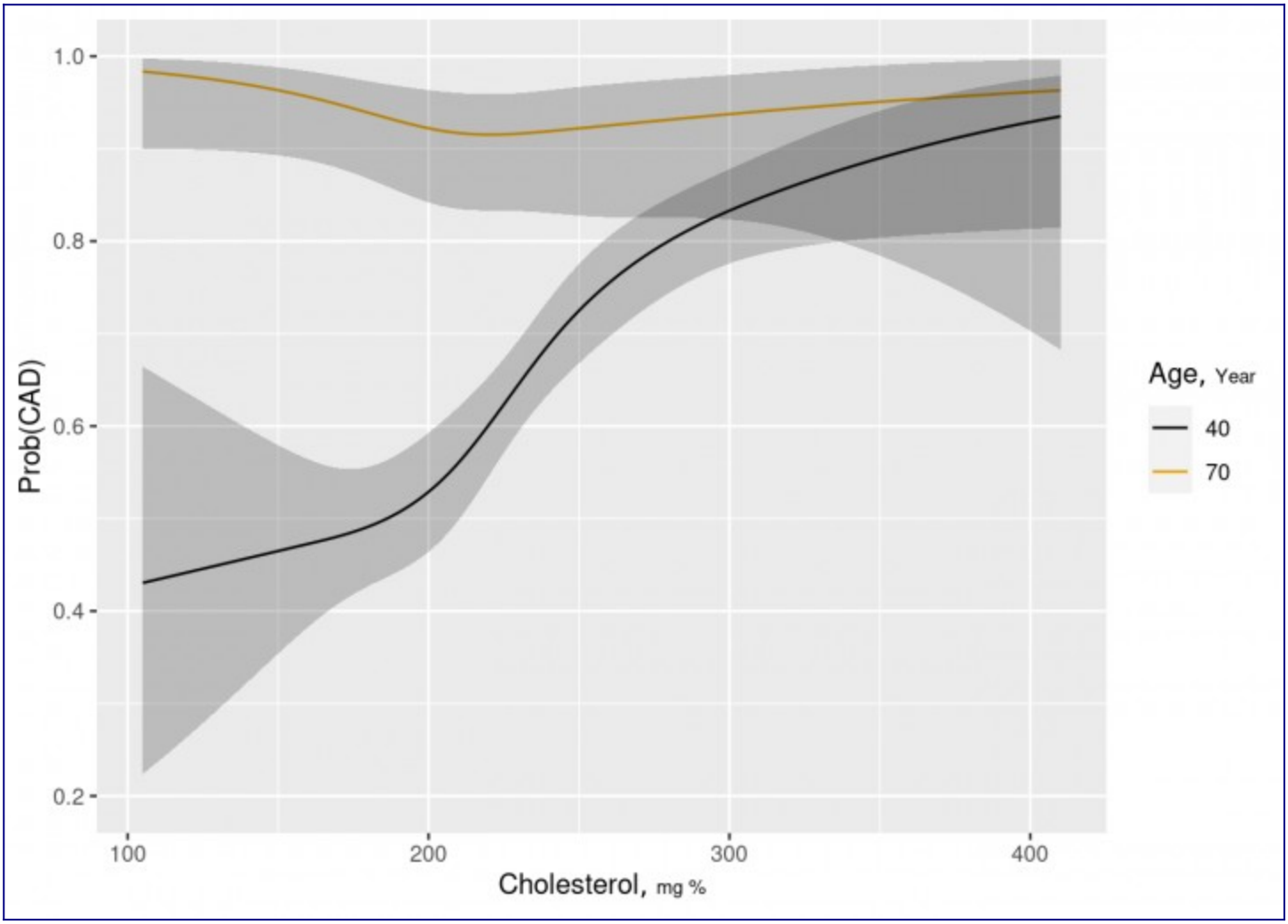
Jose L. Domingo The editor in chief of the journal Food and Chemical Toxicology (FCT) has resigned with more than a year left of his term, according to an email announcing his move to colleagues. In …

Shared by Retraction Watch

Statistical Rethinking
xcelab.net
Most recent set of free lectures: Statistical Rethinking 2022 The second edition is now out in print. Publisher information on the CRC Press page. For more detail about what is new, look here. Statis…
Shared by 黒木玄 Gen Kuroki



Statistical Thinking - Statistically Efficient Ways to Quantify Added Predictive Value of New Measurements
fharrell.com
Researchers have used contorted, inefficient, and arbitrary analyses to demonstrated added value in biomarkers, genes, and new lab measurements. Traditional statistical measures have always been up t…
Shared by Frank Harrell



[Opinion | Science Has a Nasty Photoshopping Problem]
nytimes.com
Scientists need to toughen up about preventing fabricated scientific results from being published.
Shared by Retraction Watch

On the Convergence of Epidemiology, Biostatistics, and Data Science
hdsr.mitpress.mit.edu
Epidemiology, biostatistics, and data science are broad disciplines that incorporate a variety of substantive areas. Common among them is a focus on quantitative approaches for solving intricate prob…
Shared by StateofMind

[Random sampling versus random allocation/randomization- implications for p-value interpretation](discourse.datamethods.org)
[The questions below are very basic and have probably been asked thousands of times by statistically-naive MDs- apologies in advance. Unfortunately, I keep getting hung up on them and can’t find clear…](discourse.datamethods.org)
Shared by Frank Harrell

[Science is still in its' infancy: a few simple randomized trials could save us](vinayprasadmdmph.substack.com)
[We waste money on laboratory and observational research.](vinayprasadmdmph.substack.com)
Shared by Vinay Prasad MD MPH

Association Between Price and Measures of Efficacy for Oncology Drugs Approved by FDA From 2015 to 2020

jamanetwork.com

This cross-sectional study estimates all US Food and Drug Administration anticancer approvals in recent years and evaluates if an association exists between the

Shared by Vinay Prasad MD MPH

Table. Characteristics of US Food and Drug Administration Oncology Drug Approvals From 2015 to 2020

Characteristic	Overall, No. (%)
No.	224
Annual cost, median (IQR), $	196 000 (170 000-277 000)
Mechanism	
Small molecule therapy	102 (45.5)
Biologics	99 (44.2)
Cytotoxic therapy	11 (4.9)
Hormonal therapy	7 (3.1)
Gene and oncolytic virus therapy	5 (2.2)
Tumor type	
Other	106 (47.3)
Non-small-cell lung cancer	32 (14.3)
Breast	21 (9.4)
Myeloma	14 (6.2)
Acute myeloid leukemia	12 (5.4)
Melanoma	11 (4.9)
Urothelial	10 (4.5)
Ovarian	9 (4.0)
Prostate	9 (4.0)
Primary end point	
Overall response rate	90 (40.2)
Progression-free survival	71 (31.7)

When editors confuse direct criticism with being impolite, science loses

retractionwatch.com

Jasmine Jamshidi-Naeini In January 2022, motivated by our experience with eClinicalMedicine, we wrote about mishandling of published errors by journal editors. We had noticed that the methods used fo…

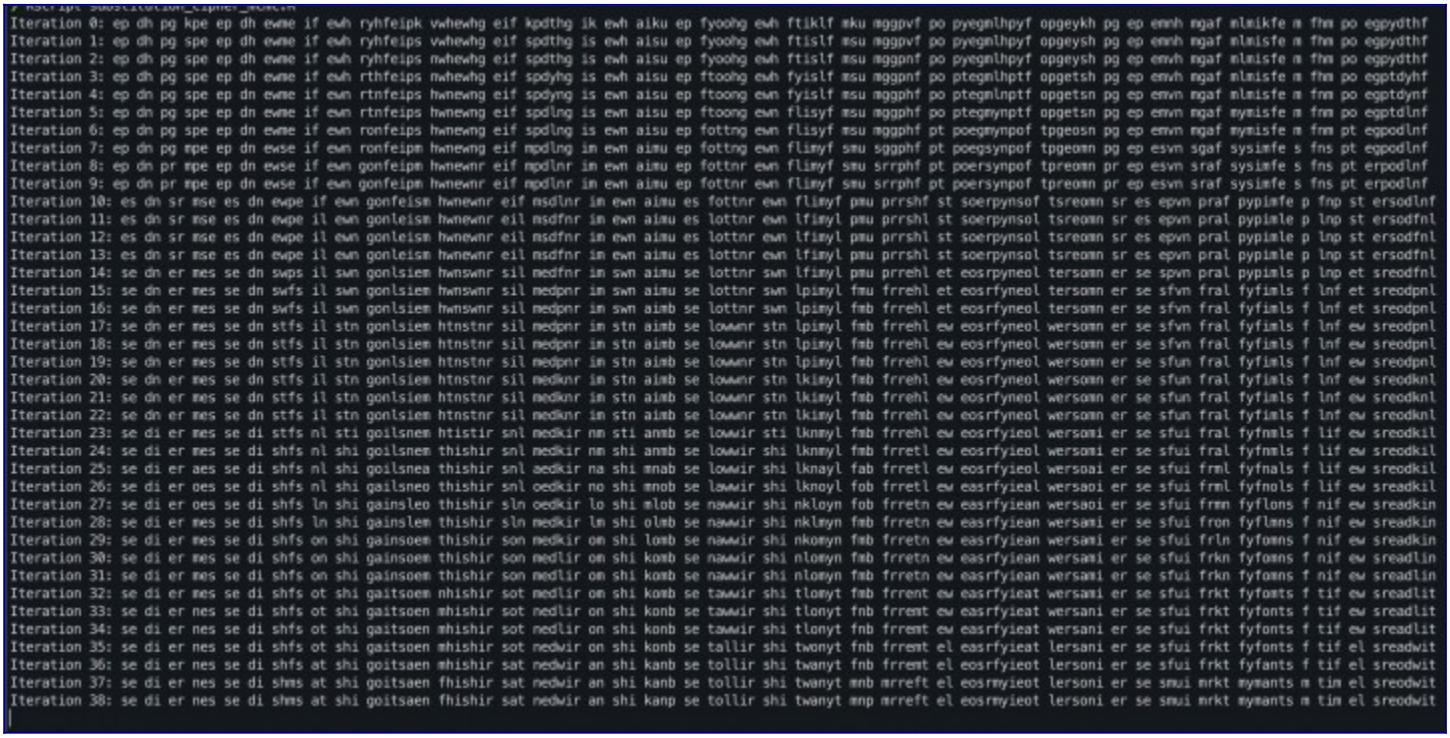
Shared by Retraction Watch

Max Rohde - Code-breaking with Markov Chain Monte Carlo (MCMC)

maximilianrohde.com

Using cryptography to demonstrate the power of MCMC techniques in computational statistics

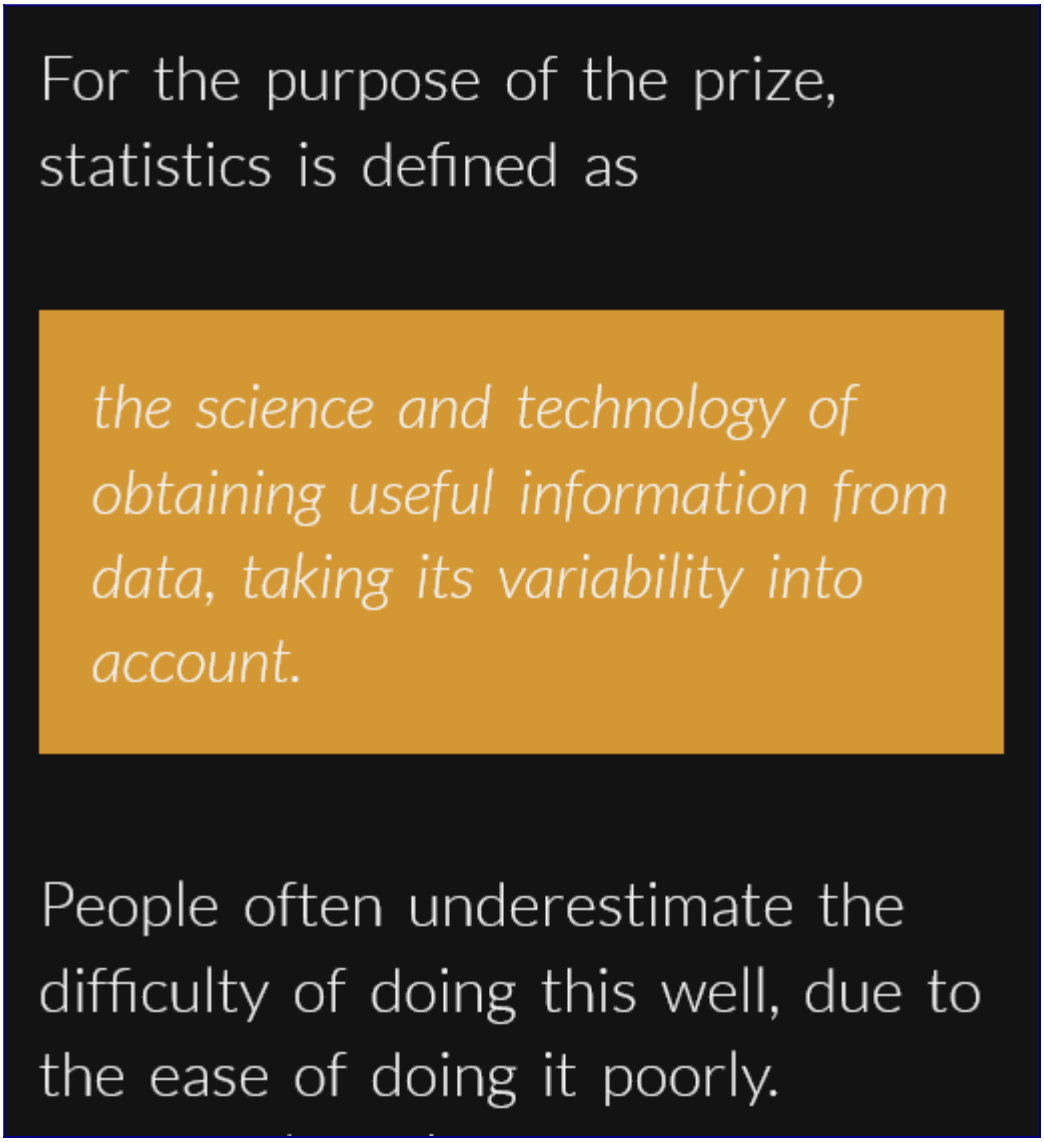
Shared by Max Rohde



Definition of statistics

discourse.datamethods.org

I am not a statistician (veterinarian) but I have an interest in improving the quality of research in my field where statisticians are infrequently part of study design or analysis. When I try to con…



[“A huge relief”: Journal takes down plagiarized paper after Retraction Watch reporting](#)
retractionwatch.com
[Andrew Colman Following a Retraction Watch story about a 2004 paper that had been copied twice since its publication, one of the journals involved has taken down its version of the article. Last mont…](#)
Shared by [Ivan Oransky](#)

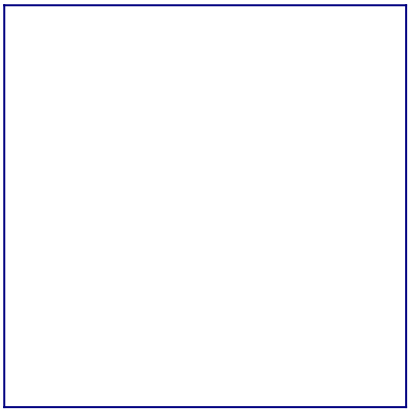
[CmdStan & Stan 2.31 release candidate](#)
discourse.mc-stan.org
[I am happy to announce that the latest release candidates of Cmdstan and Stan are now available on Github! This release cycle brings algebraic solver functions with variadic arguments, truncation syn…](#)
Shared by [Stan (bayes.club/@mcmc_stan)](#)

[Notes from a data witch - Everything I know about Mastodon](#)
[blog.djnavarro.net](#)
[A hastily written guide for data science folks trying to navigate the fediverse.](#)
Shared by [Mara Averick](#)
[He’s an Outspoken Defender of Meat. Industry Funds His Research, Files Show.](#)
[nytimes.com](#)
[A UC Davis professor runs an academic center that was conceived by a trade group, according to records, and gets most of its funding from farming interests.](#)
Shared by [Retraction Watch](#)

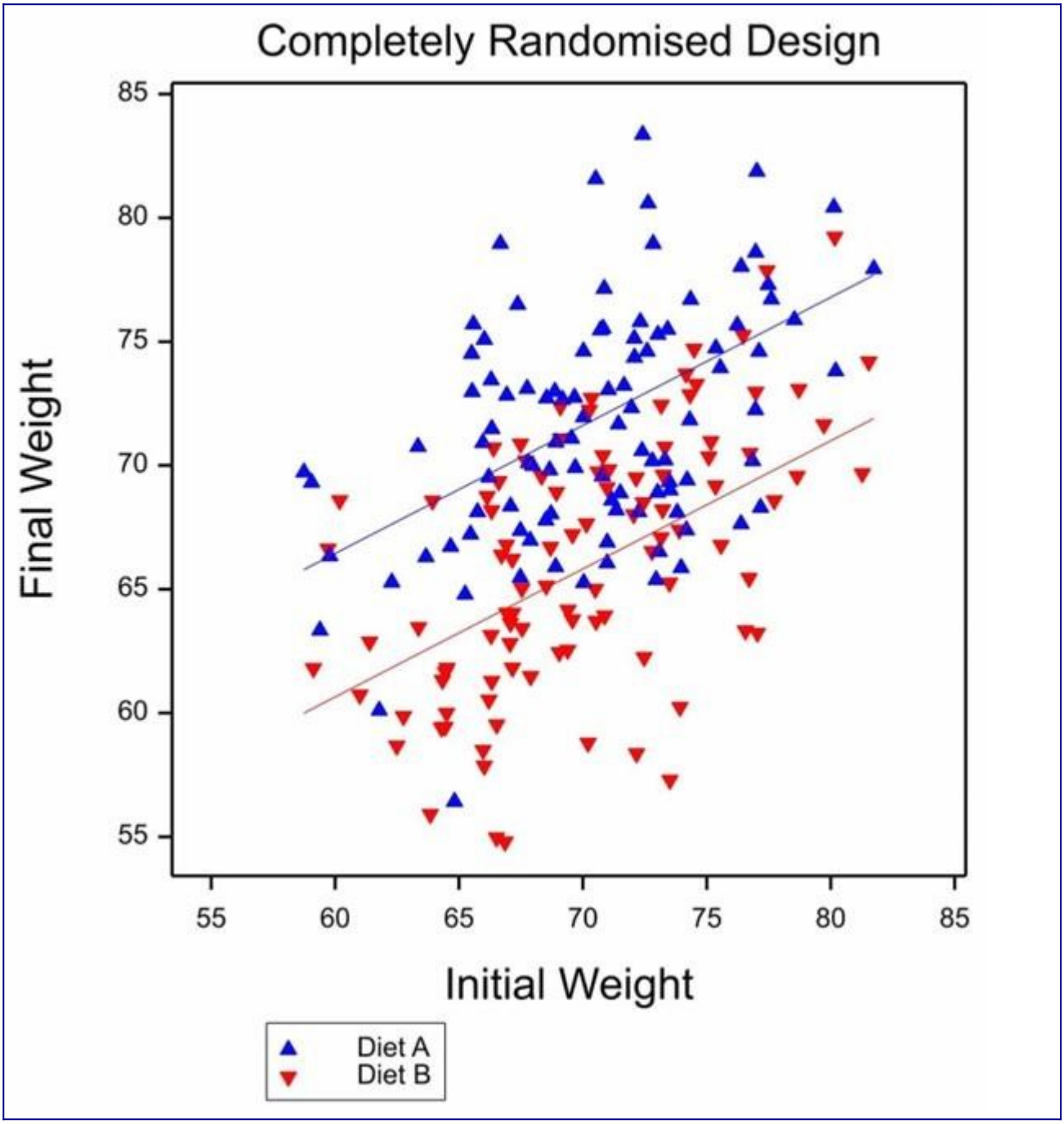
Reproducibility in Clinical Trials is Like Trying to Catch a Rabbit by the Tail
neo.life
Scientists are not routinely sloppy, clinicians don’t usually cheat, and clinical research is generally well designed. So why is it so hard to reproduce
Shared by Retraction Watch

Stability of clinical prediction models developed using statistical or machine learning methods
arxiv.org
Clinical prediction models estimate an individual's risk of a particular health outcome, conditional on their values of multiple predictors. A developed model is a consequence of the development data…
Shared by Richard Riley (R^2)

[Mixed models repeated measures (mmrm) package for R]
thestatsgeek.com
I was recently made aware of the release of the mmrm package in R. It has been developed by a group of programmers and statisticians at a number of pharmaceutical companies, led by Daniel Sabanes B…
Shared by Jonathan Bartlett



S. Senn: Lauding Lord (Guest Post)
errorstatistics.com
Stephen SennConsultant StatisticianEdinburgh, Scotland A Diet of Terms A large university is interested in investigating the effects on the students of the diet provided in the university dining hall…
Shared by Stephen John Senn



[You've decided to quit Twitter. Here's what you can use to replace it.](#)
washingtonpost.com
[The best alternatives to Twitter for people who don't want to stay on the Elon Musk-owned site.](#)
Shared by [Allen Frances](#)

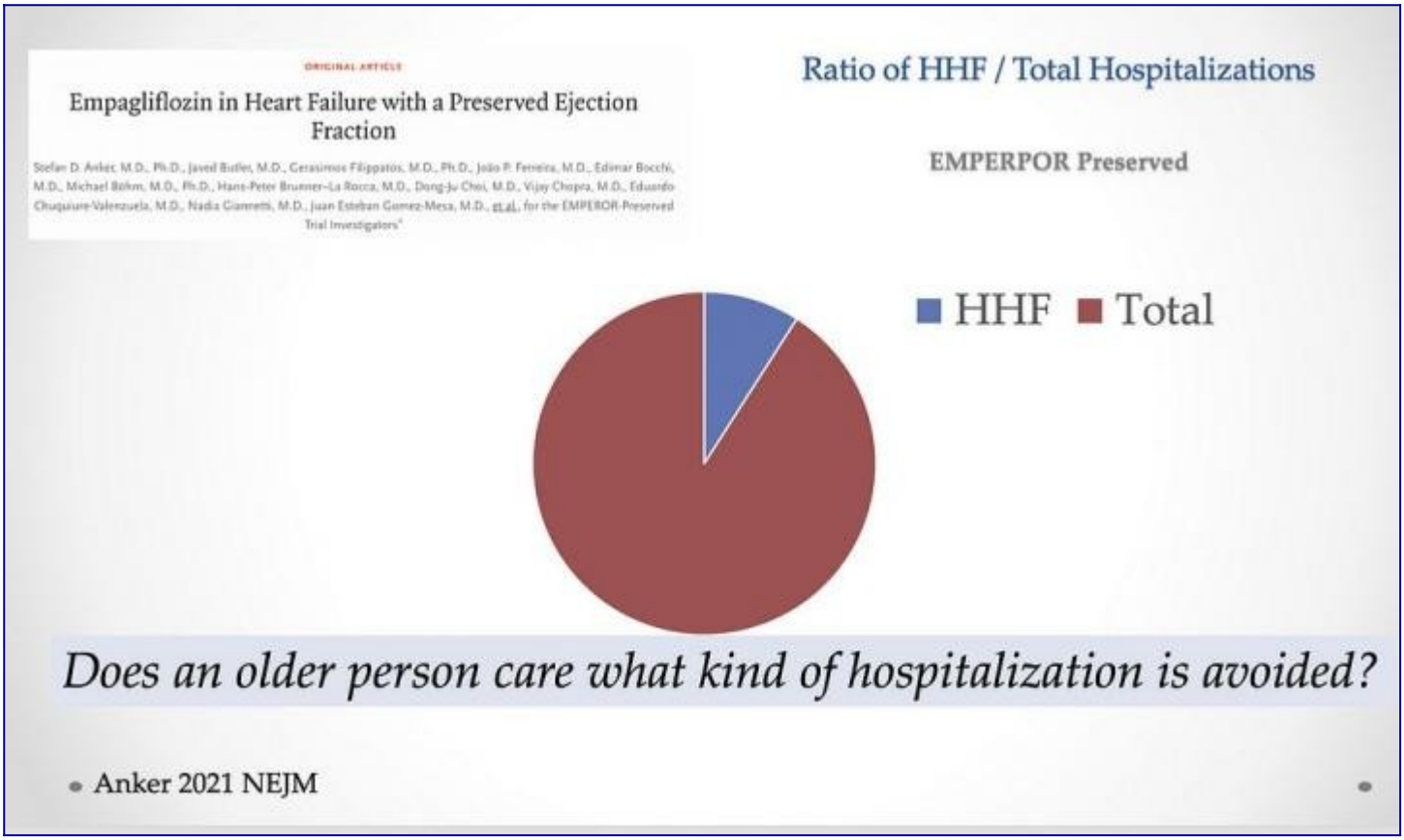
Elon buys Twitter: Good!
vinayprasadmdmph.substack.com
Can he fix it?
Shared by Vinay Prasad MD MPH

I wanted to reach out personally to share my motivation in acquiring Twitter. There has been much speculation about why I bought Twitter and what I think about advertising. Most of it has been wrong.

The reason I acquired Twitter is because it is important to the future of civilization to have a common digital town square, where a wide range of beliefs can be debated in a healthy manner, without resorting to violence. There is currently great danger that social media will splinter into far right wing and far left wing echo chambers that generate more hate and divide our society.

In the relentless pursuit of clicks, much of traditional media has fueled and catered to those polarized extremes, as they believe that is what brings in the money, but, in doing so, the opportunity for dialogue is lost.

[The All-Important Endpoint of a Medical Study]
[johnmandrola.substack.com]
[Medical studies are often hyped as "positive." Consumers of medical evidence must always be on alert for the endpoint measured. Not all positive results are the same.]
Shared by [John Mandrola, MD]

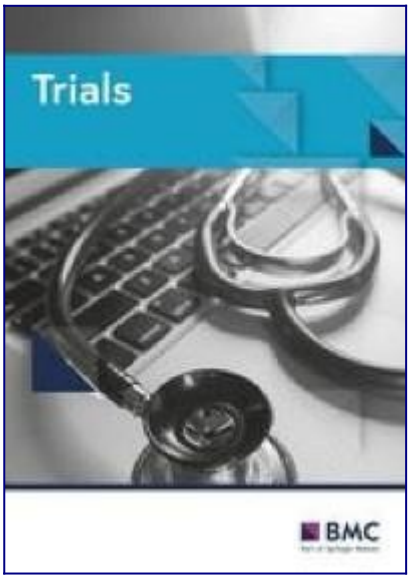


10 quick Mastodon tips
axbom.com
There has recently been a huge rush of people trying out Mastodon. This is also noticeable in the number of people reading my earlier Mastodon Guide. I wanted to take this opportunity to provide a br…
Shared by @briannosek@nerdculture.de

How a magician-mathematician revealed a casino loophole
bbc.com
When a gang of gambling cheats sussed out how to beat the house, they inadvertently highlighted a loophole from a shuffled deck. It took a magician-turned-mathematician to reveal how.
Shared by Darren Dahly https://mastodon.social/@statsepi

Introduction to Bayes for Evaluating Treatments
hbiostat.org
P-values and the p < 0.05 rule of thumb came into use before the computing revolution. Assuming the null hypothesis is true greatly simplified the model, often requiring only manual calculations.
Shared by Frank Harrell

What is the purpose of clinical trial monitoring?
trialsjournal.biomedcentral.com
The sources of information on clinical trial monitoring do not give information in an accessible language and do not give detailed guidance. In order to enable communication and to build clinical tri…
Shared by Darren Dahly https://mastodon.social/@statsepi

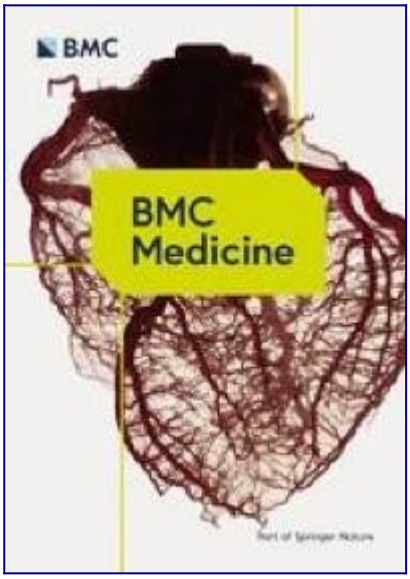


Three myths about risk thresholds for prediction models

bmcmedicine.biomedcentral.com

Clinical prediction models are useful in estimating a patient’s risk of having a certain disease or experiencing an event in the future based on their current characteristics. Defining an appropriate…

Shared by Maarten van Smeden



How Bayesian Additive Regression Tree(BART) algorithm works part1

medium.com

Abstract : Bayesian Additive Regression Trees (BART) is a popular Bayesian non-parametric regression algorithm. The posterior is a distribution over sums of decision trees, and predictions are made…

Shared by Ramsey Elbasheer

A Brief Mastodon Guide for Social Media Worriers
axbom.com
Tips and advice for getting started with Mastodon. Especially if you're used to Twitter.
Shared by @briannosek@nerdculture.de

When to check model assumptions
stats.stackexchange.com
Statistical methods are based on model assumptions. For example, an independent one-way ANOVA makes the following assumptions: Normally distributed residuals Homogeneity of variance Independence...
Shared by Frank Harrell

Regression models for 2-timepoint non-experimental data | A. Solomon Kurz
solomonkurz.netlify.app
Purpose In the contemporary longitudinal data analysis literature, 2-timepoint data (a.k.a. pre/post data) get a bad wrap. Singer and Willett (2003, p. 10) described 2-timepoint data as only “margina…
Shared by Aki Vehtari

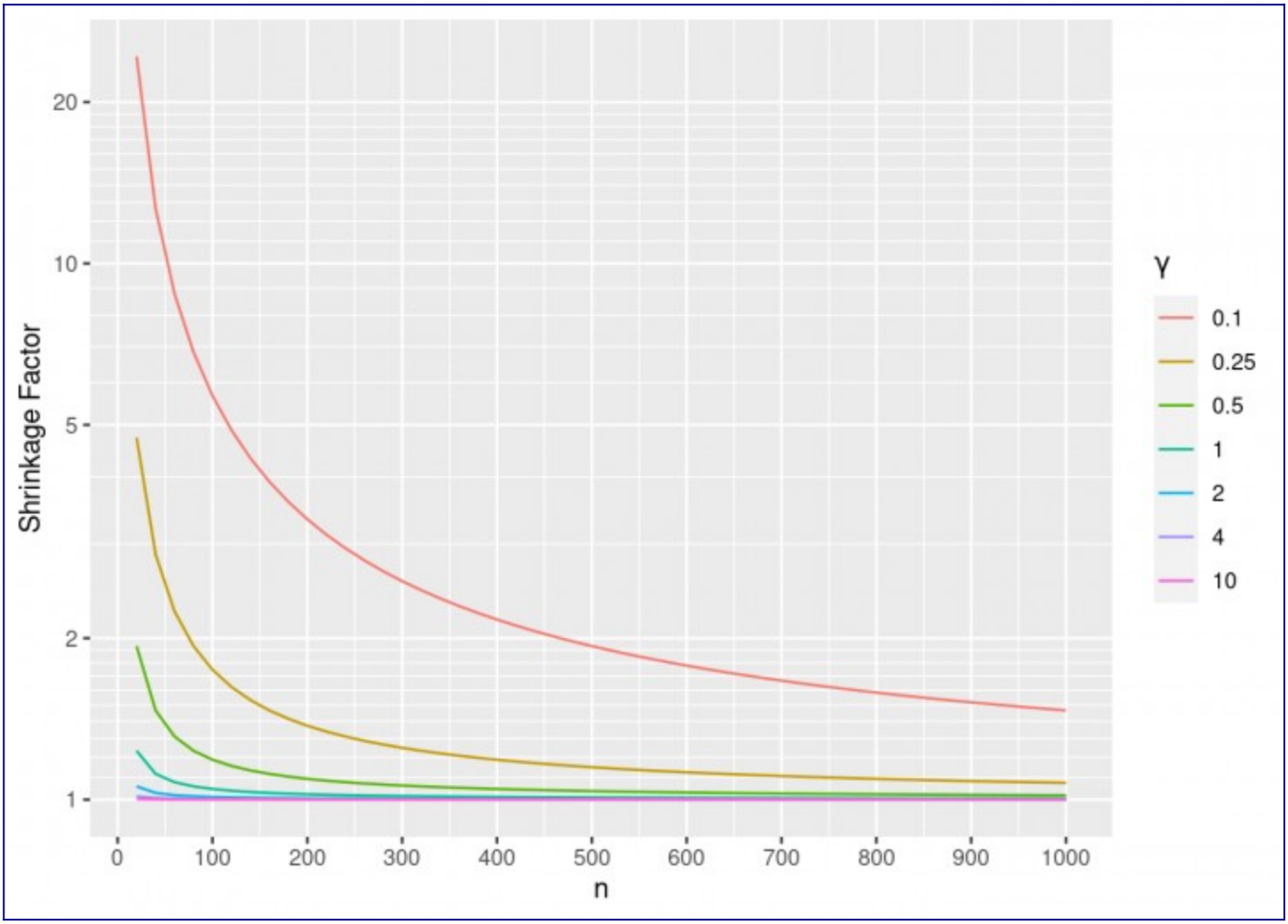
Speculating about assigning probabilities to speculative theories: Lots of problems here.
statmodeling.stat.columbia.edu
Whenever I see theories about ancient stuff I always feel it is very speculative. “This artifact is a stone ax from a hominid from c. 800,000 BP”. “The Samson story in the book of Judges is based on …
Shared by Andrew Gelman et al.

More Endpoints, More Problems: FDA Offers Advice for Multi-endpoint Trials
tctmd.com
It makes sense that trialists want their studies to answer multiple questions, but for findings to hold water, proper design is key.

Note that "but failing to account for multiplicity can lead to a type 1 error where it’s falsely concluded the drug is effective" is problematic. alpha is not the probability of making a mistake. It's P(making an assertion) when any assertion of efficacy is by def. false.
Shared by Michael O'Riordan

Biostatistics for Biomedical Research - 10 Simple and Multiple Regression Models and Overview of Model Validation
hbiostat.org
Background
Shared by Frank Harrell

Statistical Thinking - Assessing the Proportional Odds Assumption and Its Impact
fharrell.com
This article demonstrates how the proportional odds (PO) assumption and its impact can be assessed. General robustness to non-PO on either a main variable of interest or on an adjustment covariate ar…
Shared by Frank Harrell

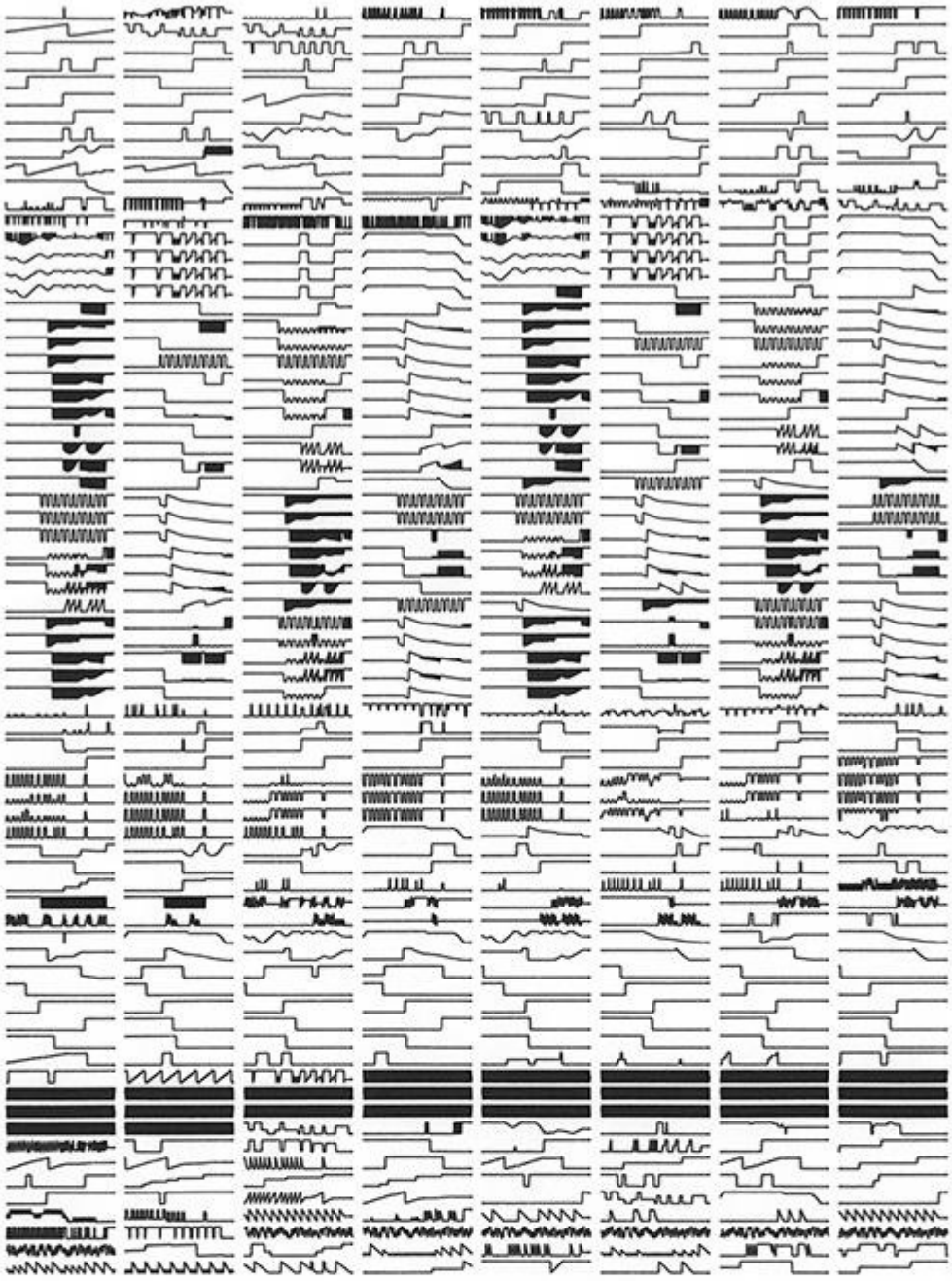


FDA explains why it won't pull Jazz's accelerated approval despite a failed confirmatory trial
endpts.com

When does a failed confirmatory trial evaluating overall survival not amount to a request from FDA to pull an accelerated approval for a cancer drug? That's the situation right now for Jazz Pharmaceu…

Shared by Vinay Prasad MD MPH

Edward Tufte forum: Sparkline theory and practice Edward Tufte
edwardtufte.com
A sparkline is a small intense, simple, word-sized graphic with typographic resolution. Sparklines mean that graphics are no longer cartoonish special occasions with captions and boxes, but rather sp…
Shared by Edward Tufte



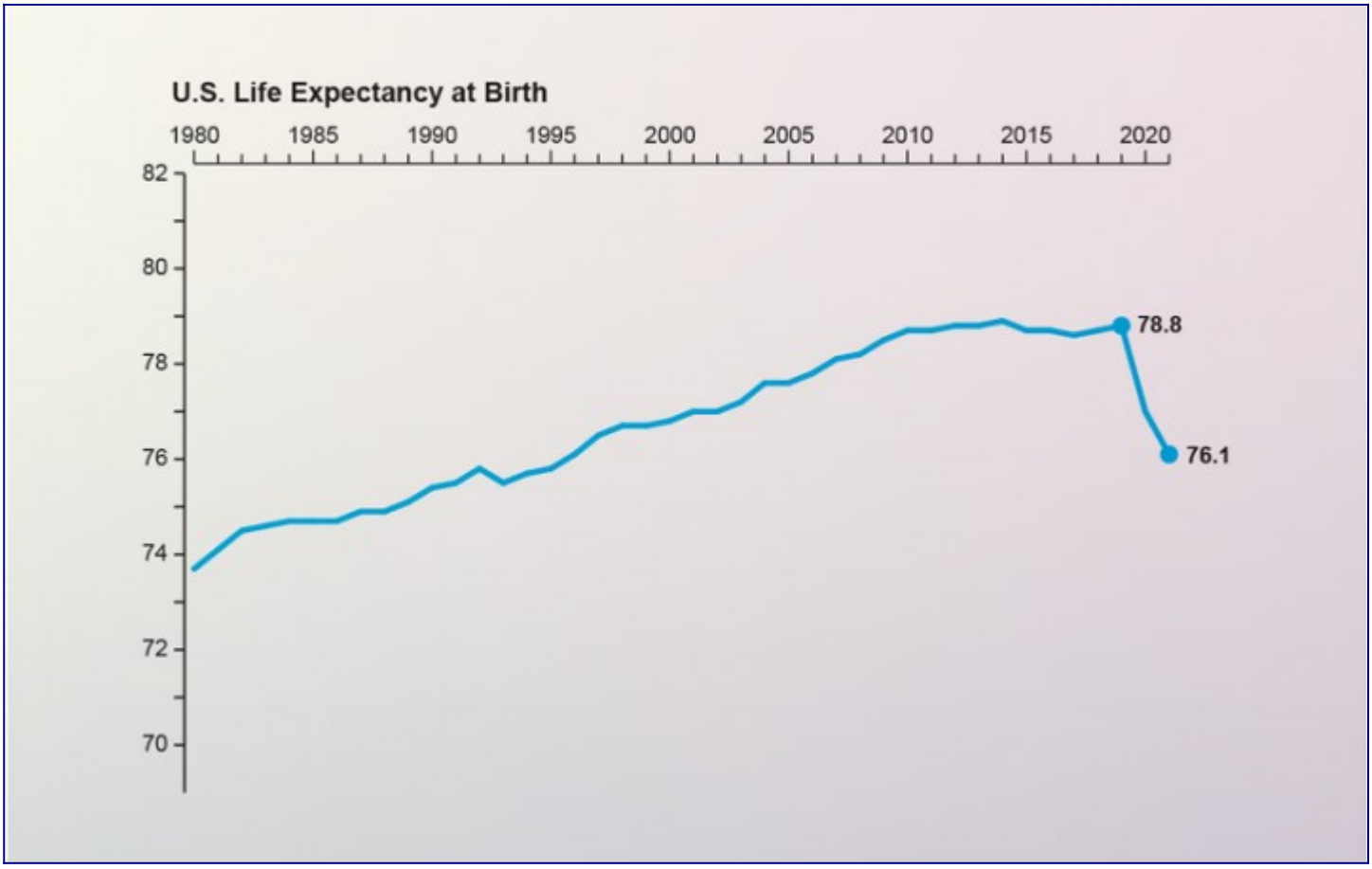
#3 Five Seconds of Super Mario Bros. MAF

NES Sparklines

For this drawing, I used an NES emulator (of my own creation) to record a snapshot of the Nintendo's RAM at each frame (60 fps) for several seconds while playing a game. The NES had just 2048 bytes of RAM! For each address in memory, I plotted its values (ranging from 0 to 255) over time as an individual sparkline. I only included addresses that changed at least once, so there are fewer than 2048 sparklines. Furthermore, I individually scaled each sparkline so as to utilize the full vertical range. Some values oscillate so rapidly that the lines can no longer be resolved when plotted - but this produces an interesting artistic effect. The sparklines are arranged in a grid-like pattern in row-major order. Because

Cheating in sports vs. cheating in journalism vs. cheating in science
statmodeling.stat.columbia.edu
Sports cheating has been in the news lately. Nothing about the Astros, but the chess-cheating scandal that people keep talking about—or, at least, people keep sending me emails asking me to blog abou…
Shared by Retraction Watch

The U.S. Just Lost 26 Years' Worth of Progress on Life Expectancy
scientificamerican.com
Scientific American is the essential guide to the most awe-inspiring advances in science and technology, explaining how they change our understanding of the world and shape our lives.
Shared by Boback Ziaeian



AI in Medicine Is Overhyped
scientificamerican.com
Scientific American is the essential guide to the most awe-inspiring advances in science and technology, explaining how they change our understanding of the world and shape our lives.
Shared by Darren Dahly, PhD oMG FFs JfC SMDH

Significance magazine - 10 women statisticians and data scientists who helped us understand Covid-19
significancemagazine.com
Significance is an official magazine and website of the Royal Statistical Society (RSS) and the American Statistical Association (ASA).
Shared by Stats & Data Science Views

Accuracy of Practitioner Estimates of Probability of Diagnosis Before and After Testing
jamanetwork.com
This survey study of physicians, nurse practitioners, and physician assistants explores practitioner understanding of diagnostic reasoning.
Shared by Andrew Vickers

Table 1. Survey Responses

Variable	No. (%) of practitioners			
	Maryland and Middle Atlantic states	Oregon and Washington	Texas	All sites
Invited to participate	390	150	183	723
No response	41 (11)	0	16 (9)	57 (8)
Refusals[a]	10 (3)	3 (2)	3 (2)	16 (2)
Not interested	2	0	2	4
Too busy or bad timing	6	2	2	10
Too difficult	2	1	0	3
Other	3	0	0	3
Agreed to participate (of all invited)	339 (87)	147 (98)	164 (90)	650 (90)
Agreed but did not complete survey	27 (7)	23 (15)	15 (8)	65
Total surveys received	312 (80)	124 (83)	149 (81)	585 (81)
Failed to complete ≥1 questions required for final analysis	7 (2)	12 (8)	13 (7)	32 (4)

[a] May list more than 1 reason for refusing to complete the survey.

Science has been in a "replication crisis" for a decade. Have we learned anything?
vox.com
Bad papers are still published. But some other things might be getting better.
Shared by Retraction Watch

Using Instrumental Variable (IV) methods to estimate the effect size in Colonoscopy Screening Trial (NordICC)
sensiblemed.substack.com
It is my pleasure to welcome this guest post by Prof Recht at Berkeley, CA. We have covered the NordICC trial of colonoscopy extensively— because it is big news, and because it matters. I recently in…
Shared by Venk Murthy, MD, PhD

Table S6: Sensitivity analysis: 10-year adjusted per-protocol risks for colorectal cancer incidence and mortality in the estimated 42% of individuals who would attend screening if invited[3].

	Screened group*		Usual care group*		Risk difference (95% CI)	Risk ratio (95% CI)
	CRC cases	10-year risk [%] (95% CI)	CRC cases	10-year risk [%] (95% CI)		
Colorectal cancer incidence	102	0.86 (0.70 to 1.02)	308	1.30 (1.01 to 1.59)	-0.44 (-0.77 to -0.12)	0.66 (0.46 to 0.86)
Colorectal cancer mortality	18	0.14 (0.07 to 0.21)	68	0.20 (0.04 to 0.36)	-0.06 (-0.22 to -0.11)	0.72 (0 to 3.70)

*sub-sample of screening and no-screening groups according to Cuzick et al[3]

When failure to correct a flawed paper could put patients' lives at risk
retractionwatch.com
Robert Speth On April 15, 2021, as COVID-19 was waning several months prior to the surge in deaths associated with arrival of the Delta variant, the journal Cell published an eye-catching paper. Titl…
Shared by Retraction Watch
In 1987, the NIH found a paper contained fake data. It was just retracted.
retractionwatch.com
Ronald Reagan was president and James Wyngaarden was director of the National Institutes of Health when a division of the agency found 10 papers describing trials of psychiatric drugs it had funded h…
Shared by Ivan Oransky
Screening Colonoscopy Misses the Mark in its First Real Test
sensiblemed.substack.com
In today's Study of the Week, Vinay and I offer commentary on one of the most important trials of the year--the NordICC trial of screening colonoscopy.
Shared by Vinay Prasad MD MPH

ORIGINAL ARTICLE

Effect of Colonoscopy Screening on Risks of Colorectal Cancer and Related Death

Michael Bretthauer, M.D., Ph.D., Magnus Løberg, M.D., Ph.D., Paulina Wieszczy, Ph.D., Mette Kalager, M.D., Ph.D., Louise Emilsson, M.D., Ph.D., Kjetil Garborg, M.D., Ph.D., Maciej Rupinski, M.D., Ph.D., Evelien Dekker, M.D., Ph.D., Manon Spaander, M.D., Ph.D., Marek Bugajski, M.D., Ph.D., Øyvind Holme, M.D., Ph.D., Ann G. Zauber, Ph.D., et al., for the NordICC Study Group*

Experts slam Florida surgeon general's warning on coronavirus vaccines
washingtonpost.com

"If you submitted that to a peer-reviewed journal ... it would get rejected," vaccine safety expert says about study on which it is based.

Shared by Boback Ziaeian

Bad Infographics: The Worst Infographics of 2020 - Venngage

venngage.com

With the pandemic, bad infographics are now a regular part of public life, creating confusion and mistrust. Here are the worst infographics from 2020.

Shared by Rafael Irizarry

Statistical Thinking - How to Do Bad Biomarker Research
hbiostat.org
This article covers some of the bad statistical practices that have crept into biomarker research, including setting the bar too low for demonstrating that biomarker information is new, believing tha…
Shared by Frank Harrell

Research

Original Investigation

Toward a Neuroimaging Treatment Selection Biomarker for Major Depressive Disorder

Callie L. McGrath, BA; Mary E. Kelley, PhD; Paul E. Holtzheimer III, MD; Boadie W. Dunlop, MD; W. Edward Craighead, PhD; Alexandre R. Franco, PhD; R. Cameron Craddock, PhD; Helen S. Mayberg, MD

IMPORTANCE Currently, fewer than 40% of patients treated for major depressive disorder achieve remission with initial treatment. Identification of a biological marker that might improve these odds could have significant health and economic impact.

OBJECTIVE To identify a candidate neuroimaging "treatment-specific biomarker" that predicts differential outcome to either medication or psychotherapy.

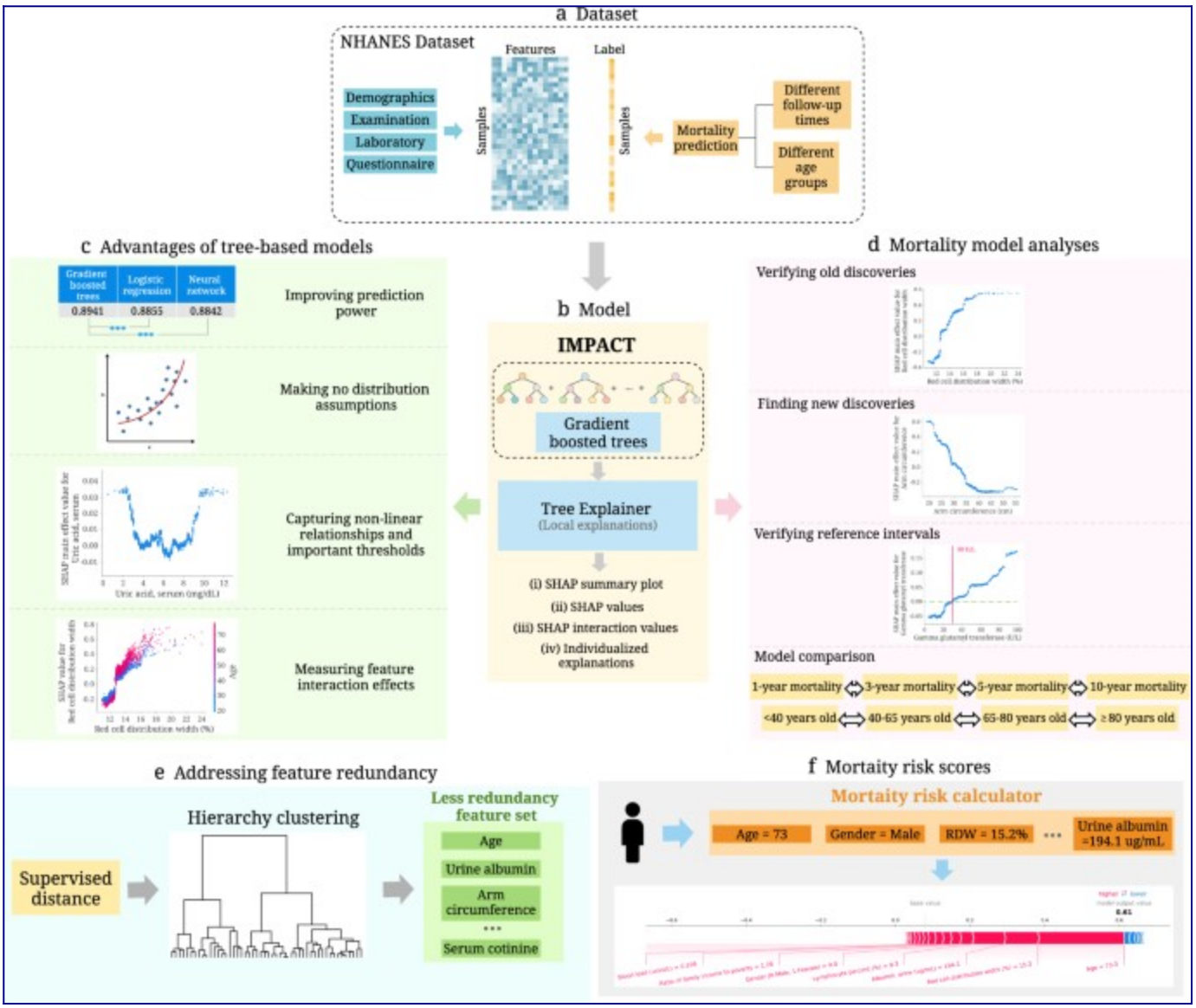
Author Audio Interview at jamapsychiatry.com

Critical appraisal of artificial intelligence-based prediction models for cardiovascular disease
academic.oup.com
Abstract. The medical field has seen a rapid increase in the development of artificial intelligence (AI)-based prediction models. With the introduction of such
Shared by Maarten van Smeden
Reference Collection to push back against "Common Statistical Myths"
discourse.datamethods.org
Note: This topic is a wiki, meaning that this main body of the topic can be edited by others. Use the Reply button only to post questions or comments about material contained in the body, or to sugge…
Shared by Jess Butler
Neuroscientists unravel the mystery of why you can’t tickle yourself
wired.com
New study shows how tickling, playfulness can address key questions about the brain.
Shared by Thomas Lumley

Interpretable machine learning prediction of all-cause mortality
nature.com
Unlike linear models which are traditionally used to study all-cause mortality, complex machine learning models can capture non-linear interrelations and provide opportunities to identify unexplored …

This paper is recommended for how NOT to do a comparative study. It is based on false information about linear models.

Shared by Communications Medicine



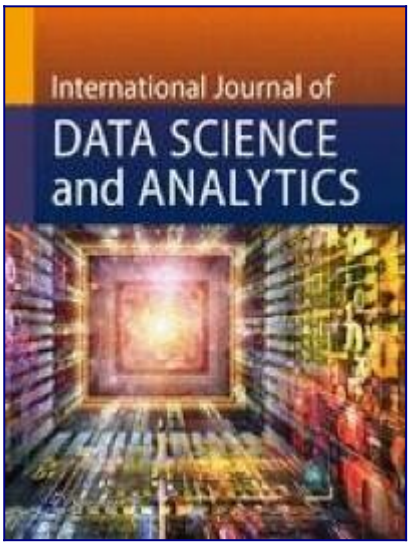
[Observations on Big Data, Precision Health, and Machine Learning](discourse.datamethods.org)
discourse.datamethods.org
I have for some time been pessimistic about the yield to patients and to public health of the incredible amount of funding that has been put into high-dimensional data, electronic health record resea…
Shared by Andrew Vickers

[At N.Y.U., Students Were Failing Organic Chemistry. Who Was to Blame?](nytimes.com)
nytimes.com
Maitland Jones Jr., a respected professor, defended his standards. But students started a petition, and the university dismissed him.

Shared by Boback Ziaeian

What can machines learn about heart failure? A systematic literature review
link.springer.com
This paper presents a systematic literature review with respect to application of data science and machine learning (ML) to heart failure (HF) datasets wit

Shared by Alicja Jasinska Piadlo

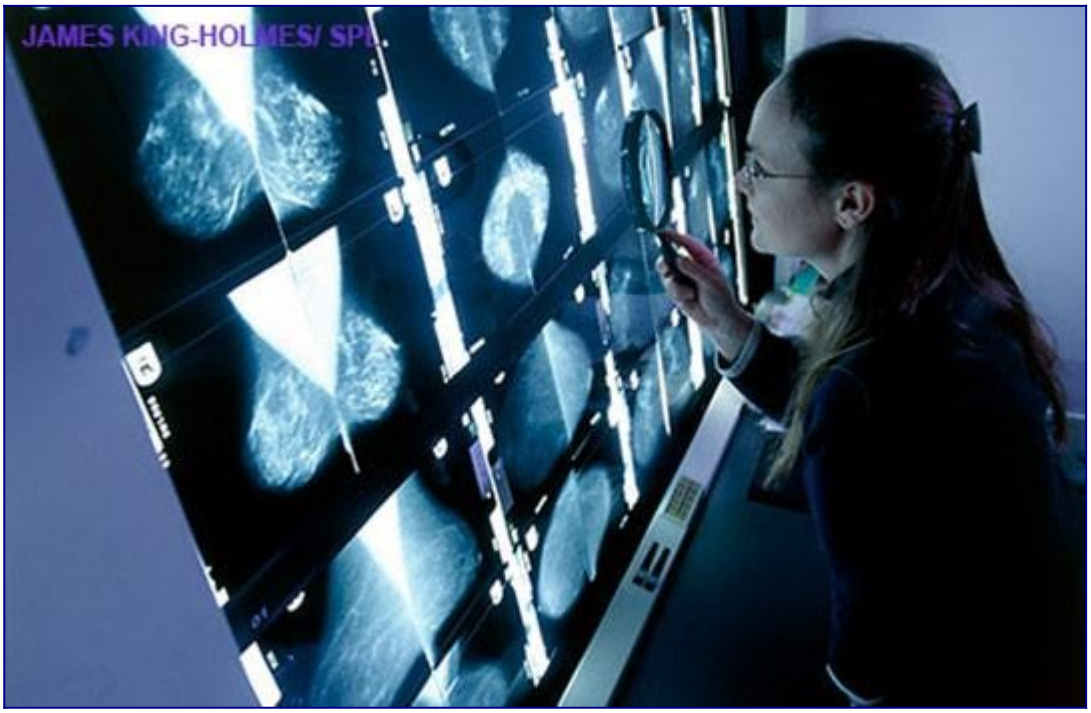


Ordinal Regression as a Model for Signal Detection
blog.msbstats.info
Preface I was basically done with this blog post when I came across Matti Vuorre's post on the same exact topic. Matti goes into all the details, and really the present post can be seen as a brief ac…

Shared by Paul Bürkner

use of baseline covariates in crossover studies

academic.oup.com
Abstract. It is our experience that in many settings, crossover trials that have within-period baseline measurements are analyzed wrongly. A “conventional” anal
Shared by Stephen John Senn
I avoided authorship discussions with collaborators—until I learned some hard lessons
science.org
Proactive conversations help ensure fairness and protect relationships, this researcher found
Shared by Retraction Watch
Why cancer screening has never been shown to “save lives”—and what we can do about it
bmj.com
The claim that cancer screening saves lives is based on fewer deaths due to the target cancer. Vinay Prasad and colleagues argue that reductions in overall mortality should be the benchmark and call …
Shared by Vinay Prasad MD MPH



A Letter to Health Journalism Students
sensiblemed.substack.com
Drs. Alderighi and Rasoini are back with some tips for budding health journalists. Their short essay helps everyone learn key tips and tricks of reading observational studies
Shared by John Mandrola, MD

Book Review: Who's Counting: Uniting Numbers And Narratives With Stories From Pop Culture, Puzzles, Politics, And More

forbes.com

New this September from Prometheus Books, Who's Counting is a collection of original essays and repurposed columns that incorporate some jaw-dropping examples of innumeracy in our 21st century world.

Shared by John Allen Paulos

Language regarding confidence in "negative" results

discourse.datamethods.org

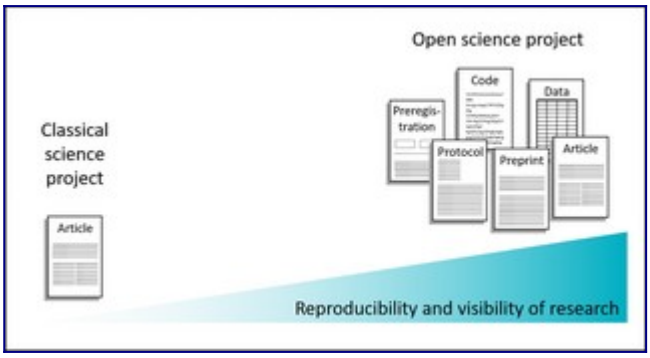
Exposure E is known to be a rare cause of a common condition C, which has many other possible causes. When C is caused by E, this is defined as a new clinical entity EC. To confirm EC an invasive and…

Biostatistics for Biomedical Research - 17 Modeling for Observational Treatment Comparisons
hbiostat.org
The randomized clinical trial is the gold standard for developing evidence about treatment effects, but on rare occasion an RCT is not feasible, or one needs to make clinical decisions while waiting…
Shared by Frank Harrell

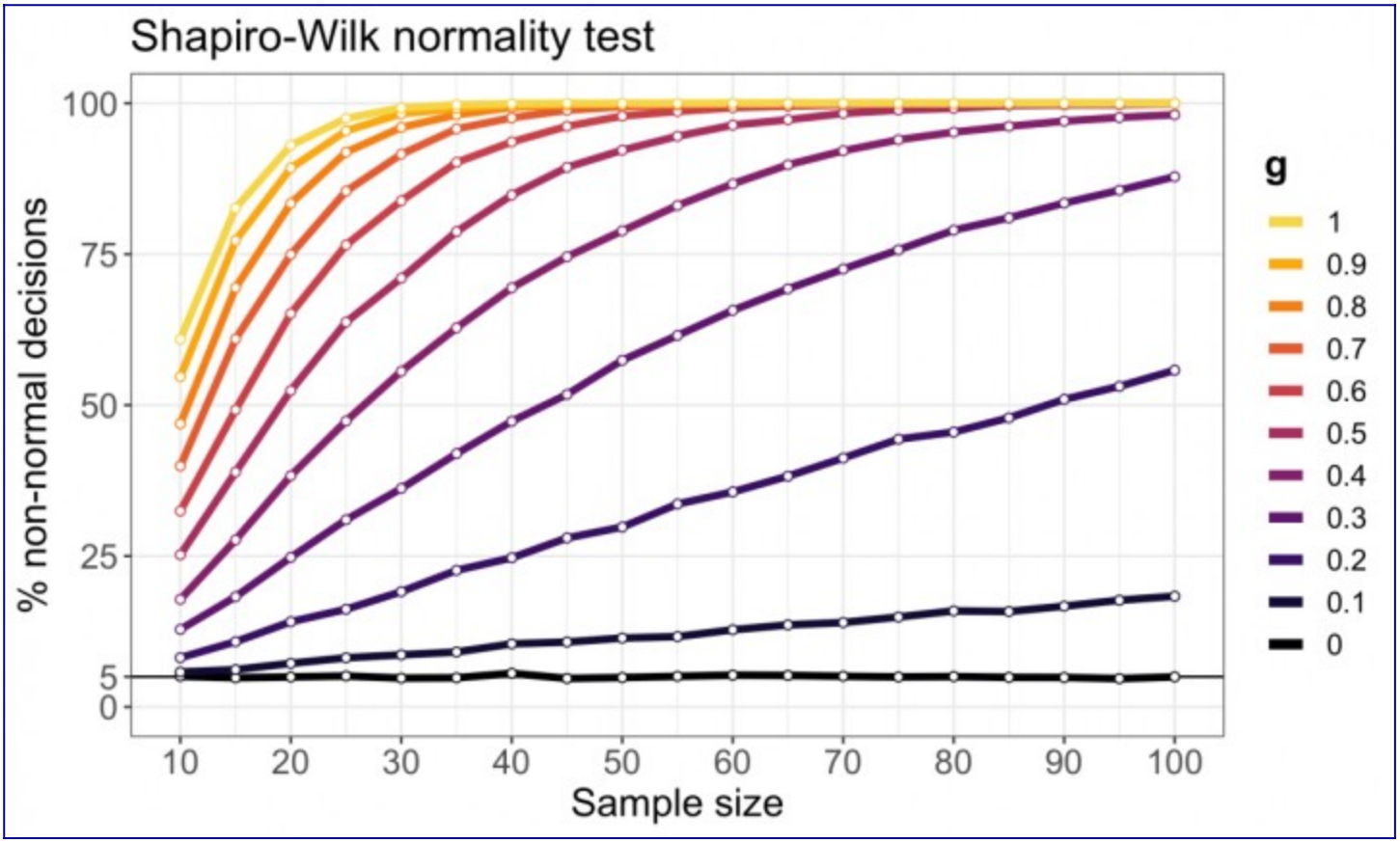
6 Productivity Hacks for Quarto
rstudio.com
With Quarto, you can render plain text and mixed formats into different types of content. We highlight six productivity hacks that may be useful to you.
Shared by RStudio

'A new standard': The woman behind the Johns Hopkins COVID dashboard wins major prize
latimes.com
The creator of the Johns Hopkins COVID-19 dashboard has won the Lasker-Bloomberg Public Service Award.
Shared by Vinay Prasad MD MPH

A guide to open science practices for animal research
journals.plos.org
Open science has become a buzzword in the scientific community that too often misses the practical application for individual researchers. This Essay, provides a guide to choosing the most appropriat…
Shared by Darren Dahly, PhD oMG FFs JfC SMDH



Why normality tests are great…
garstats.wordpress.com
Why normality tests are great as a teaching example and should be avoided in research.
Shared by Andrew Althouse

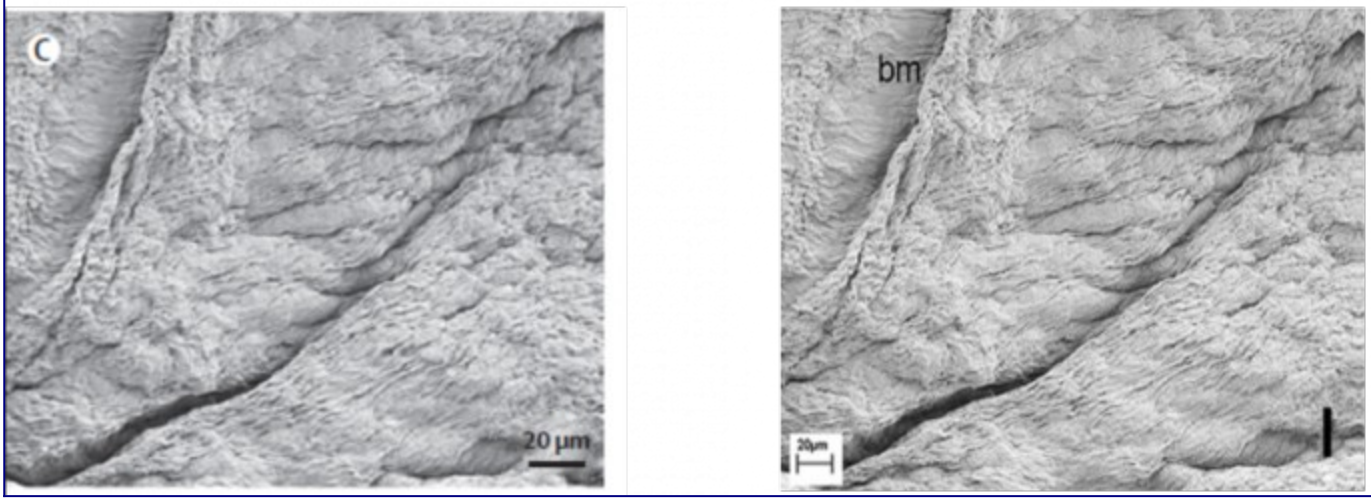


Is the Lancet complicit in research fraud?
drpeterwilmshurst.wordpress.com
This blog was written jointly by Patricia Murray, Professor of Stem Cell Biology and Regenerative Medicine, University of Liverpool, UK and Peter Wilmshurst. The editor of a medical journal that char…
Shared by Retraction Watch

Identical samples in Fig.6C of Gonfiotti et al (2014) and Fig IL of Baiguera et al (2010)

Fig. 6C in Gonfiotti et al (2014). Transmission electron microscopy showing **external surface** of airway.

Fig. 1l in Baiguera et al (2010). Transmission electron microscopy showing **internal surface** of airway.



Study of the Week -Professor Sanjay Kaul Reviews The POST-PCI Trial
sensiblemed.substack.com
Professor Kaul explains why the POST-PCI trial of stress testing vs no stress testing after high-risk coronary stenting failed to answer its question. It's a tour-de-force in critical appraisal.
Shared by John Mandrola, MD

ORIGINAL ARTICLE

Routine Functional Testing or Standard Care in High-Risk Patients after PCI

Duk-Woo Park, M.D., Do-Yoon Kang, M.D., Jung-Min Ahn, M.D., Sung-Cheol Yun, Ph.D., Yong-Hoon Yoon, M.D., Seung-Ho Hur, M.D., Cheol Hyun Lee, M.D., Won-Jang Kim, M.D., Se Hun Kang, M.D., Chul Soo Park, M.D., Bong-Ki Lee, M.D., Jung-Won Suh, M.D., et al., for the POST-PCI Investigators*

Can you explain what these 1,500 papers are doing in this journal?
retractionwatch.com
James Heathers The Internet of Things. Computer science. Botany. COVID-19. All worthwhile subjects, to be sure. But what do they have to do with materials science? That's what James Heathers, who wil…

Shared by Retraction Watch

Adversarial Collaboration: An EDGE Lecture by Daniel Kahneman

edge.org

[EDITOR'S NOTE: In marking this year's 25th anniversary of Edge, we are presenting original lectures from eminent scientists and other thinkers in the empirical world who are changing the way we thin…

Shared by Brian Nosek

Table 2 Fallacy: Presenting and Interpreting Confounder and Modifier Coefficients

academic.oup.com

Abstract. It is common to present multiple adjusted effect estimates from a single model in a single table. For example, a table might show odds ratios for one

Shared by Maarten van Smeden

Doctors' Observations about COVID-19 Used to Be Embraced

sensiblemed.substack.com

While rounding on an inpatient service, Dr. Anil Makam noticed positive news about COVID-19. Many of his online colleagues did not appreciate these upbeat observations

Shared by Vinay Prasad MD MPH

Universities Say They Want More Diverse Faculties. So Why Is Academia Still So White?
fivethirtyeight.com
When she was hired as a professor by Harvard University in 2013, Lorgia García Peña was the only Black Latina on a tenure track in the university's Faculty of A…
Shared by Darren Dahly, PhD oMG FFs JfC SMDH

DELTA2 guidance on choosing the target difference and undertaking and reporting the sample size calculation for a randomised controlled trial
bmj.com
Randomised controlled trials are considered to be the best method to assess comparative clinical efficacy and effectiveness, and can be a key source of data for estimating cost effectiveness. Central…
Shared by Victoria Cornelius

Leveraging electronic health records for data science: common pitfalls and how to avoid them
thelancet.com
Analysis of electronic health records (EHRs) is an increasingly common approach for studying real-world patient data. Use of routinely collected data offers several advantages compared with other stu…
Shared by The Lancet Digital Health

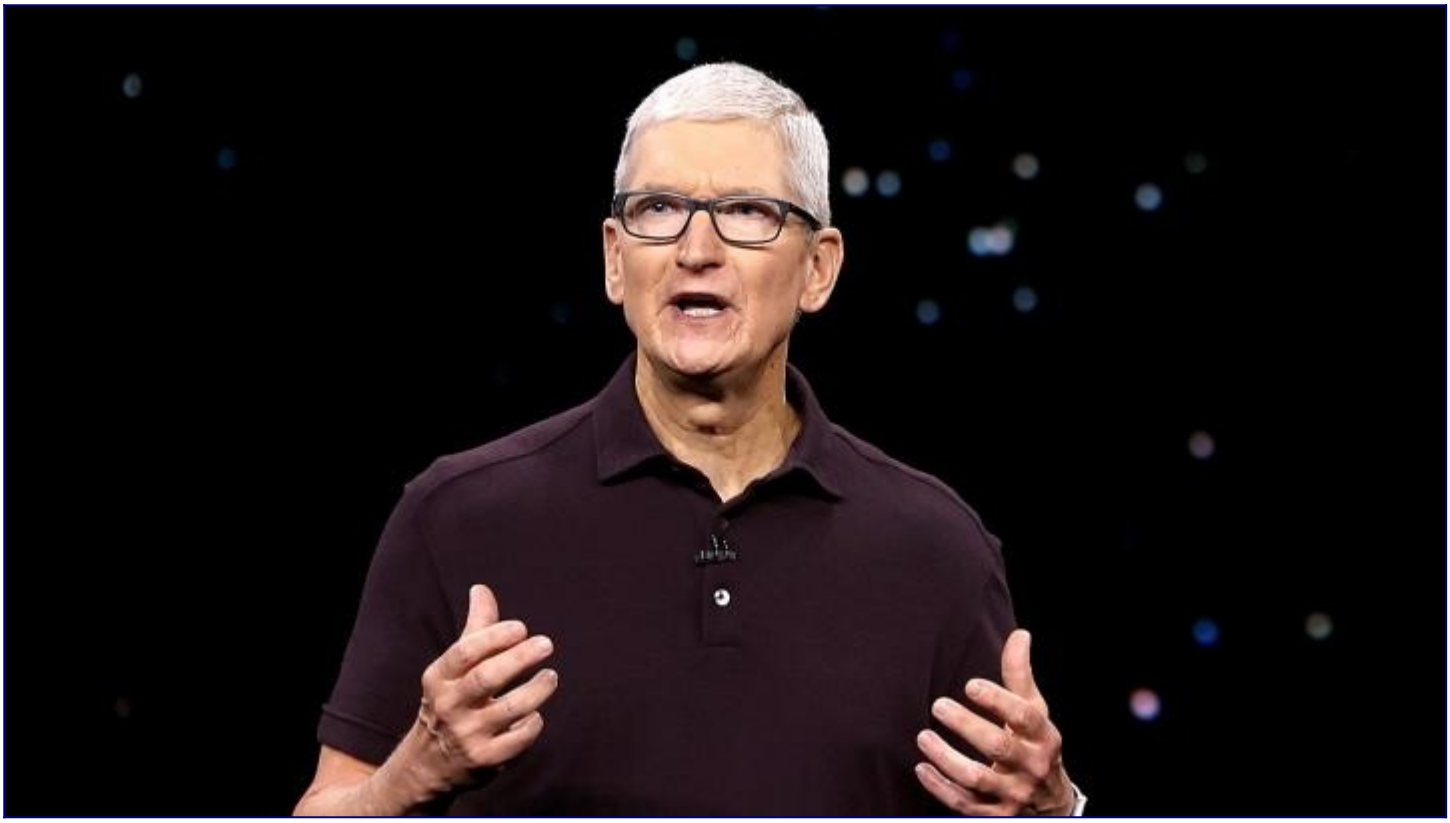
Steps	Identified pitfalls	Potential solutions
Cohort building	1) Sample selection bias	Discuss design choices with an experienced clinician Use gold-standard cohort definitions Perform sensitivity analyses
Definitions of variables	2) Imprecise variable definitions	Develop definitions in a multidisciplinary team, involving clinicians, epidemiologists, statisticians, social scientists, and patients
Feature selection	3) Limitations to deployment	Check if registration time is aligned with result time Only include information that is available at baseline Ensure there is no patient overlap between the training dataset and the test dataset
Feature selection	4) No adjustment for association between frequency of measurements and severity of disease	Consult experts and adjust variables using epidemiological or statistical strategies such as weighting, multiple imputation, or removal of unreliable variables
Study design	5) Subjective treatment allocation affecting causal inference studies	Aim to adjust for interphysician and intraphysician differences in treatment allocation
Study or results validation	6) Model overfitting and reduced generalisability to other settings	Question how local policies might affect generalisability Potentially adjust effect size estimates on the basis of known differences in prevalence

Time doesn’t flow like a river. So why do we feel swept along?
psyche.co
Physics tells us that time doesn’t flow like a river, as Heraclitus claimed. Why then do we feel like we’re swept along?
Shared by Aeon+Psyche

Reviewers award higher marks when a paper’s author is famous
science.org
“Matthew effect” is powerful, unusually large study finds
Shared by Retraction Watch

Overlooked Shortcomings of Observational Studies of Interventions in Coronavirus Disease 2019: An Illustrated Review for the Clinician
academic.oup.com
An illustrated review of shortcomings of observational studies of interventions in COVID-19 with focus on treatment selection bias due to “goals of care” and me
Shared by ID Journal Club

Apple Announces New iPhones Will No Longer Be Compatible With Human Hand
theonion.com
CUPERTINO, CA—With the latest versions of its most popular product scheduled for release at the end of the week, Apple announced Tuesday that iPhones would no longer be compatible with the human hand…
Shared by Boback Ziaeian

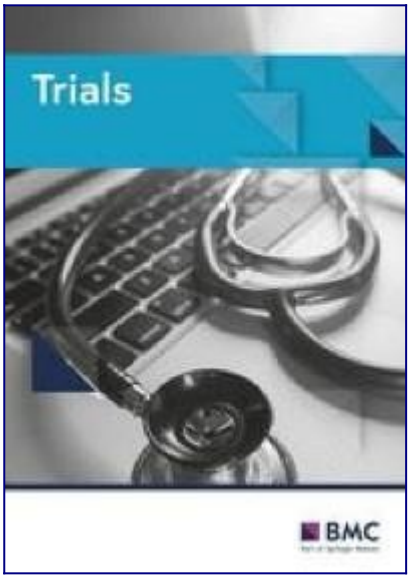


Re-analysis on the statistical sampling biases of a mask promotion trial in Bangladesh: a statistical replication

trialsjournal.biomedcentral.com

A recent randomized trial evaluated the impact of mask promotion on COVID-19-related outcomes. We find that staff behavior in both unblinded and supposedly blinded steps caused large and statisticall…

Shared by Vinay Prasad, MD MPH



You’re Descended from Royalty and So Is Everybody Else

nautil.us

Anybody you can name from ancient history is in your family tree.

Shared by Darren Dahly, PhD oMG FFs JfC SMDH

Consumer group says drugmakers abuse U.S. patent system to keep prices high
reuters.com
Makers of the top selling drugs in the United States are costing patients billions of dollars and worsening a drug pricing crisis by abusing the U.S. patent system to stifle competition and inflate p…

Shared by Boback Ziaeian

Why a cancer ‘moonshot’ is unlikely to find us a cure
washingtonpost.com
Politicians don’t appear to have the stomach for steadily increasing funding for science, which is how breakthroughs happen.

Shared by Vinay Prasad, MD MPH

[“Merchants of doubt” operating within organized science | Statistical Modeling, Causal Inference, and Social Science]
statmodeling.stat.columbia.edu
I came across this post from 2018, where Dorothy Bishop wrote: In Merchants of Doubt, Eric Conway and Naomi Oreskes describe how raising doubt can be used as an effective weapon against inconvenient …
Shared by Andrew Gelman et al.

Statistical Thinking - Randomized Clinical Trials Do Not Mimic Clinical Practice, Thank Goodness
hbiostat.org
Randomized clinical trials are successful because they do not mimic clinical practice. They remain highly clinically relevant despite this.
Shared by Frank Harrell

Paul Brown • 4 years ago

thanks for this piece. I wonder what advocates of "real world evidence" think about a lack of adherence when it's in the control group eg if they're switching to the active group. I saw a clinical trial of daily sunscreen-use where participants randomised to the 'no sunscreen' group stayed indoors or wore long-sleeve t-shirts (no effect of sunscreen was shown for nonmelanoma skin cancer)

^ | v • Reply • Share ›

Frank Harrell Mod ➚ Paul Brown • 4 years ago

That's a good question that has been inadequately discussed. And yesterday's publication of a paper showing that evaluation of a workplace health program from 'real world evidence' is terribly biased when compared with a randomized trial certainly provides food for thought. I hope we get more discussion about all this.

^ | v • Reply • Share ›

Alex Greenfield • 5 years ago

Thank you very much for the great read! In general, your blog posts confirm many of my intuitions/beliefs, but with much more rigor, which I greatly appreciate. In the way of background, I'm a computational biologist currently working in exploratory biomarker analysis of clinical trial data.

Regarding a clinicians view of RCT not being representative - I think in many cases thats due to the biased sample of the population that is present in the RCT. For example, in oncology trials, which is where most of the RCT data I work with comes from, there is a demographic and socioeconomic bias - i.e. only relatively well off people, who live close to centers that run the trial and have the time to get there can afford to do so. Thus, the sample of the population in the RCT is different from the "real-world" population. One of my physician friends described it succinctly as: "I care much less about what happened on average between two groups in an RCT, I care about what will happen for MY two patients".

It seems that what many physicians are asking for is patient-specific predictions of response. It's clear that this is challenging largely due to confounding - we don't know which signals, molecular (i.e. mechanism of action), or clinical, are needed to fully explain response. We can have an entire separate conversation regarding what is "response" and why many current measures of response are flawed.

I would appreciate your comments on what type of data, sample size, and statistical approaches would be required in order to enable a "patient response prediction dashboard"

^ | v • Reply • Share ›

Frank Harrell Mod ➚ Alex Greenfield • 5 years ago

Thanks for the comments and good questions. I've dealt with some of this at http://www.fharrell.com/pos... . To your first point, I hark back to the intent of RCTs: not to estimate how patients in treatment A will fare but to estimate the difference in A and B. You don't need "real world" for that; that's the beauty of RCTs. As in the link just above, if you combine relative effects from RCTs with absolute risk estimates from observational cohort studies, you have what you need. In this sense I think your physician friend is only partially

Nobel Prize winner Gregg Semenza retracts four papers
retractionwatch.com
Gregg Semenza A Johns Hopkins researcher who shared the 2019 Nobel Prize in Medicine or Physiology has retracted four papers from the Proceedings of the National Academy of Sciences (PNAS) for concer…
Shared by Retraction Watch
The Dark Side of Research: Scientific Fraud
davidson.weizmann.ac.il
Elisabeth Bik, who gave up a research career to fight scientific fraud, tells us about the battle for scientific integrity
Shared by Retraction Watch

95) Numbers and Narratives
podcasttheway.com
For decades, New York Times best-selling author John Allen Paulos has enlightened readers by showing how to make sense of the numbers and probabilities behind real-world events, political calculation…
Shared by John Allen Paulos

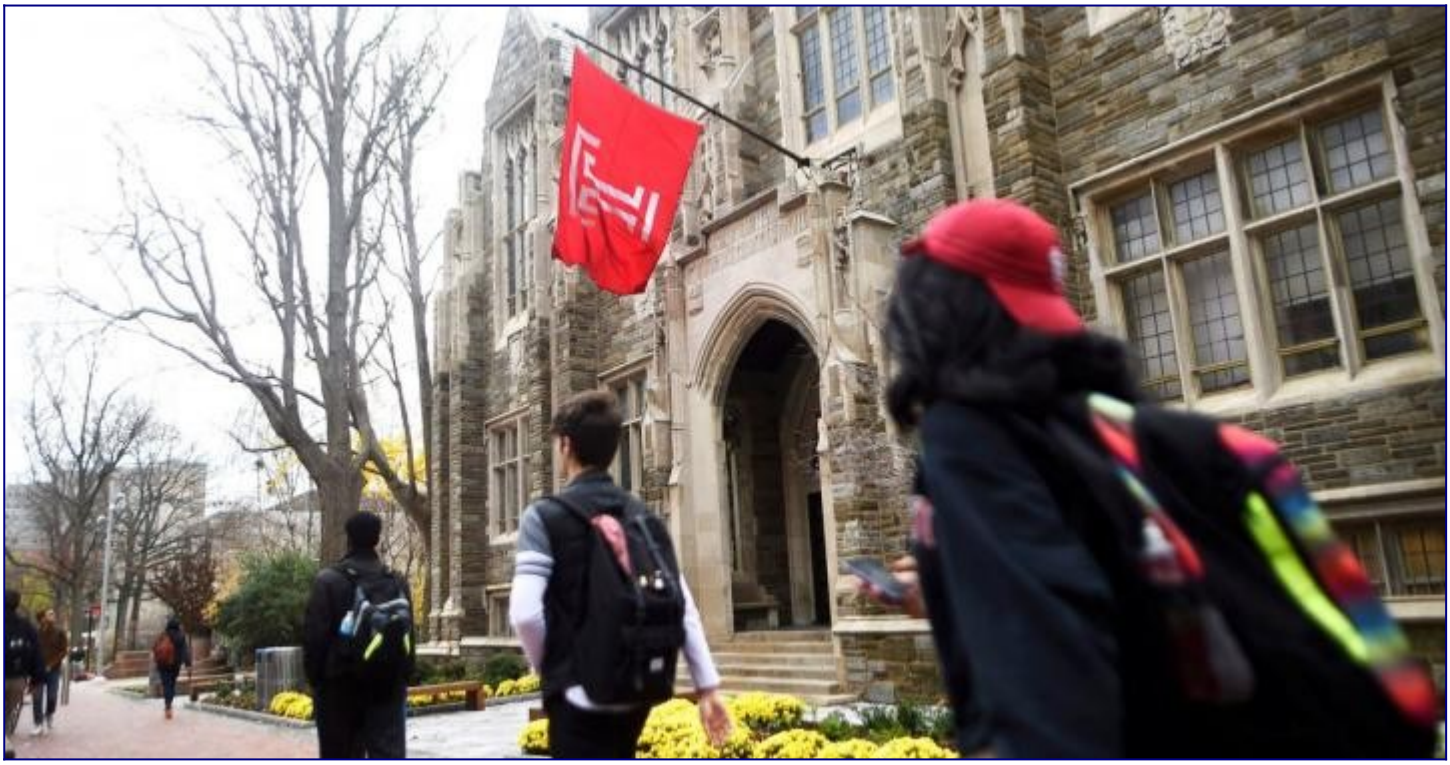
Descriptive Statistics vs Statistical Models

discourse.datamethods.org

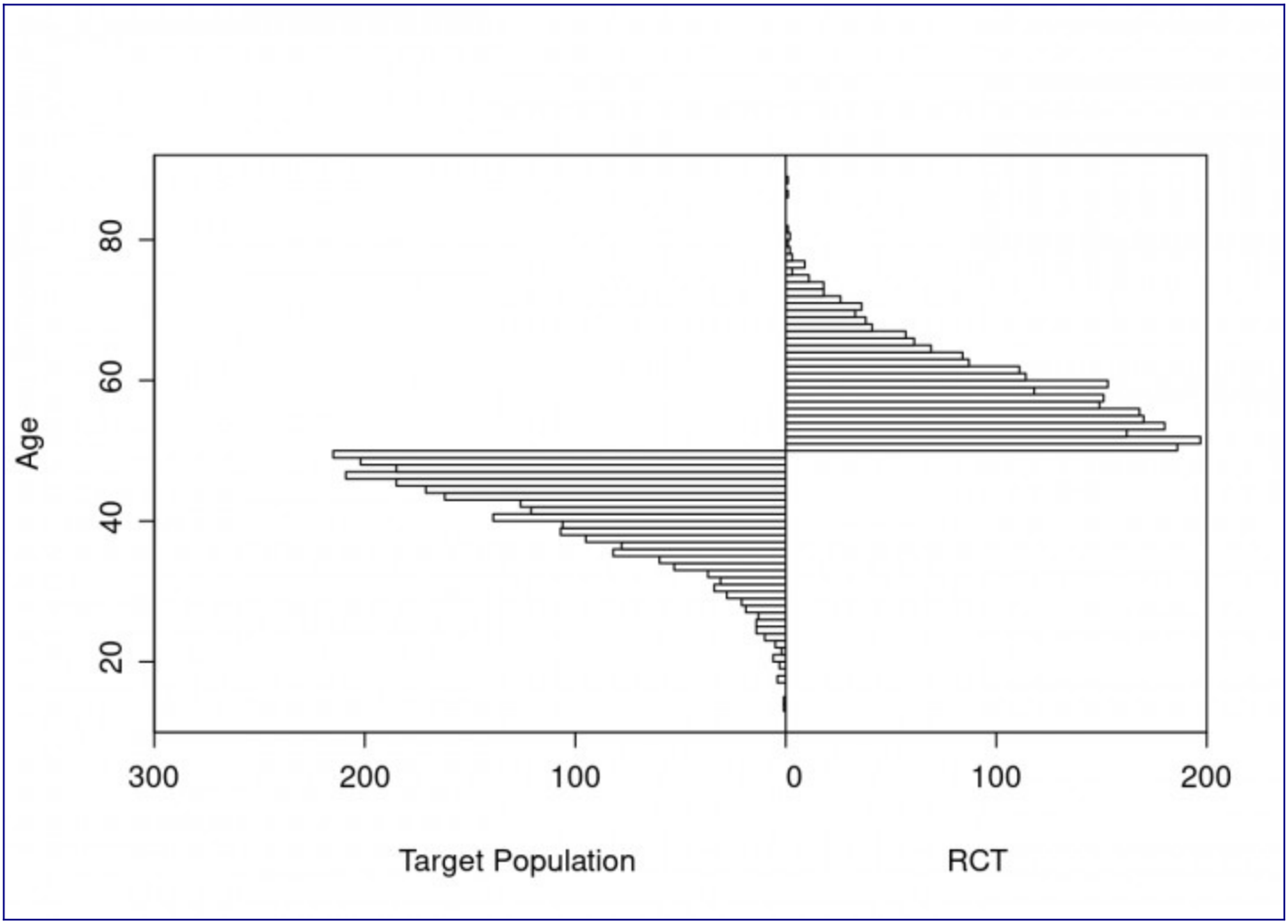
Most of us know when to use simple Statistical Models and when Machine Learning may be suitable. I am curious to know if there are situations where just descriptive statistics is good enough and use …

Planck's principle
en.wikipedia.org
In sociology of scientific knowledge, Planck's principle is the view that scientific change does not occur because individual scientists change their mind, but rather that successive generations of s…
Shared by Venk Murthy, MD, PhD

Exclusive: Medical journals broaden inquiry into potential heart research misconduct
reuters.com
Three medical journals recently launched independent investigations of possible data manipulation in heart studies led by Temple University researchers.
Shared by Boback Ziaeian



Statistical Thinking - Implications of Interactions in Treatment Comparisons
hbiostat.org
This article explains how the generalizability of randomized trial findings depends primarily on whether and how patient characteristics modify (interact with) the treatment effect. For an observatio…
Shared by Frank Harrell

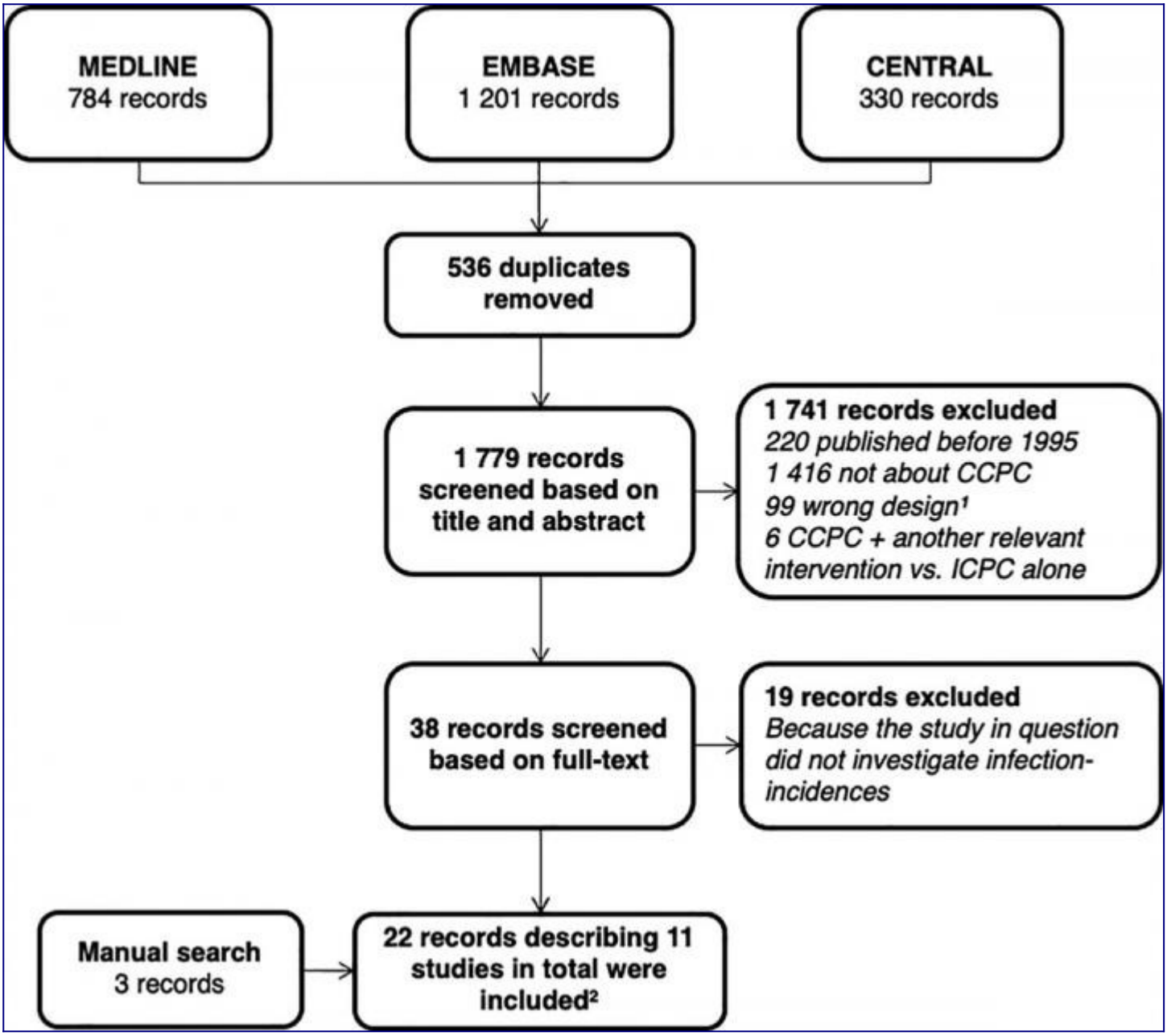


Effectiveness of Continuous Cuff Pressure Control in... : Critical Care Medicine

journals.lww.com

OBJECTIVES: Microaspiration of subglottic secretions is the main pathogenic mechanism for ventila

Shared by SOCAMICYUC



Why nonhuman beings should be granted personhood rights | Aeon Essays

aeon.co

She has deep emotions, complex social needs and a large, elephant brain. Her legal personhood should be recognised too

Shared by Aeon+Psyche

[Critical Appraisal of the DANCAVAS study of Cardiac Screening]
[sensiblemed.substack.com]
[This Study of the Week column is a tour de force of critical appraisal of one the most important trials presented at the recent ESC conference]
Shared by [Vinay Prasad, MD MPH]

[Nobel Prize winner Gregg Semenza retracts four papers]
[retractionwatch.com]

Gregg Semenza A Johns Hopkins researcher who shared the 2019 Nobel Prize in Medicine or Physiology has retracted four papers from the Proceedings of the National Academy of Sciences (PNAS) for concer…

Shared by Retraction Watch

Some models are useful, but how do we know which ones? Towards a unified Bayesian model taxonomy

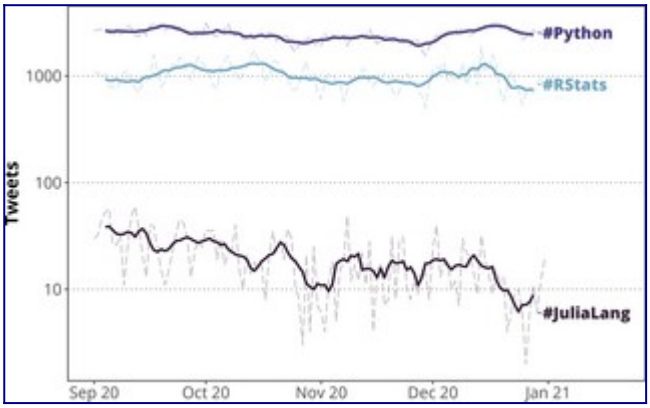
arxiv.org

Probabilistic (Bayesian) modeling has experienced a surge of applications in almost all quantitative sciences and industrial areas. This development is driven by a combination of several factors, inc…

Shared by Paul Bürkner

Bayesian Data Analysis course - Aalto 2022
avehtari.github.io
The first lecture is 5.9.2022 in hall A, Otakaari 1 Aalto 2022 course can be taken online except for the final project presentation. The lectures will be given on campus, but recorded and the recordi…
Shared by Aki Vehtari
Ten simple rules for teaching yourself R
journals.plos.org

Cultivating and maintaining an approach to coding and a system of organization for your projects will help your code to be clear and consistent for you and others. When beginning coding projects, con…
Shared by Gary Collins



Data Literacy as a Tool for Community Health and Social Justice | Amstat News
magazine.amstat.org
Data is one of the world's most valuable resources and an important tool for improving community health. Therefore, knowledge of how to understand, interpret, and share data is necessary for personal…
Shared by ASA

Didier Raoult papers earn expressions of concern as criminal investigation gets underway
retractionwatch.com
Didier Raoult A leading microbiology society has issued expressions of concern for four papers from a group in France led by the controversial scientist Didier Raoult, whose lab is under investigatio…
Shared by Retraction Watch
Exploratory and confirmatory data analysis | Statistical Modeling, Causal Inference, and Social Science

statmodeling.stat.columbia.edu

Valentin Amrhein points us to a recent article, “Exploratory hypothesis tests can be more compelling than confirmatory hypothesis tests,” published in the journal Philosophical Psychology. The articl…

Shared by Andrew Gelman et al.

When an independent replication isn’t really independent

retractionwatch.com

Matt Warman My laboratory at Boston Children’s Hospital and Harvard Medical School studies genetic diseases that affect the skeletal system. We became interested in the protein osteocalcin after Gera…

Shared by Retraction Watch

There Are No Magic Outcome Variables

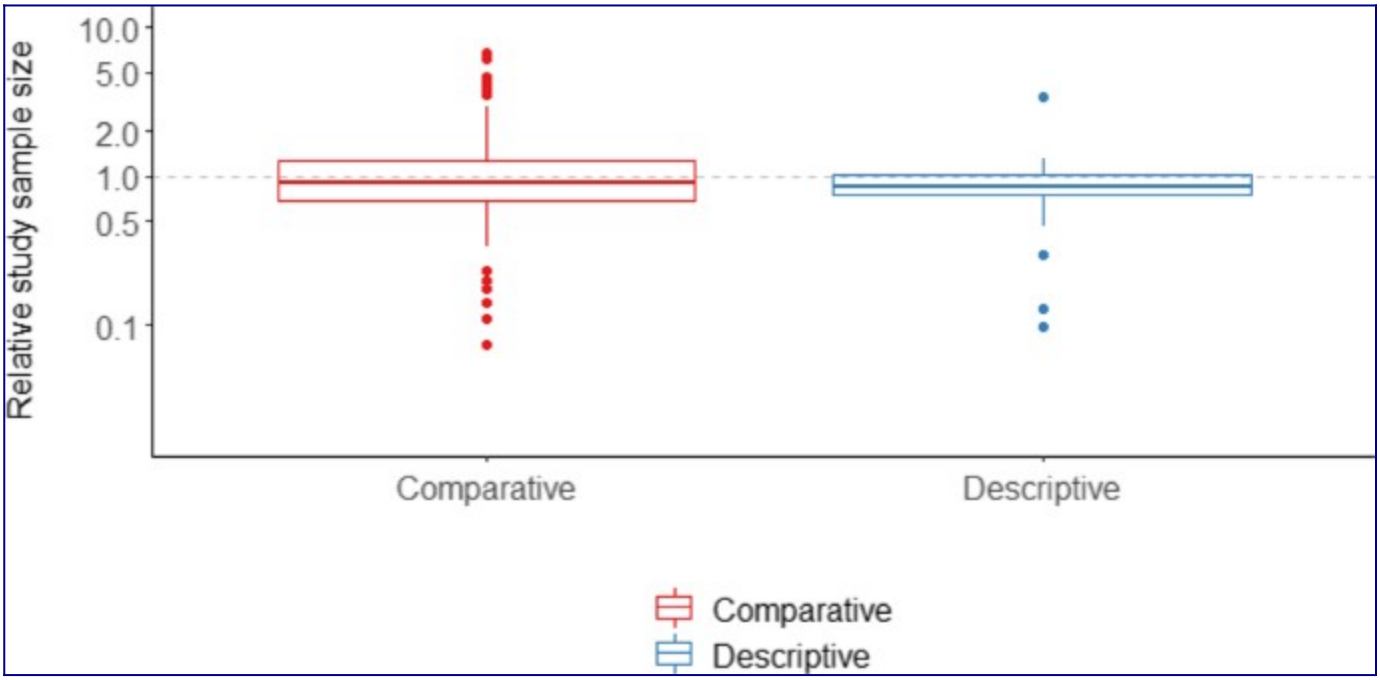
elevanth.org

I'm very sorry but I am going to write about statistics and causal inference again. I'd much rather be doing science. I'll make it brief. I was reading a preprint about the relationship between ki…

Shared by Richard McElreath

[The AAP (American Academy of Pediatrics) is broken, failed organization](#)
sensiblemed.substack.com
[Kids need a new advocate who puts Kids first](#)
Shared by [Vinay Prasad, MD MPH](#)

[Reproducibility of real-world evidence studies using clinical practice data to inform regulatory and coverage decisions]
nature.com
[Studies that generate real-world evidence on the effects of medical products through analysis of digital data collected in clinical practice provide key insights for regulators, payers, and other hea…]
Shared by [Brian Nosek]

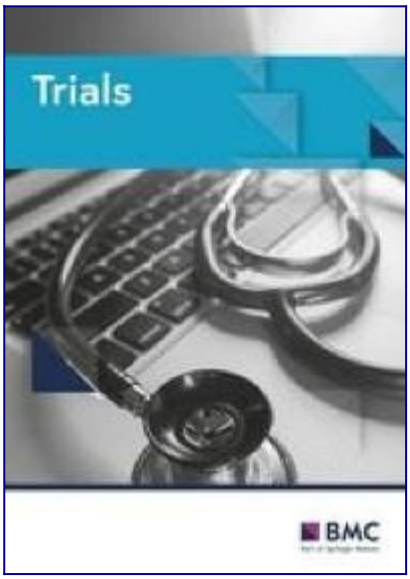


[On model-based time trend adjustments in platform trials with non-concurrent controls]
bmcmedresmethodol.biomedcentral.com

Platform trials can evaluate the efficacy of several experimental treatments compared to a control. The number of experimental treatments is not fixed, as arms may be added or removed as the trial pr…
Shared by Martin Posch

How to Stop the Unknowing Citation of Retracted Papers
pubs.asahq.org
Imagine you were involved in a court case. Your attorney submits a clear and convincing argument that the lawsuit should be dismissed. But it turns out your lawyer has cited precedents that have been…
Shared by Retraction Watch

Principles and recommendations for incorporating estimands into clinical study protocol templates
trialsjournal.biomedcentral.com
Clinical study protocols are the foundation of good clinical studies. Prospective and multidisciplinary collaboration that pays attention to the design of all components of the study protocol can ens…
Shared by Darren Dahly, PhD oMG FFs JfC SMDH



Undue reliance on I 2 in assessing heterogeneity may mislead
bmcmedresmethodol.biomedcentral.com

The heterogeneity statistic I 2, interpreted as the percentage of variability due to heterogeneity between studies rather than sampling error, depends on precision, that is, the size of the studies i…

Shared by Richard Riley (R^2)

Editors-in-chief of aging journal resign en masse after ‘impasse with the Anatomical Society and Wiley’

retractionwatch.com

A journal regarded as the leader in its field is without editors after they resigned as a group earlier this month in a dispute over their workload and compensation. On August 11, the four editors-in…

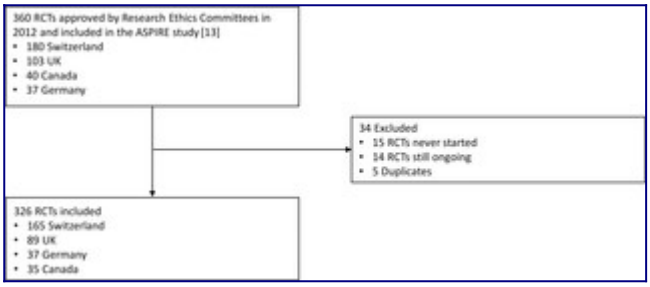
Shared by Retraction Watch

Nonregistration, discontinuation, and nonpublication of randomized trials: A repeated metaresearch analysis

journals.plos.org

Benjamin Speich and colleagues investigate whether rates of trial completion and publication have increased over the past decade, the extent to which non-published trials can be identified in registr…

Shared by Beni



How Important Are Study Designs? A Simulation Assessment of Vaccine Effectiveness Estimation Bias with Time-Varying Vaccine Coverage, and Heterogeneo…

medrxiv.org

medRxiv - The Preprint Server for Health Sciences

Shared by Jing Lian Suah 敬廉

Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal

bmj.com

Objective To review and appraise the validity and usefulness of published and preprint reports of prediction models for prognosis of patients with covid-19, and for detecting people in the general po…

Shared by Gary Collins

Panel: Trump staffers pushed unproven COVID treatment at FDA

apnews.com

WASHINGTON (AP) — Officials in the Trump White House tried to pressure U.S. health experts into reauthorizing a discredited COVID-19 treatment , according to a congressional investigation that provid…

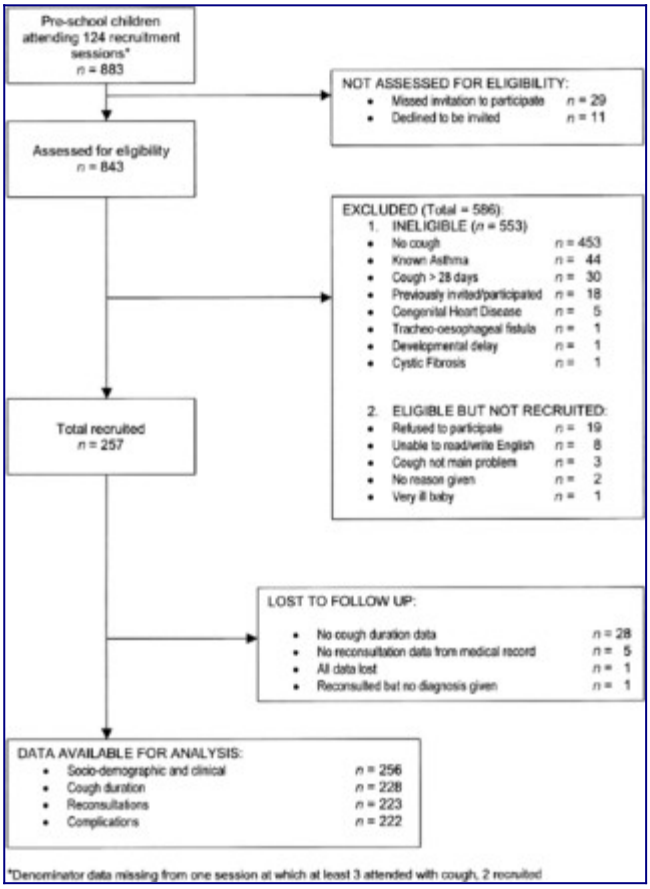
Shared by Boback Ziaeian

What happened when we tried to get a paper claiming ‘billions of lives are potentially at risk’ from COVID-19 vaccines retracted
retractionwatch.com
In February, the editor-in-chief of Food and Chemical Toxicology published an editorial calling for “Papers on potential toxic effects of COVID-19 vaccines.” Following this call, in April 2022, the j…
Shared by Retraction Watch
How a tweet sparked an investigation that led to a PhD student leaving his program
retractionwatch.com
Leslie McIntosh Leslie McIntosh, like many other denizens of Science Twitter, saw a tweet from a pseudonymous account in mid-March that bemoaned a journal’s lack of action after the owner of the acco…
Shared by Ivan Oransky
harm of class imbalance corrections for risk prediction models: illustration and simulation using logistic regression
academic.oup.com
AbstractObjective. Methods to correct class imbalance (imbalance between the frequency of outcome events and nonevents) are receiving increasing interest for de
Shared by Gary Collins
Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration

journals.plos.org

In this explanatory and elaboration document Mattias Egger and colleagues provide the meaning and rationale of each checklist item on the STROBE Statement.

Shared by Charlotte Jackson



Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies

bmj.com

Poor reporting of research hampers assessment and makes it less useful. An international group of methodologists, researchers, and journal editors sets out guidelines to improve reports of observatio…

Shared by Gary Collins

The PRISMA 2020 statement: an updated guideline for reporting systematic reviews
bmj.com
The Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement, published in 2009, was designed to help systematic reviewers transparently report why the review was done, w…
Shared by Gary Collins

Tracking down a Real Data Set(tm)

notstatschat.rbind.io
As you may have noticed, I've been writing software for handling multiple-response data: a package for data manipulation and another one for modelling. Ivy Liu sent me two data sets she had used in a…
Shared by Thomas Lumley
A tale of (3)2 retraction notices: On publishers, paper mill products, and the sleuths that find them
retractionwatch.com
Should publishers acknowledge the work of sleuths when their work has led to retractions? We were prompted to pose the question by a recent retraction from International Environmental Agreements: Pol…
Shared by Retraction Watch
Can I use prevalence estimate to update predictive model intercept to achieve calibration-in-the-large in a different population?
discourse.datamethods.org
When applied to a different population than that where it was developed, it is common for prediction models to show bad calibration-in-the-large, meaning that the average predicted probability doesn't…

'It's time to devise a more efficient solution': Science editor in chief wants to change the retraction process
retractionwatch.com
Holden Thorp On the heels of a high-profile retraction that followed deep investigations by the Science news team, Holden Thorp, the editor in chief of the journal, says it's time to improve the proc…
Shared by Retraction Watch
A Very Short List of Common Pitfalls in Research Design, Data Analysis, and Reporting
journals.stfm.org
Performing scientific research without falling victim to one of the many research design, analysis, and reporting pitfalls can be challenging. As a medical statistician with research experience in a …
Shared by Maarten van Smeden

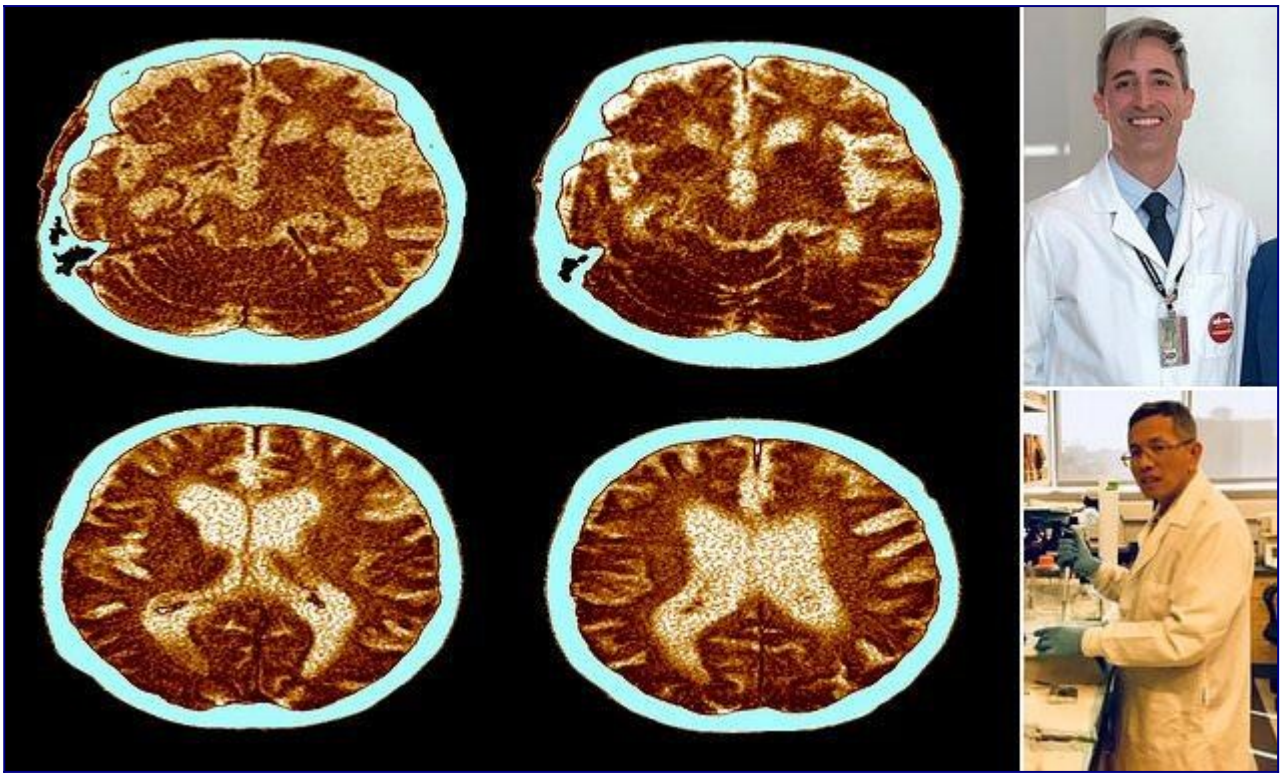
Table 1: Glossary of Terms

Term	Definition
Regression to the mean	When an observation is made that is distanced far from the mean value, the next observation is likely to be closer to the mean. Within individuals, a shift towards the mean can easily be confused for the effect of, for instance, a drug therapy.
Causal inference methodology	Research methodology that aims to structure analyses, identify, and quantify the evidence for one or more causal effects in data. Often applied in nonexperimental data.
Overfitting	The situation where idiosyncrasies in the data at hand are confused for generalizable associations or patterns. The risk of overfitting can be notably high when the number of model parameters is high relative to the size of the data set.
Imprecision	Lack of precision in the estimation, eg, reflected by wide confidence intervals for parameter(s) of interest.
Power	Statistical power often relates to a (binary) hypothesis test. It is the probability that the test correctly rejects the null hypothesis when a specific alternative hypothesis is true.
Initial data analysis	The process of data inspection after data have been collected but before formal statistical analyses.
Data dredging	Performing many analyses to find associations or patterns in data and subsequently ignoring the analyses in which no evidence was found when reporting the study. High risk of false positive findings.
Dichotomania	The tendency to dichotomize variables (eg, into "high" and "low" values) that were original measured on a continuous or a categorical scale with more than two categories.
Noisy data fallacy	The misconception that in data with measurement error associations are always attenuated (ie, biased towards less extreme effects).
Multiple imputation	A popular and flexible approach to deal with missing values in analyses, for which it is generally well known in which circumstances it leads to valid analyses and which circumstances it does not.
Point-estimate-is-the-effect-ism	The tendency to ignore the uncertainty of an estimated association and focus only on a single-point estimate to evaluate the strength of the association.
Confounding	In causal inference studies that evaluate the relation between a variable of interest (eg, a treatment) and an outcome, confounding can be caused by other variables that are a "common cause" of the variable of interest and outcome. Failure to statistically correct (eg, through multivariable regression) for important confounders can lead to residual confounding bias.
Collider	In causal inference studies that evaluate the relation between a variable of interest (eg, a treatment) and an outcome, a collider is a variable that is caused by the variable of interest and outcome. Unlike confounders, statistically correcting for colliders can introduce bias rather than solve it.
Multivariate regression	Regression with an outcome with multiple outcomes (eg, as in multivariate analysis of variance, MANOVA). Often confused with multivariable regression (multiple covariates in the regression).
Table 2 fallacy	The misconception that the associations between confounders and the outcome can be interpreted as valid estimates of causal associations between each confounder and the outcome.

Bayesian inference continues to completely solve the multiple comparisons problem
statmodeling.stat.columbia.edu
I saw you re-posted your Bayes-solves-multiple-testing demo. Thanks for linking to my paper in the PPS! I think it would help people's understanding if you explicitly made the connection with your ob…
Shared by Andrew Gelman et al.

Bring Your Own Binary Packages with RSPM
rstudio.com
RStudio Package Manager 2022.07 now lets you publish your own binary packages, saving you time and frustration.
Shared by RStudio

'Fraudulent' Alzheimer's study gave false hope to families
dailymail.co.uk
It's alleged a pivotal study on the cause of Alzheimer's disease contained manipulated results, potentially leading other scientists down a blind alley.
Shared by Retraction Watch



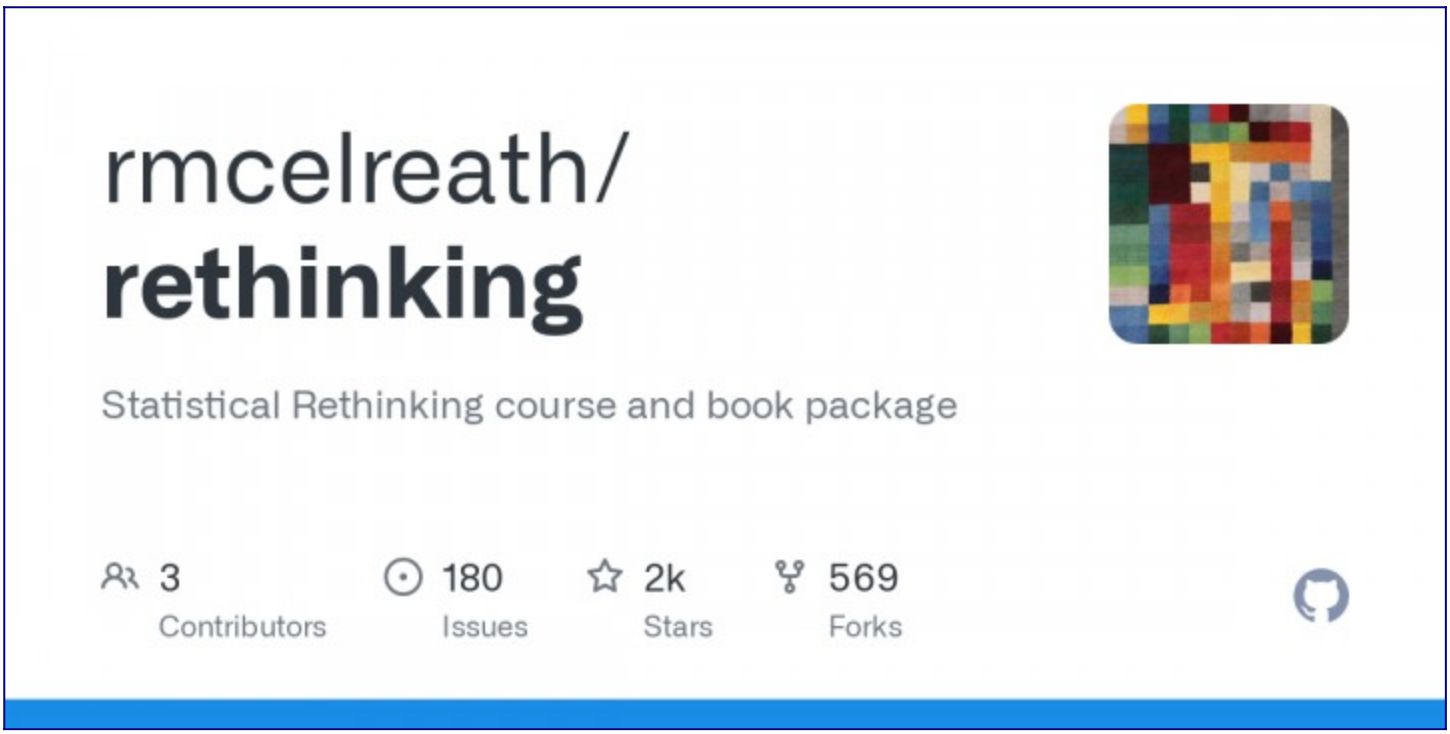
Statistically Efficient Ways to Quantify Added Predictive Value of New Measurements | Statistical Thinking
fharrell.com
Researchers have used contorted, inefficient, and arbitrary analyses to demonstrated added value in biomarkers, genes, and new lab measurements. Traditional statistical measures have always been up t…
Shared by Frank Harrell
S. Senn: Randomisation is not about balance, nor about homogeneity but about randomness (Guest Post)
errorstatistics.com
Stephen Senn Consultant Statistician Edinburgh The intellectual illness of clinical drug evaluation that I have discussed here can be cured, and it will be cured when we restore intellectual primacy …
Shared by Stephen John Senn

	Period				
Sequence	1	2	3	4	5
1	B	D	A	C	G
2	A	C	G	B	F
3	G	B	F	A	E
4	E	G	D	F	C
5	C	E	B	D	A
6	F	A	E	G	D
7	D	F	C	E	B
8	C	G	A	E	F
9	B	F	G	D	E
10	A	E	F	C	D
11	F	C	D	A	B
12	D	A	B	F	G
13	G	D	E	B	C
14	E	B	C	G	A
15	D	C	A	G	E
16	C	B	G	F	D
17	B	A	F	E	C
18	G	F	D	C	A
19	E	D	B	A	F
20	A	G	E	D	B
21	F	E	C	B	G

Company’s Alzheimer’s treatment study earns a flag
retractionwatch.com
Paul Sanberg A journal has issued an expression of concern for a federally-funded paper on Alzheimer’s disease after a sleuth on PubPeer noted potentially duplicated figures in the article. We should…
Shared by Ivan Oransky

Dubious findings about food from the ‘Nutrition Researchers Guild’: How can we learn from this unscientific manipulation of statistics?
geneticliteracyproject.org
Are you confused about conflicting “research” findings on certain foods’ effects on our health? It would hardly be surprising. First, butter is the enemy; then, it’s solid margarine. Is caffeine good…
Shared by MyMartialArtsTV

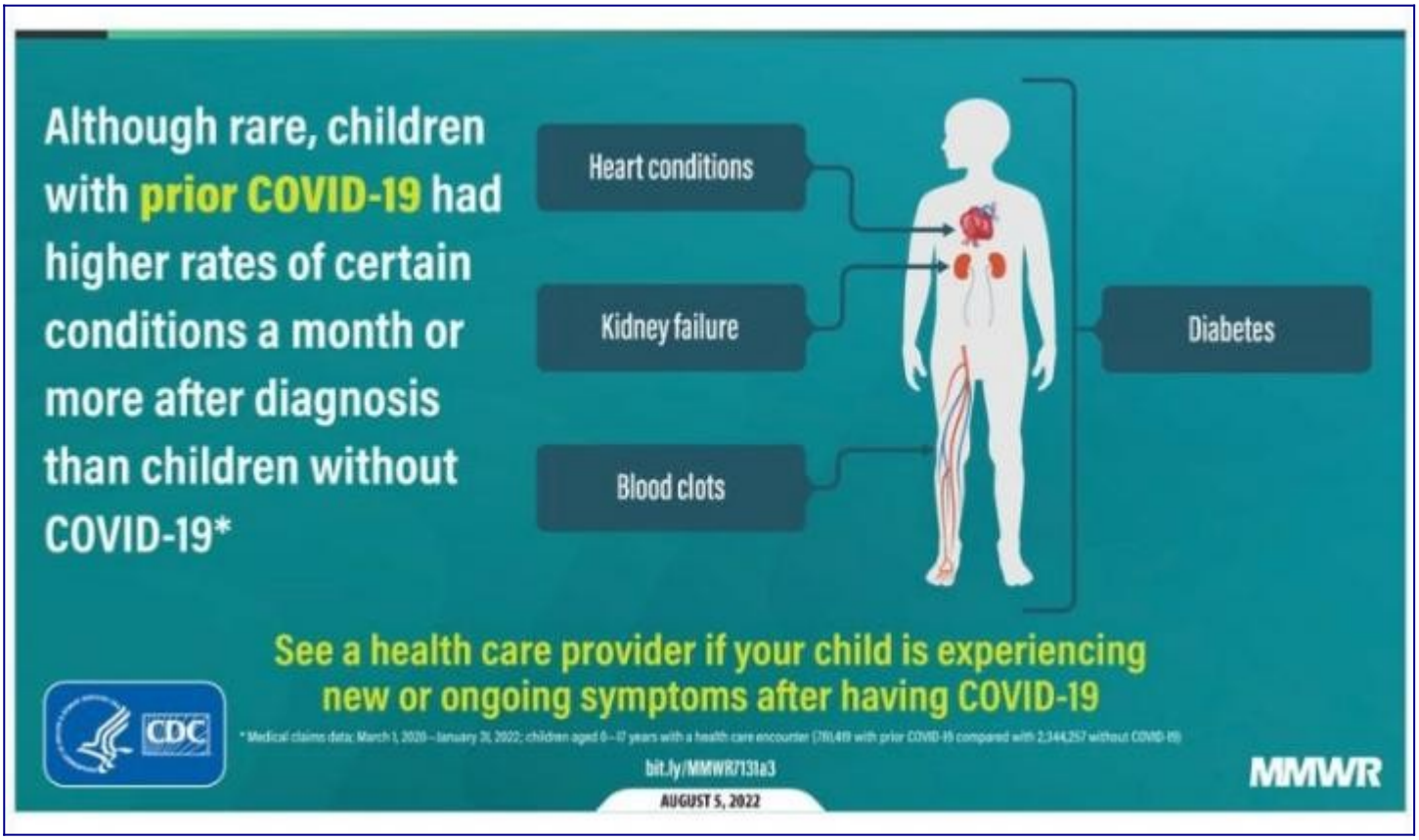
GitHub - rmcelreath/rethinking at Experimental
github.com
Statistical Rethinking course and book package. Contribute to rmcelreath/rethinking development by creating an account on GitHub.
Shared by Richard McElreath



Exclusive: PLOS ONE to retract more than 100 papers for manipulated peer review
retractionwatch.com
In March, an editor at PLOS ONE noticed something odd among a stack of agriculture manuscripts he was handling. One author had submitted at least 40 manuscripts over a 10-month period, much more than…
Shared by Boback Ziaeian

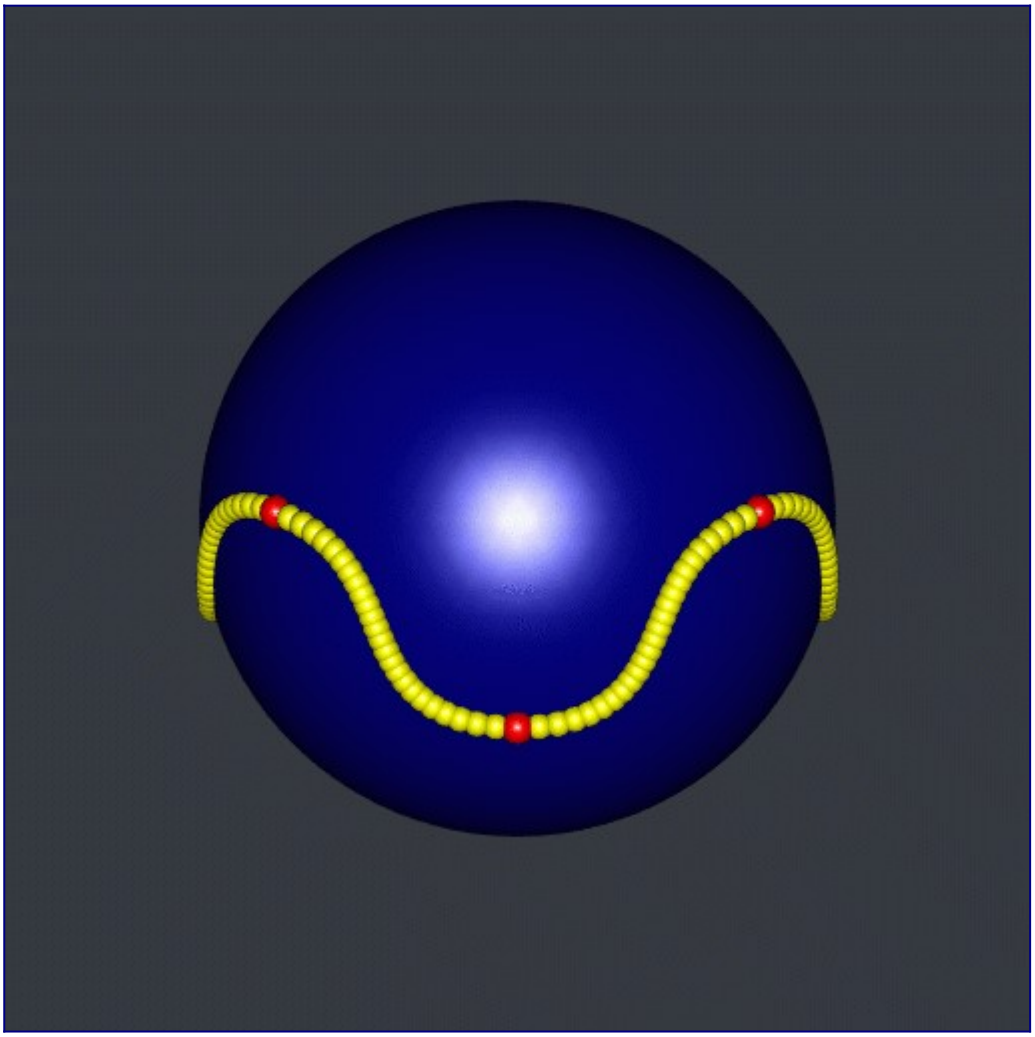
Simply Statistics: Is Code the Best Way to Represent a Data Analysis?
simplystatistics.org
Code is a useful representation of a data analysis for the purposes of transparency and openness. But code alone is often insufficient for evaluating the quality of a data analysis and for determini…
Shared by Roger D. Peng

Another bad CDC study on Long COVID in Kids
vinayprasadmdmph.substack.com
If you tweet bad science long enough; I have to conclude you are a bad scientist.
Shared by Vinay Prasad, MD MPH



Quarto
marketplace.visualstudio.com
Extension for Visual Studio Code - Extension for the Quarto scientific and technical publishing system.
Shared by Quarto
If the outcome is that rare, then nothing much can be learned from pure statistics.
statmodeling.stat.columbia.edu
You may have heard of this recent controversial study on the efficacy of colchicine to reduce the number of hospitalisations/deaths due to covid. It seems to be the opposite of the pattern usually re…
Shared by Andrew Gelman et al.
June 2022: "Top 40" New CRAN Packages
rviews.rstudio.com
An R community blog edited by RStudio

Shared by RStudio



Exclusive: Astronaut Jessica Watkins Becomes First Person To Publish A Science Paper From Space
iflscience.com
We spoke to Dr Watkins, no stranger to making history, about doing it again.
Shared by Retraction Watch

More than a year ago, an editor agreed a paper should be retracted. It hasn’t been.
retractionwatch.com
Eighteen months after the editor in chief of a Springer Nature journal received allegations of plagiarism – and more than a year after the editor apparently decided to retract it – the article remain…
Shared by Ivan Oransky

An editor invited me to submit a commentary, then he rejected it – and named and blamed me in an editorial
retractionwatch.com
Brad Rodu The American Journal of Public Health is the flagship publication of the American Public Health Association, which has more than 25,000 members worldwide. The AJPH boasts that it is “a high…
Shared by Retraction Watch

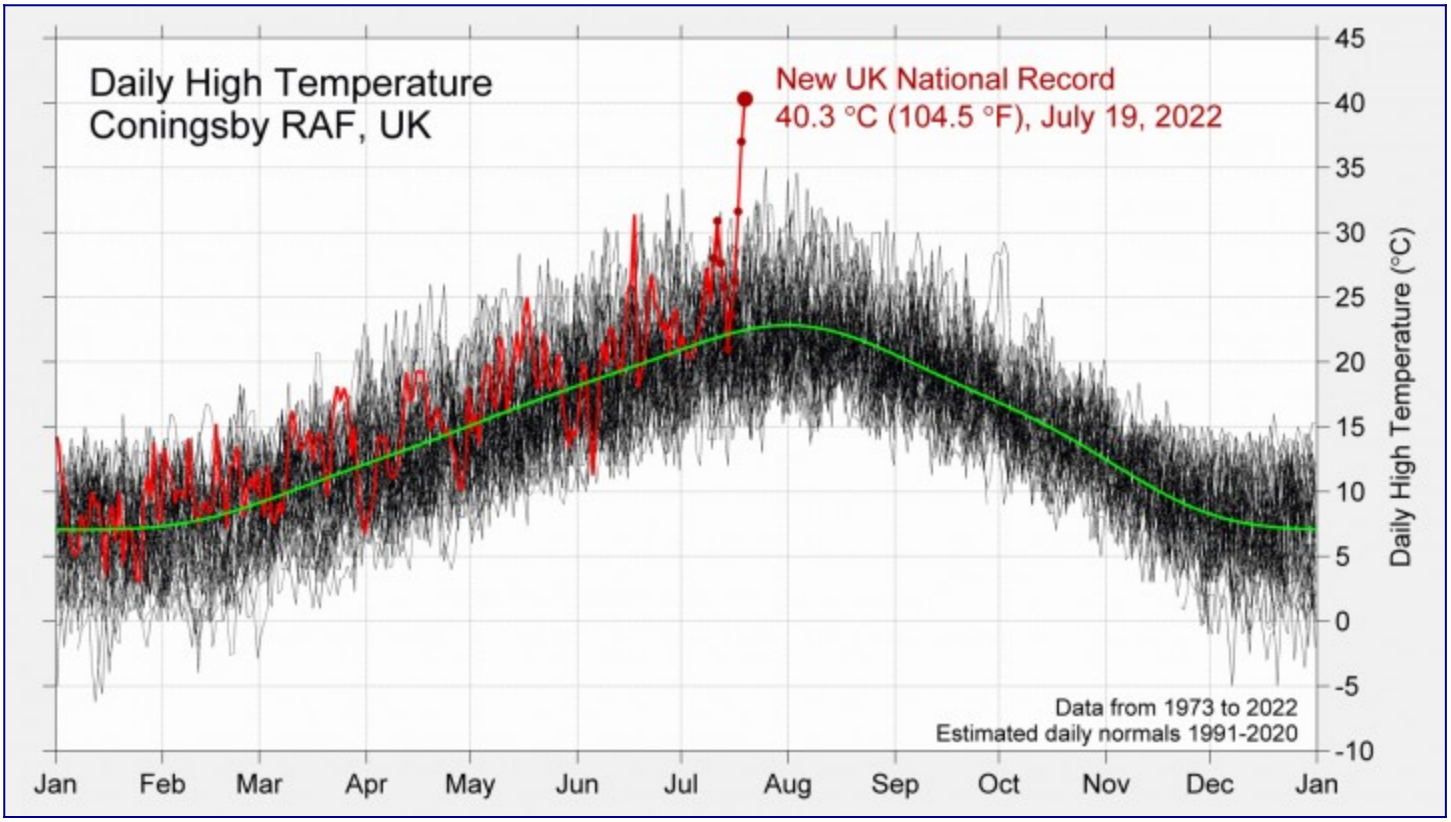
What an Alzheimer’s Controversy Reveals About the Pressures of Academia
theatlantic.com
A prominent research paper is under review for possible fraud. Why is it so hard to correct the record?
Shared by Retraction Watch

Announcing Quarto, a new scientific and technical publishing system
rstudio.com
Today we're excited to announce Quarto, the next generation of R Markdown.
Shared by RStudio

Fisher, Bradford Hill, and randomization
academic.oup.com
In the 1920s RA Fisher presented randomization as an essential ingredient of his approach to the design and analysis of experiments, validating significance tes
Shared by John Mandrola, MD

Target Validity and the Hierarchy of Study Designs
academic.oup.com
Abstract. In recent years, increasing attention has been paid to problems of external validity, specifically to methodological approaches for both quantitative
Shared by huerta.rodrigo

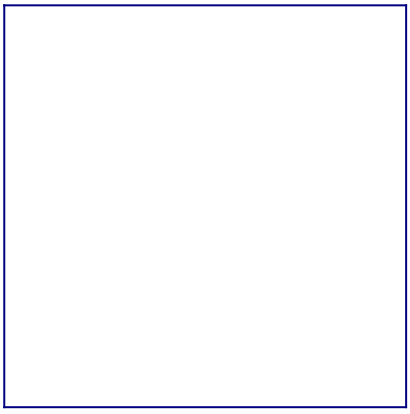
Analysis | Britain’s freakish heat demolished records. Here’s what happened.
washingtonpost.com
The magnitude and extent of Tuesday’s record temperatures underscored the exceptional nature of the European heat wave.
Shared by Edward Tufte



[Risk Difference with: Normal likelihood + identity link + robust variance estimator?]
discourse.datamethods.org
[In this interesting paper, authors argued: For the risk difference, one may use a GLM with a Gaussian (i.e., normal) distribution and identity link function, or, equivalently, an ordinary least squar…]

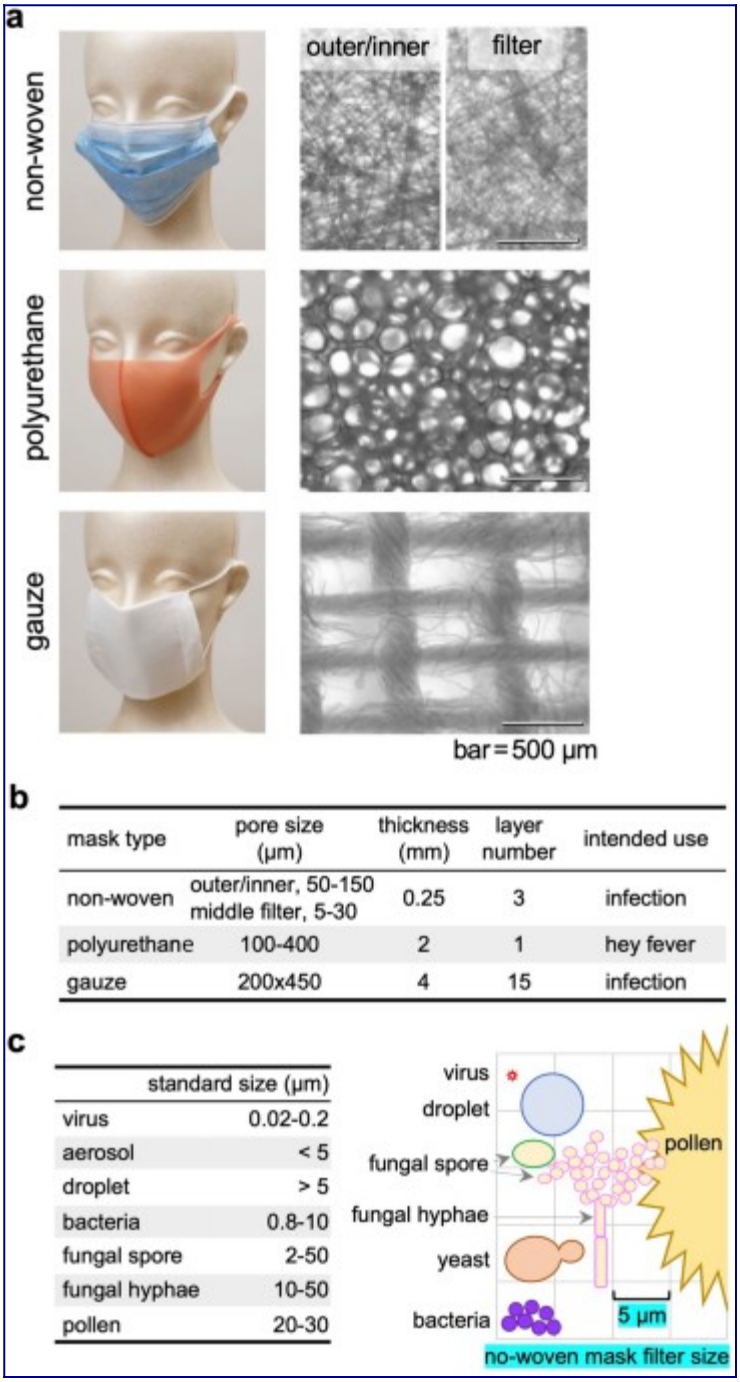
[Evaluating methods for Lasso selective inference in biomedical research: a comparative simulation study]
bmcmedresmethodol.biomedcentral.com
[Variable selection for regression models plays a key role in the analysis of biomedical data. However, inference after selection is not covered by classical statistical frequentist theory, which assu…]
Shared by [Georg Heinze]

Release of CmdStan 2.29
blog.mc-stan.org
We are very happy to announce that the 2.29.0 release of CmdStan is now available on Github! As usual, the release of CmdStan is accompanied by new releases of Stan Math, core Stan and Stanc3. A PySt…
Shared by Aki Vehtari



Papers in Croce case with “blatantly obvious” problems still aren’t retracted after misconduct investigation: sleuth
retractionwatch.com
Carlo Croce This week, Nature reported on two institutional reports that found scientists in Carlo Croce’s cancer research lab at The Ohio State University had committed research misconduct including…
Shared by Retraction Watch

Bacterial and fungal isolation from face masks under the COVID-19 pandemic
nature.com
The COVID-19 pandemic has led people to wear face masks daily in public. Although the effectiveness of face masks against viral transmission has been extensively studied, there have been few reports …
Shared by Vinay Prasad, MD MPH



mask type	pore size (μm)	thickness (mm)	layer number	intended use
non-woven	outer/inner, 50-150 middle filter, 5-30	0.25	3	infection
polyurethane	100-400	2	1	hey fever
gauze	200x450	4	15	infection

	standard size (μm)
virus	0.02-0.2
aerosol	< 5
droplet	> 5
bacteria	0.8-10
fungal spore	2-50
fungal hyphae	10-50
pollen	20-30

Comparative analysis of retracted pre-print and peer-reviewed articles on COVID-19
medrxiv.org
medRxiv - The Preprint Server for Health Sciences
Shared by Retraction Watch

Think about it – your existence is utterly astonishing | Aeon Essays

aeon.co

Life can be better appreciated when you remember how wonderfully and frighteningly unlikely it is that you exist at all

Shared by Aeon+Psyche

Correlation without a cause: an epidemiological odyssey

academic.oup.com

Background. In the 1980s debate intensified over whether there was a protective effect of high-density lipoprotein cholesterol (HDL-C) or an adverse effect of t

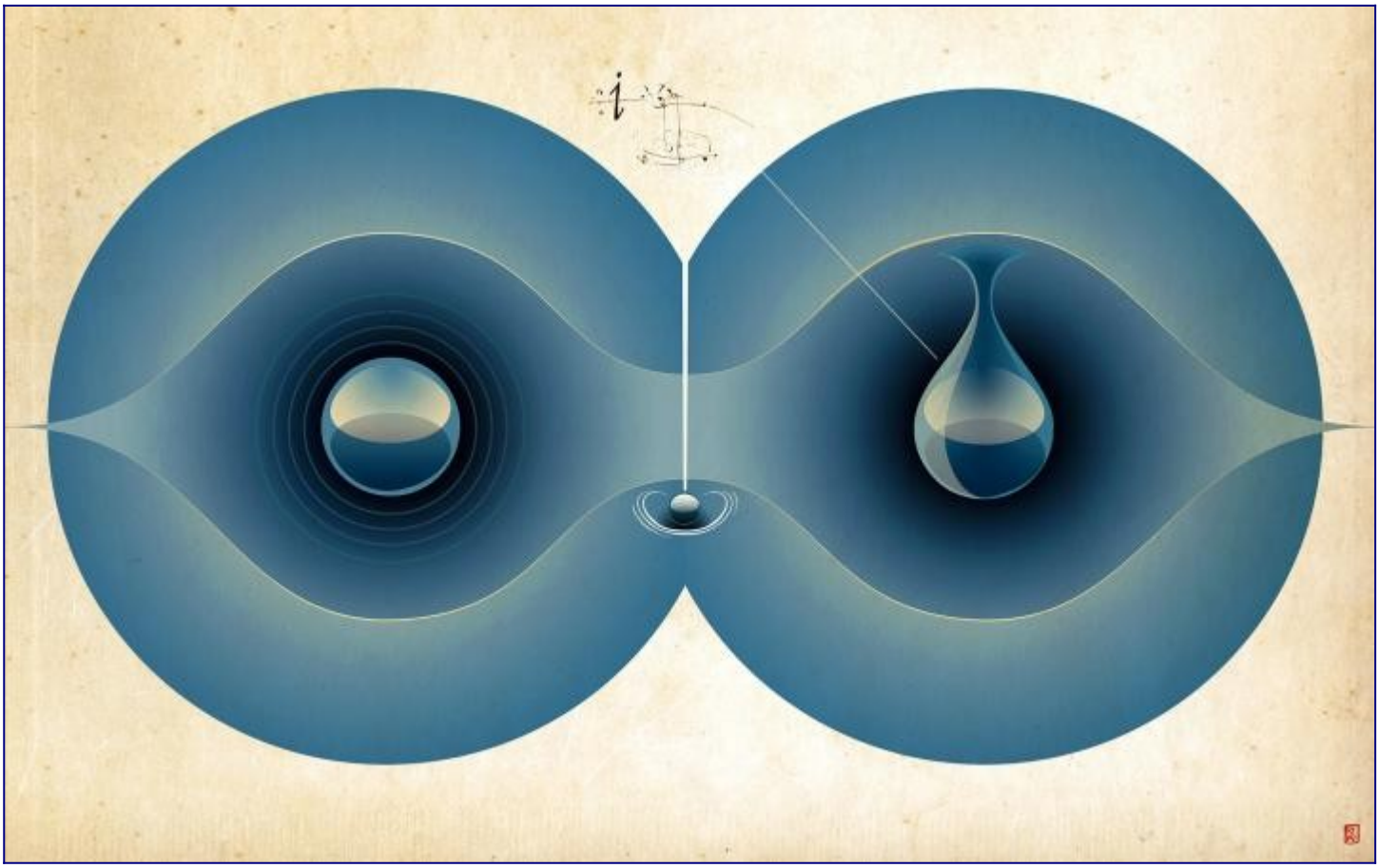
Shared by george davey smith

How imaginary numbers describe the fundamental shape of nature | Aeon Essays

aeon.co

These odd values were long dismissed as bookkeeping. Now physicists are proving that they describe the hidden shape of nature

Shared by Aeon+Psyche



Gaurav Sood’s review of the book Noise by Kahneman et al.: In writing about noise in human judgment, the authors didn’t wrestle with the problem of n…

statmodeling.stat.columbia.edu

Here it is. This should interest some of you. Gaurav makes a convincing case that: 1. The main topic of the book—capriciousness in human judgment—is important, it’s worth a book, and the authors (Kah…

Shared by Andrew Gelman et al.

harm of class imbalance corrections for risk prediction models: illustration and simulation using logistic regression

academic.oup.com

AbstractObjective. Methods to correct class imbalance (imbalance between the frequency of outcome events and nonevents) are receiving increasing interest for de

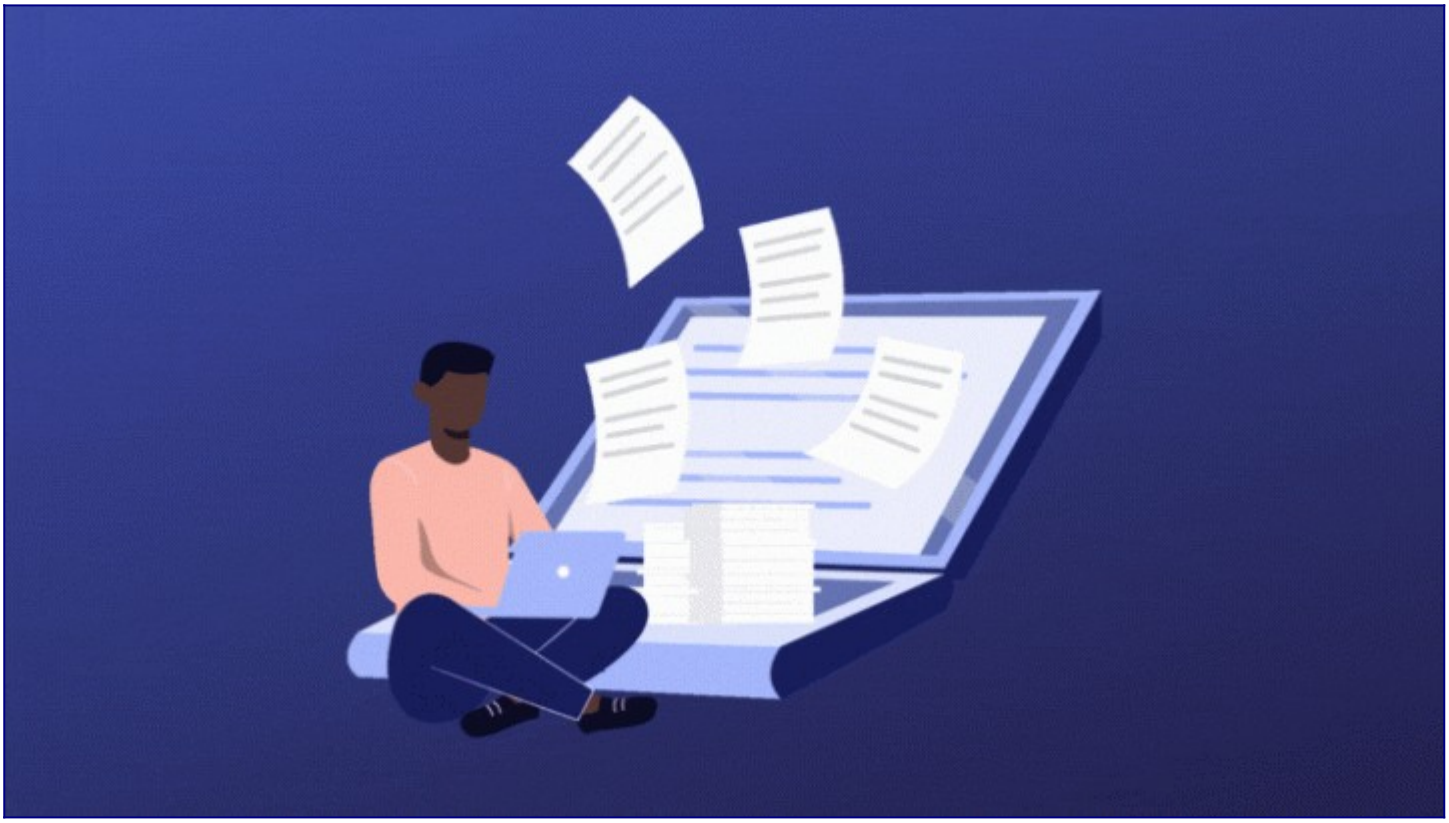
Shared by Maarten van Smeden

The deconstruction of paradoxes in epidemiology

blog.oup.com

If a “revolution” in our field or area of knowledge was ongoing, would we feel it and recognize it? And if so, how? I think a methodological “revolution” is probably going on in the science of epidem…
Shared by Judea Pearl

Making sense of preprints by adding context – The Publish Your Reviews initiative
blogs.lse.ac.uk
Improving scientific publishing is often framed as an issue of openness and speed and less often as one of context. In this post, Ludo Waltman and Jessica Polka make the case for a more contextuali…
Shared by Retraction Watch



U.S. Public Health Agencies Aren't ‘Following the Science,’ Officials Say
commonsense.news
‘People are getting bad advice and we can’t say anything.’
Shared by Vinay Prasad, MD MPH

Target Validity and the Hierarchy of Study Designs
academic.oup.com
Abstract. In recent years, increasing attention has been paid to problems of external validity, specifically to methodological approaches for both quantitative
Shared by Alex Turner
Reference Collection to push back against "Common Statistical Myths"
discourse.datamethods.org
Note: This topic is a wiki, meaning that this main body of the topic can be edited by others. Use the Reply button only to post questions or comments about material contained in the body, or to sugge…
Shared by Brenton Wiernik
R Workflow
hbiostat.org
R Workflow for Reproducible Data Analysis and Reporting
Shared by Frank Harrell
Why we need data-literate citizens | Royal Society
royalsociety.org
Envision: Why we need data-literate citizens. Adrian Smith for the Royal Society
Shared by David Spiegelhalter

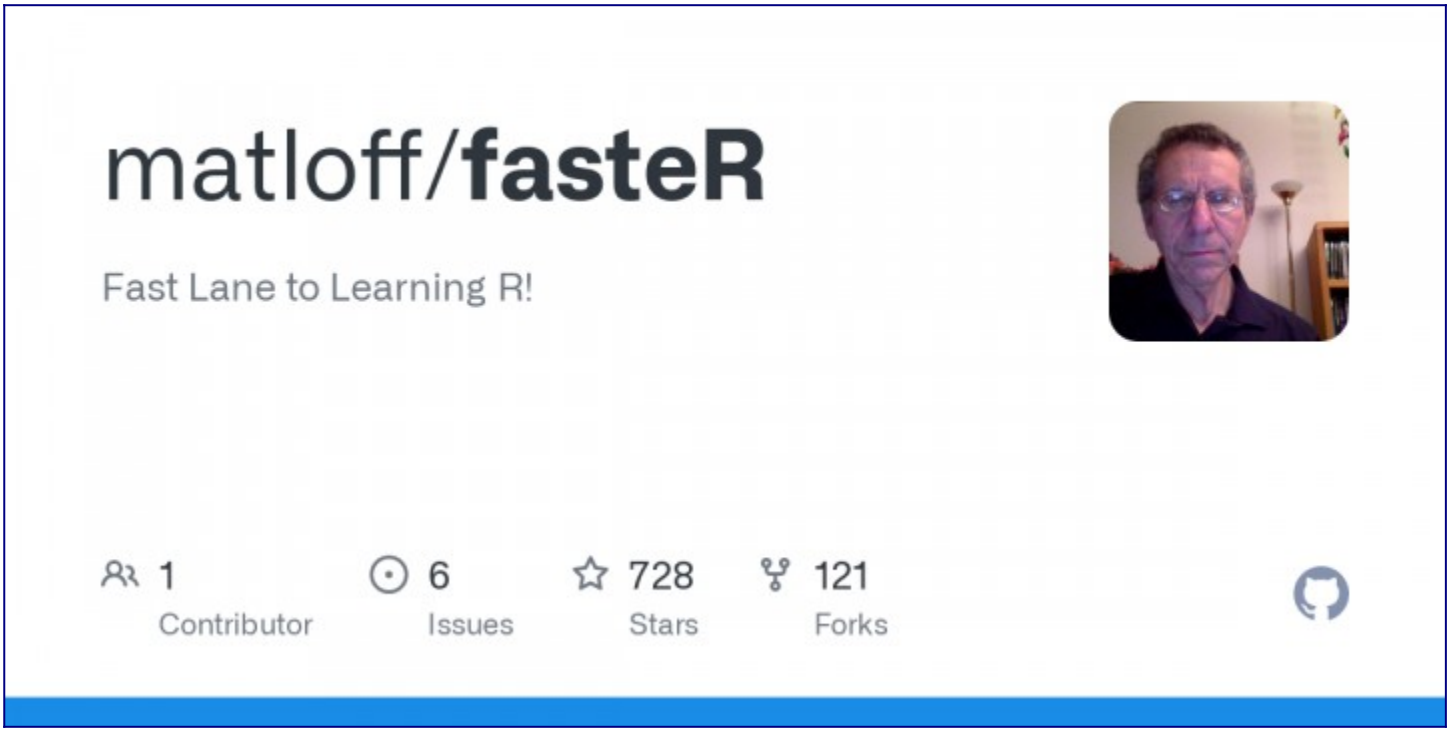
[Performance comparison of dynamic prediction based on joint models and…](#)
kcl.ac.uk
The Prediction Modelling Group provides a forum for researchers at the IoPPN and clinicians at SLaM NHS Foundation Trust, who are interested in prediction modelling applications for precision medicin…
Shared by Richard Riley (R^2)

[GitHub - matloff/fasteR: Fast Lane to Learning R!](#)

github.com

Fast Lane to Learning R! Contribute to matloff/fasteR development by creating an account on GitHub.

Shared by Frank Harrell

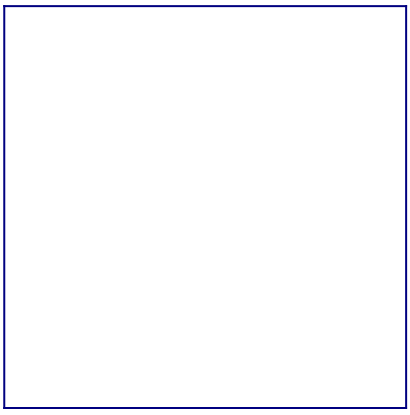


Release of CmdStan 2.30

blog.mc-stan.org

We are very happy to announce that the 2.30.0 release of CmdStan is now available on Github! As usual, the release of CmdStan is accompanied by new releases of Stan Math, core Stan, and Stanc3. A PyS…

Shared by Stan



An Elsevier journal said it would retract 10 papers two years ago. It still hasn't.

retractionwatch.com

Andrew Grey An Elsevier journal has sat for two years on its decision to retract 10 papers by researchers with known misconduct issues, according to emails seen by Retraction Watch. The Journal of th…

Shared by Retraction Watch

Women in Science Receive Less Credit for Their Contributions

nyu.edu

Study first to use new dataset to show gender differences in scientific attribution

Shared by ASA Science Policy

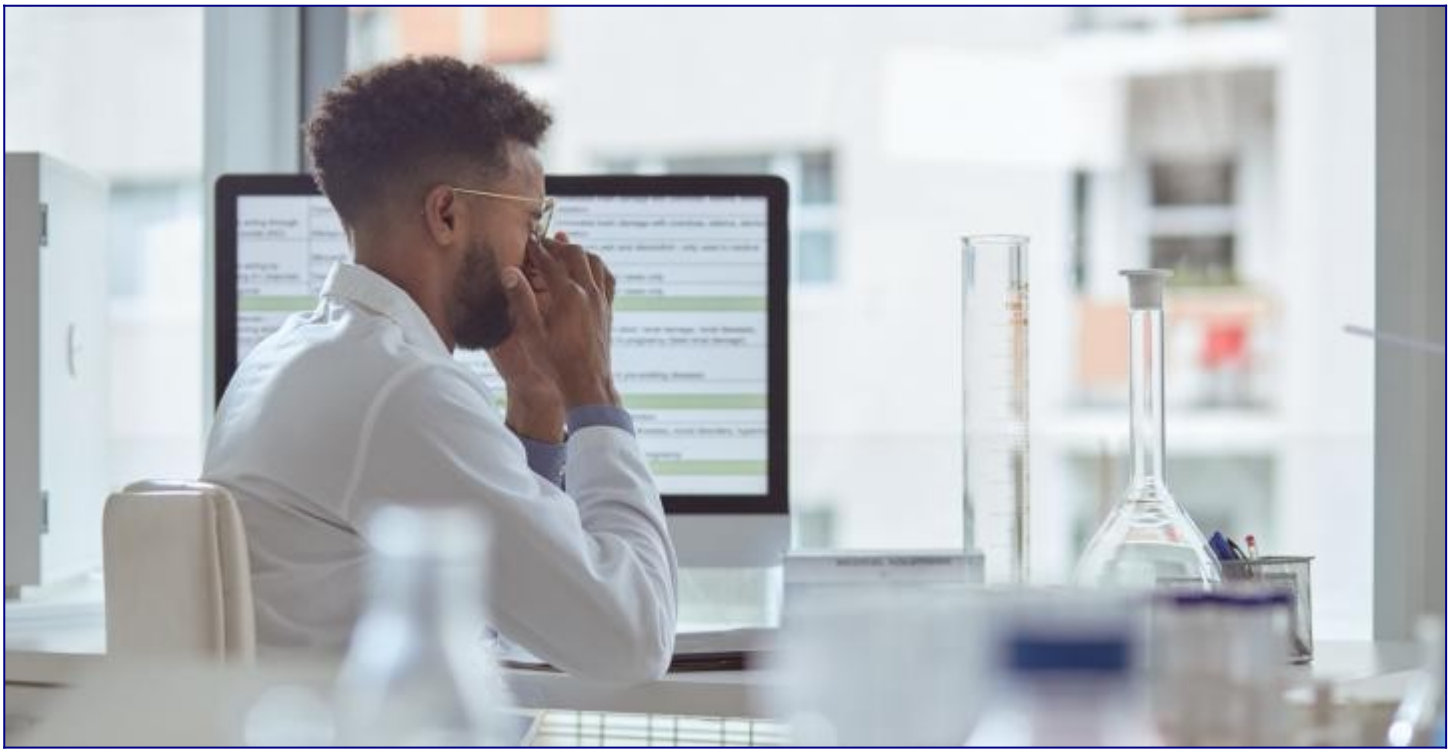
Years after Harvard scandal, U.S. pours millions into tainted field

reuters.com

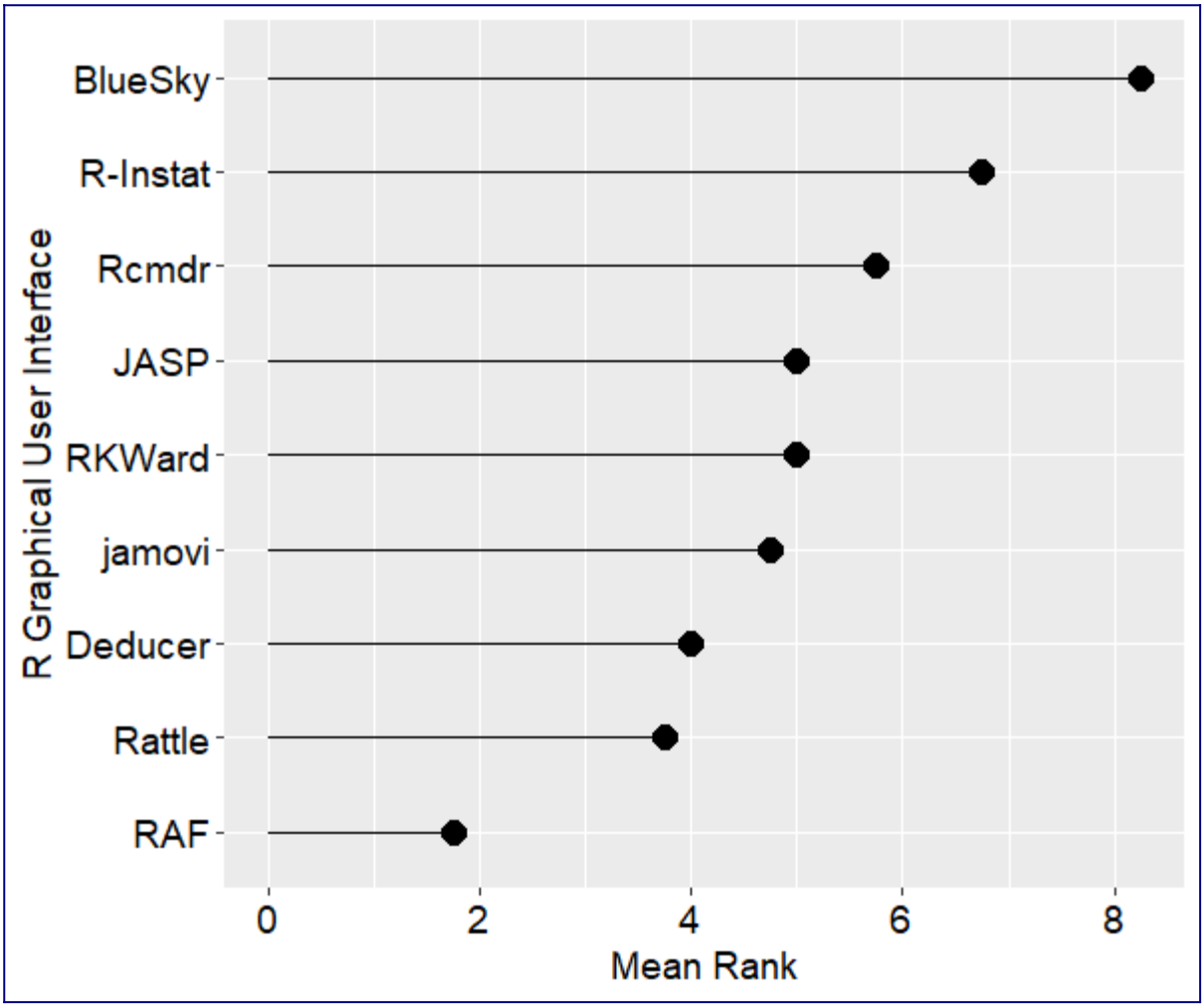
Faked heart studies by a once-obscure scientist duped the U.S. government and medical establishment for years. Washington is still paying for it.

Shared by Retraction Watch

[Why Won't Academia Let Go of 'Publish or Perish'?](#)
undark.org
Opinion | From graduate students to full professors, scientists seem to be too entrenched in the current system to change it.
Shared by Retraction Watch

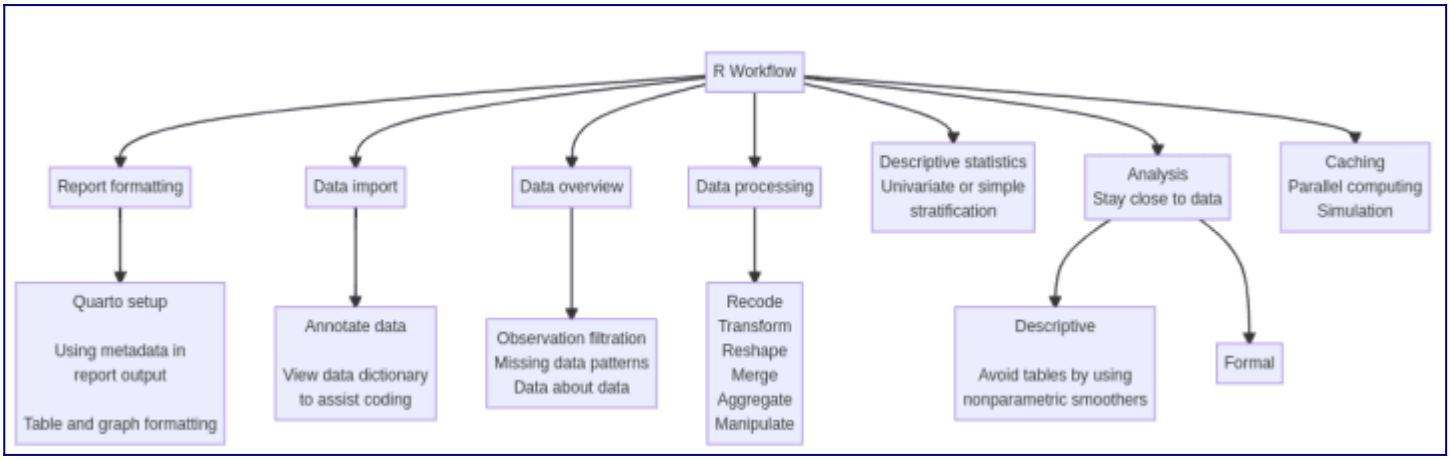


[Updated Comparison of R Graphical User Interfaces](#)
r4stats.com
By Robert A. Muenchen, 6/20/2022 I have just updated my detailed reviews of Graphical User Interfaces (GUIs) for R, so let's compare them again. It's not too difficult to rank them based on the numbe…
Shared by Bob Muenchen

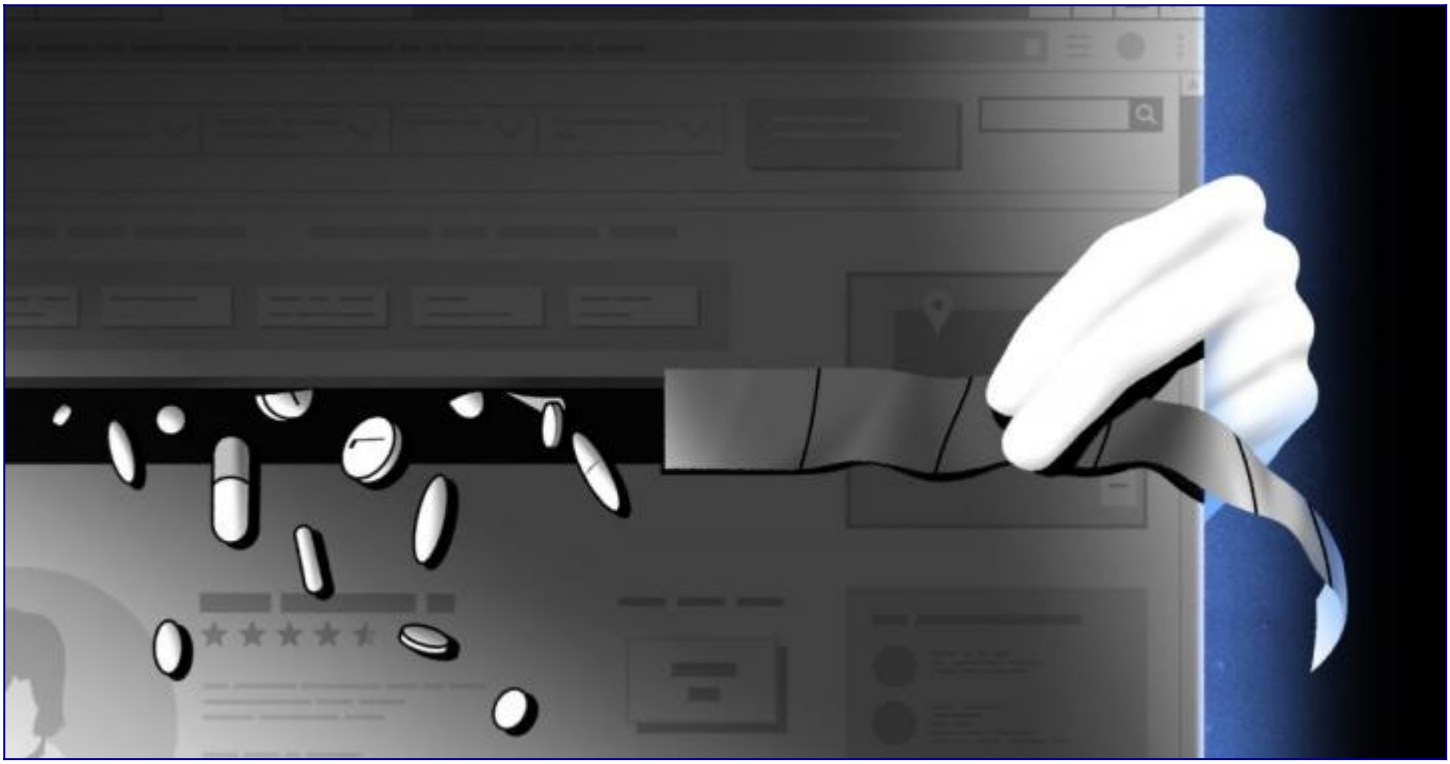


Clinical research: What will be the effect of closing down “unviable” trials?
bmj.com
The UK government’s move to shut down clinical trials that have little chance of being completed has prompted a mixed reaction, as Shaun Griffin hears The Department of Health and Social Care for Eng…
Shared by Retraction Watch

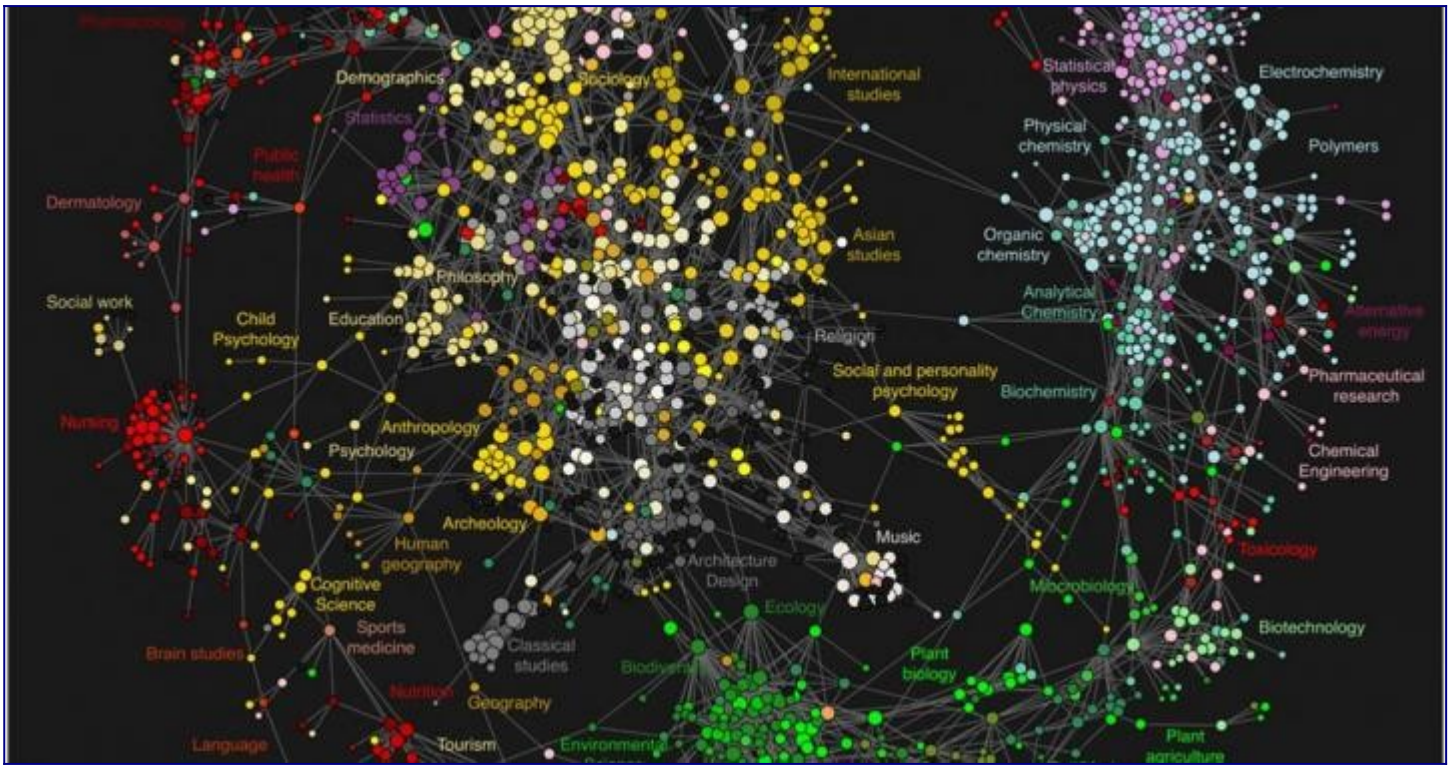
R Workflow
hbiostat.org
This work is intended to foster best practices in reproducible data documentation and manipulation, statistical analysis, graphics, and reporting.
Shared by Frank Harrell



Statistics and science reform: My conversation with economist Noah Smith
statmodeling.stat.columbia.edu
N.S.: In the past few years, you've become sort of the scourge of bad statistical papers. I've heard it said that the scariest seven words in academia are: "Andrew Gelman just blogged about your pape…
Shared by Andrew Gelman et al.
Facebook Is Receiving Sensitive Medical Information from Hospital Websites
themarkup.org
Experts say some hospitals' use of an ad tracking tool may violate a federal law protecting health information
Shared by Harlan Krumholz



[How Science Fuels a Culture of Misinformation](#)
[openmindmag.org](#)
[We tend to blame the glut of disinformation in science on social media and the news, but the problem often starts with the scientific enterprise itself.](#)
Shared by [Retraction Watch](#)



[Journal run by new AMA president-elect caught in special issue scam](#)

retractionwatch.com

Jesse Ehrenfeld A med-tech journal whose editor-in-chief is the president-in-waiting of the American Medical Association has retracted six papers for compromised peer review and related problems. The…

Shared by Retraction Watch

Article(s): Coffee Benefits and Continuous vs. Categorical Variable Analysis

discourse.datamethods.org

Saturday morning, feeling guilty about making a second (or was it third?) cup of coffee, I was mucking around on PubMed to see if I could see if we really know if it benefits health. You can’t read t…

Do thought experiments really uncover new scientific truths? | Aeon Essays

aeon.co

Thought experiments played a crucial role in the history of science. But do they tell us anything about the real world?

Shared by Aeon+Psyche

Peer-review crisis creates problems for journals and scholars

insidehighered.com

The peer-review system, which relies on unpaid volunteers, has long been stressed. COVID-19 is making it worse—a lot worse. Possible solutions include paying reviewers or limiting revise-and-resubmit…

Shared by Retraction Watch

Ivermectin for Treatment of Mild-to-Moderate COVID-19 in the Outpatient Setting: A Decentralized, Placebo-controlled, Randomized, Platform Clinical T…

medrxiv.org

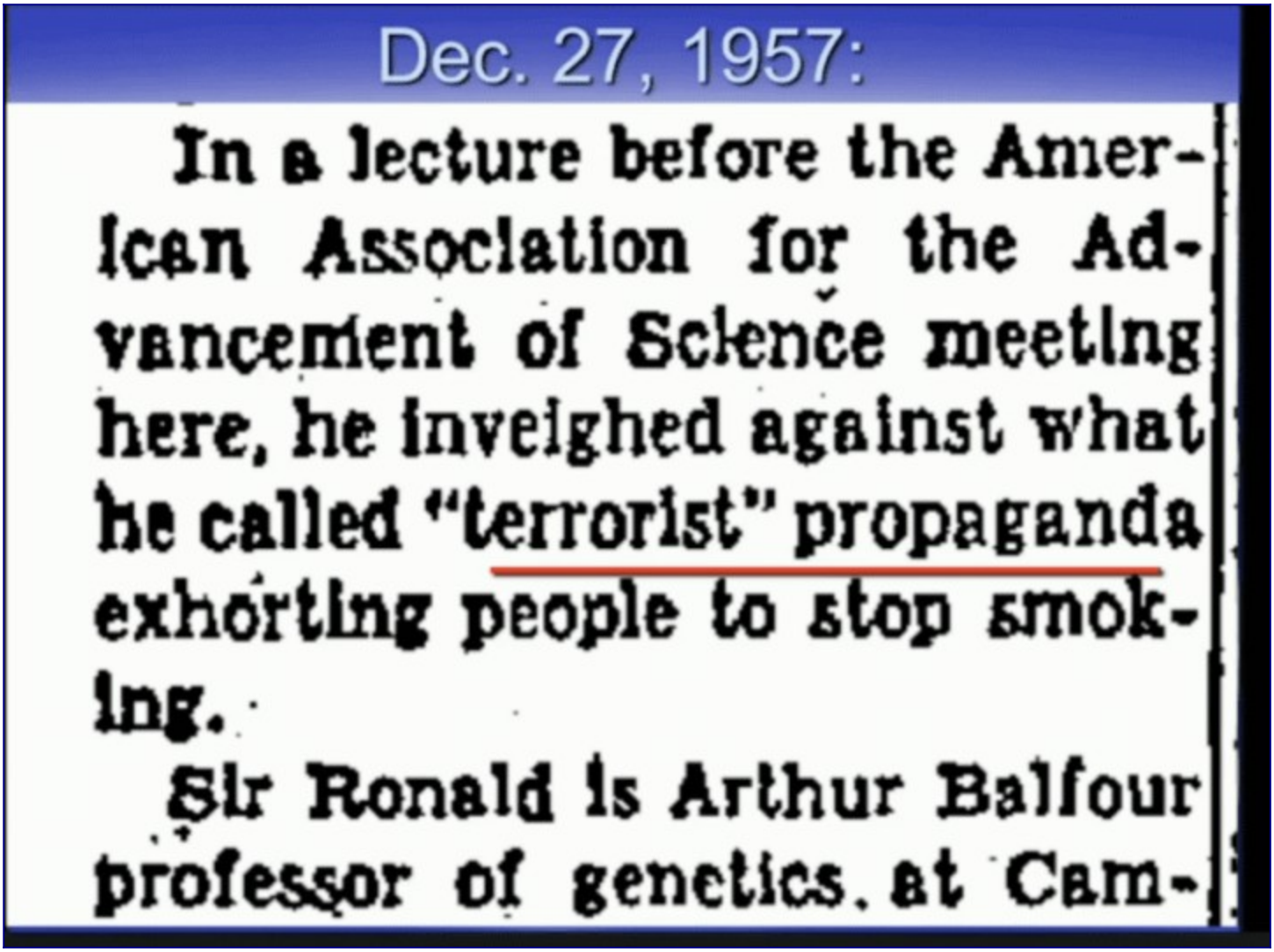
medRxiv - The Preprint Server for Health Sciences

Shared by Harlan Krumholz

Bayesian workflow book - Digits
avehtari.github.io
A workflow for deciding how many digits to report when summarizing the posterior distribution, and how to check how many independent Monte Carlo (MC) draws or dependent Markov chain Monte Carlo (MCMC…
Shared by Aki Vehtari

Everything you always wanted to know about evaluating prediction models (but were too afraid to ask)
ncbi.nlm.nih.gov
Many urology papers describe the development of statistical prediction models, what are often described as “nomograms”. Indeed, a simple search of Medline for urology prediction models retrieves clos…
Shared by Maarten van Smeden

How can a top scientist be so confidently wrong? (R. A. Fisher and smoking example)
statmodeling.stat.columbia.edu
Robert Proctor points us to this lecture, “Why did Big Tobacco love (and fund) eugenicists like R.A. Fisher?” It’s pretty good! He talks about material from his 2011 book, Golden Holocaust: Origins o…
Shared by Andrew Gelman et al.



Disambiguating Impact
blogs.lse.ac.uk
Outside of specific institutional and organizational settings discussions about 'impact' often descend into confusion, as the term holds several meanings for researchers. Responding to a provocatio…
Shared by Retraction Watch

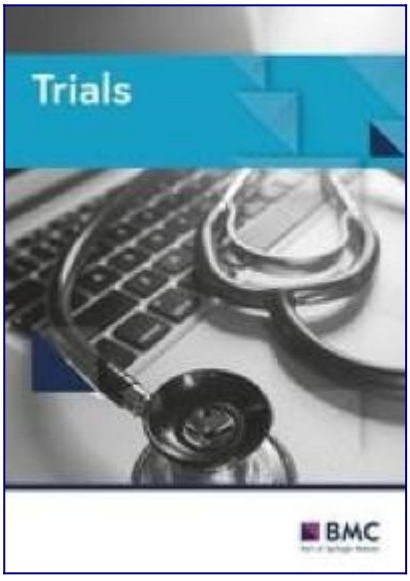
[Why “stealth research” is the bane of scientific inquiry]
salon.com
First Theranos, and now AI researchers, are being opaque about their work. Should we trust them?
Shared by Retraction Watch

Cornell food marketing researcher who retired after misconduct finding is publishing again retractionwatch.com

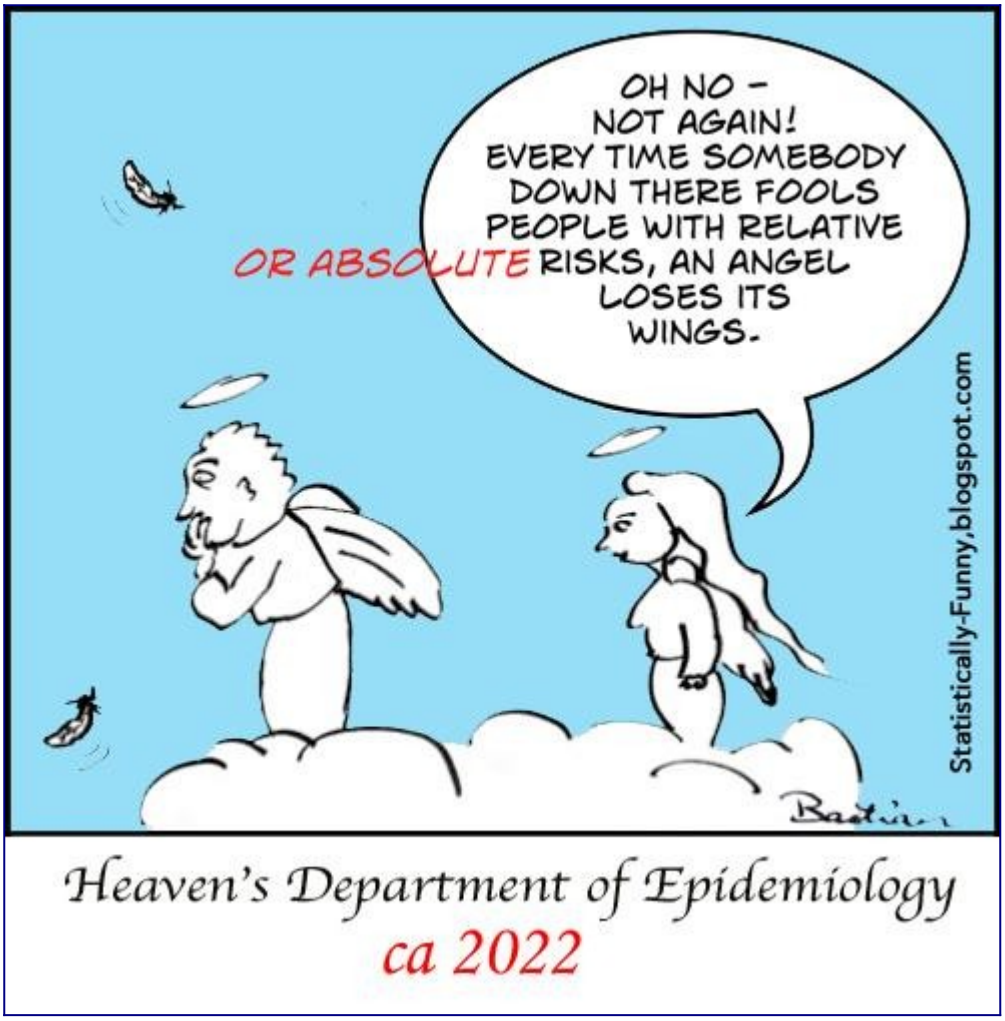
Brian Wansink Brian Wansink, the food marketing researcher who retired from Cornell in 2019 after the university found that he had committed academic misconduct, has published two new papers. The art…

Shared by Retraction Watch

Tolerating bad health research: the continuing scandal
trialsjournal.biomedcentral.com
At the 2015 REWARD/EQUATOR conference on research waste, the late Doug Altman revealed that his only regret about his 1994 BMJ paper ‘The scandal of poor medical research’ was that he used the word ‘…
Shared by Maarten van Smeden



Another Bit of Pandemic Fallout: The Weaponization of Absolute Risk Statistics
absolutelymaybe.plos.org
Back in the good old days, we only had to tear our hair out when people used relative risks without context, thereby…
Shared by Darren Dahly



UK shuts down all ‘unviable’ clinical trials
transparimed.org
“The capacity of the NHS to support research delivery remains under continued pressure from workload and workforce issues, as well as the need to reduce the elective backlog. Limited capacity and res…
Shared by Michelle Gallaher

Improving research integrity: a framework for responsible science communication
bmcresnotes.biomedcentral.com
Research integrity, an essential precept of scientific inquiry and discovery, comprises norms such as Rigor, Reproducibility, and Responsibility (the 3R’s). Over the past decades, numerous issues hav…
Shared by Retraction Watch

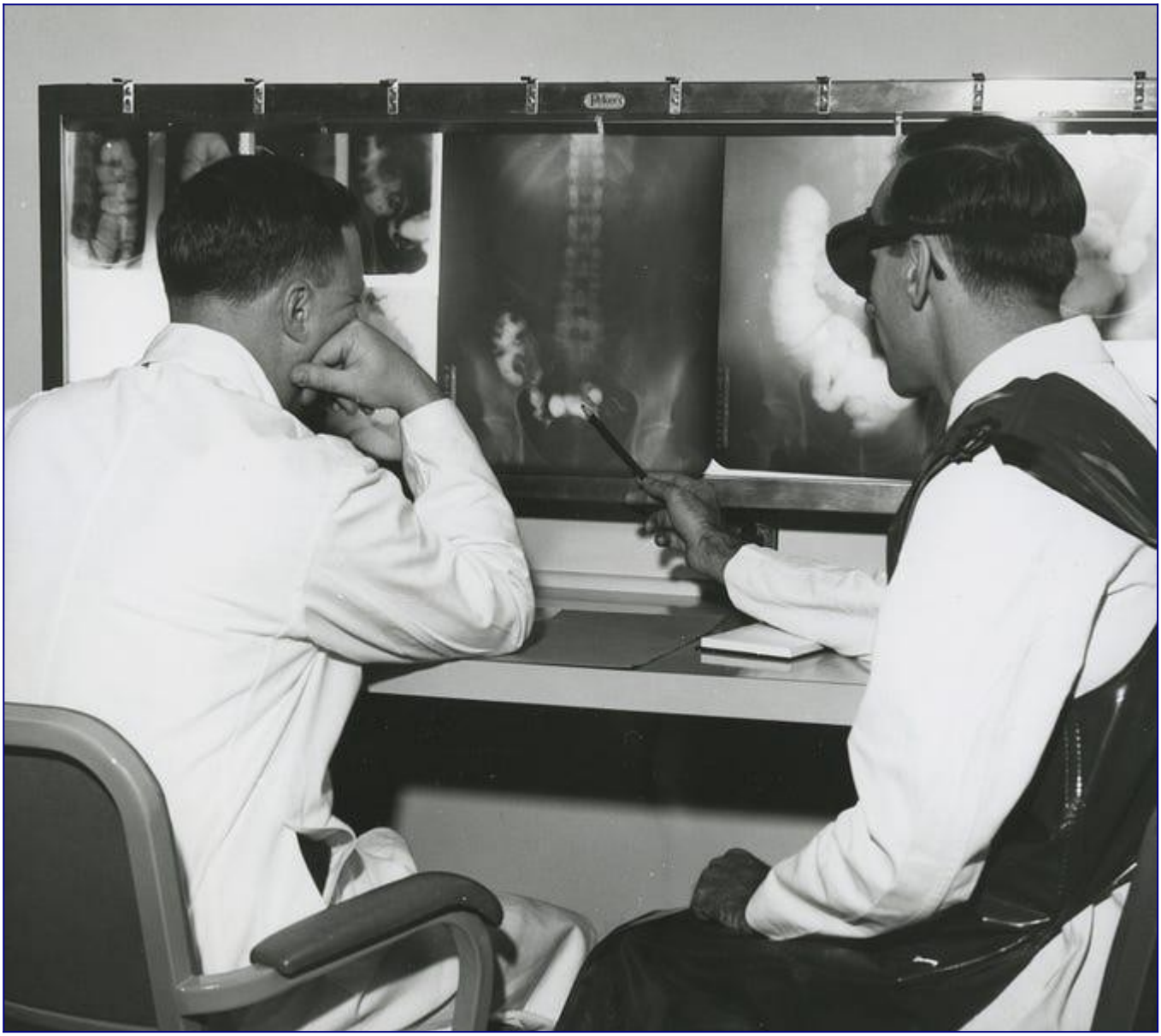
My students cheated... A lot
crumplab.com
May 26th, 2022 - Lot's of cheating happened. It's a long story with ups and downs. I tried my best to turn things around. I'm blogging about it for catharsis and personal growth.
Editor's note: This is a fantastic read
Shared by EJ Wagenmakers
Real-Time Collaborative Editing on RStudio Cloud
rstudio.com
We are excited to announce real-time collaborative editing on RStudio Cloud. Users can join the same project, edit code, and immediately see each other’s changes.
Shared by RStudio

Study of reproducibility issues points finger at the mice
arstechnica.com
Experiments may have a certain amount of randomness that can't be eliminated.
Shared by Retraction Watch

R Workflow
fharrell.com
This article outlines analysis project workflow found to be efficient in making reproducible research reports using R with Quarto.
Shared by Frank Harrell
Stop aggregating away the signal in your data
stackoverflow.blog
By aggregating our data in an effort to simplify it, we lose the signal and the context we need to make sense of what we’re seeing.
Shared by Edward Tufte

[“Causal” is like “error term”: it’s what we say when we’re not trying to model the process.
statmodeling.stat.columbia.edu
After my talks at the University of North Carolina, Cindy Pang asked me a question regarding causal inference and spatial statistics: both topics are important in statistics but you don’t often see t…]
Shared by Andrew Gelman et al.

[AI for medicine is overhyped right now
moreisdifferent.substack.com
Should we update our timelines to AGI as a result?]
Shared by Thomas Speidel



Software Development Resources for Data Scientists
rstudio.com
Many tools used routinely by software developers can also be useful to data scientists.
Shared by RStudio

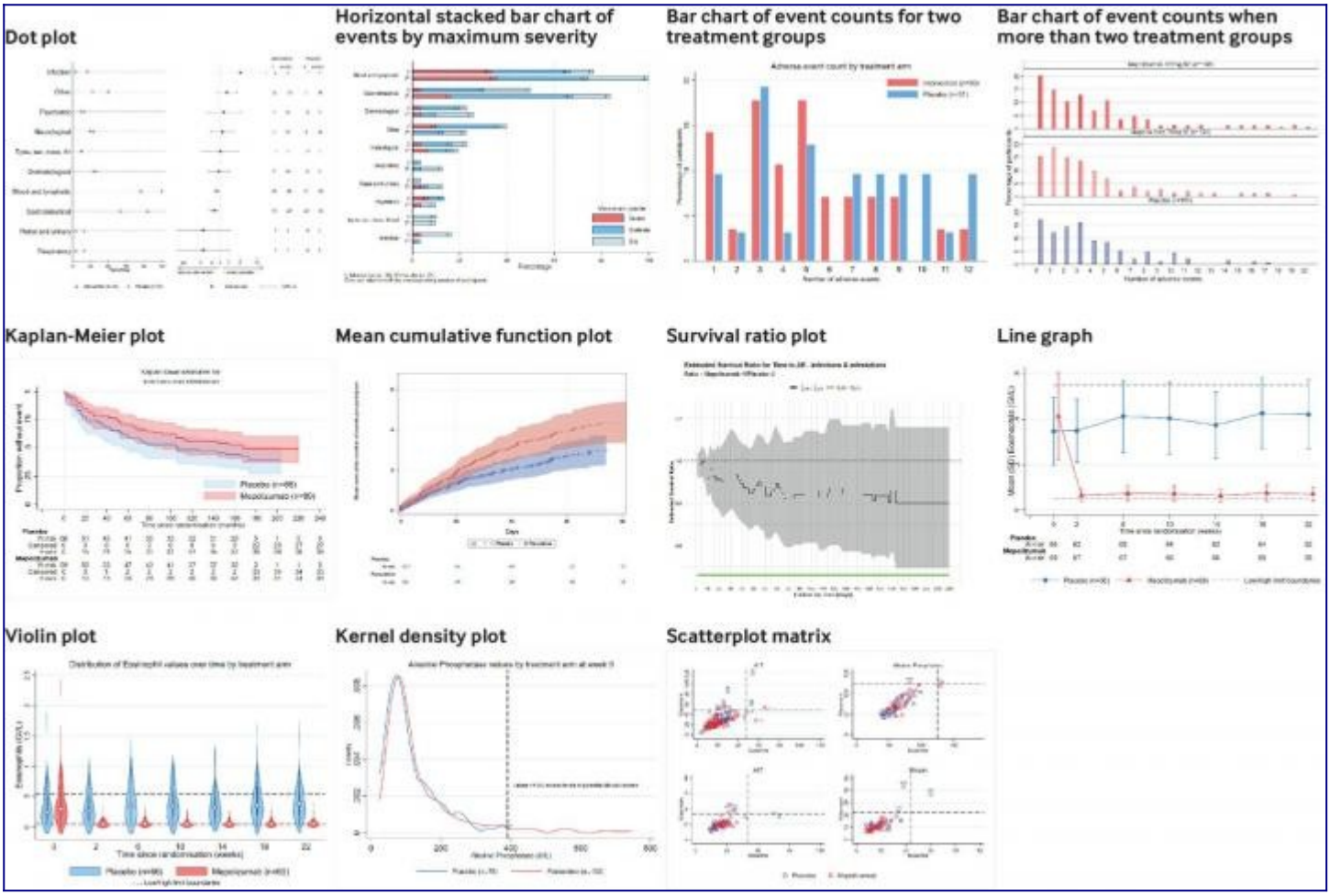
No, They Didn’t Find the Cause of SIDS
theatlantic.com
An inspiring study went viral for the wrong reasons.
Shared by Venk Murthy, MD, PhD

Real world evidence: Lessons learned from an FDA pilot show the limits of emulating RCTs
endpts.com
Only about half of a select group of clinical trials could be well-emulated with available real world evidence, according to the newly discussed results of an FDA pilot program. The FDA-funded progra…

Shared by Vinay Prasad, MD MPH

Visualising harms in publications of randomised controlled trials: consensus and recommendations
bmj.com
Objective To improve communication of harm in publications of randomised controlled trials via the development of recommendations for visually presenting harm outcomes. Design Consensus study. Settin…

Shared by Medical Research Library of Brooklyn



[Between two stools: preclinical research, reproducibility, and statistical design of experiments](bmcresnotes.biomedcentral.com)
[Translation of animal-based preclinical research is hampered by poor validity and reproducibility issues. Unfortunately, preclinical research has ‘fallen between the stools’ of competing study design…]
Shared by [Brian Nosek]

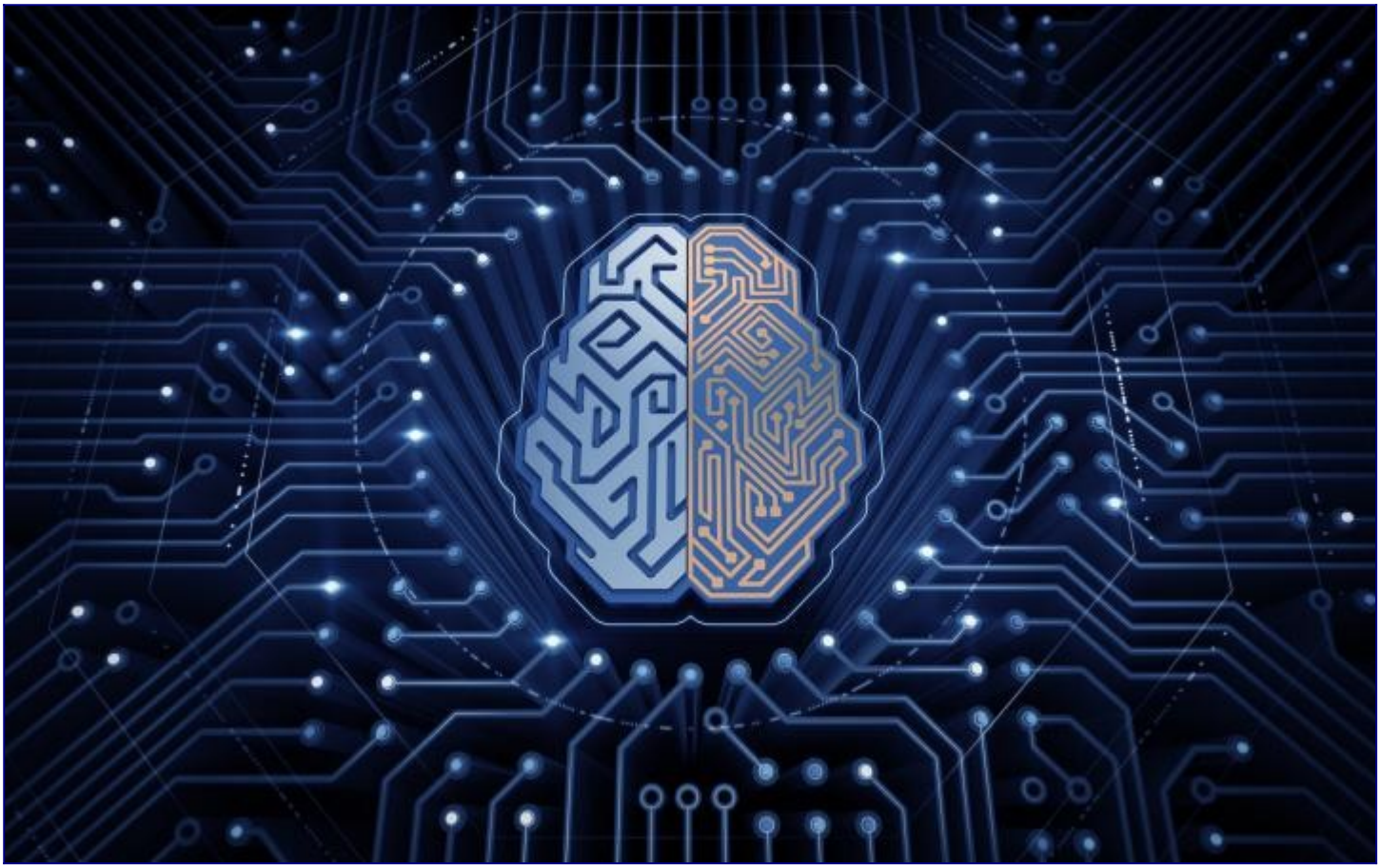
[Another Firing Among Google’s A.I. Brain Trust, and More Discord]

nytimes.com
The researchers are considered a key to the company’s future. But they have had a hard time shaking infighting and controversy over a variety of issues.
Shared by Retraction Watch

Unfair Comparison?
linkedin.com
Causal inference vs. statistical inference by Stephen Senn
Shared by Frank Harrell
Lessons learnt when accounting for competing events in the external validation of time-to-event prognostic models
academic.oup.com
AbstractBackground. External validation of prognostic models is necessary to assess the accuracy and generalizability of the model to new patients. If models ar
Shared by Maarten van Smeden
An expert’s opinion: Interview with Ivan Oransky on the perils of scientific publishing
sciencemediahub.eu
Medical writer Ivan Oransky, co-founder of the blog Retraction Watch and Editor-in-Chief of the autism research news website Spectrum, speaks about the
Shared by Retraction Watch

Necessary/recommended level of theory for developing statistical intuition
discourse.datamethods.org
I brought this up after Day 1's lecture but I thought it might warrant further discussion/elaboration. (Apologies for the length!) I am of the opinion that a "strong" background/foundation in theory …

AI: Still Just Curve Fitting, Not Finding a Theory of Everything
mindmatters.ai
Computers are much better than humans at curve fitting but still far worse at devising models that help us understand and predict the world.
Shared by Maarten van Smeden



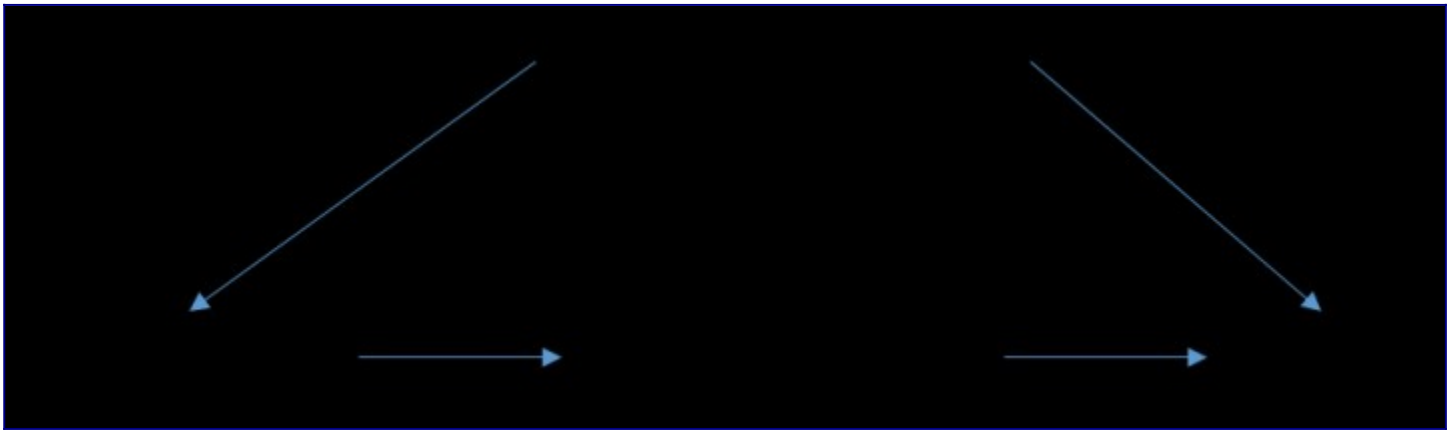
Guernsey McPearson Prose
senns.uk
Continuitis, Dichotomania and the Tetrachoric Coefficient of Correlation Balancing Act: Guernsey Scrapes the Barrel OK, so you might think that in this age of air travel, channel tunnels and the like…
Shared by Stephen John Senn

‘A terrifying experience’: A team of researchers does the right thing when they find an error
retractionwatch.com
Mitch Brown Mitch Brown was preparing last August to launch a follow-up study to a 2021 paper on coalitions when he found something in his computer coding that sent his stomach to his shoes. As Brown…
Shared by Retraction Watch

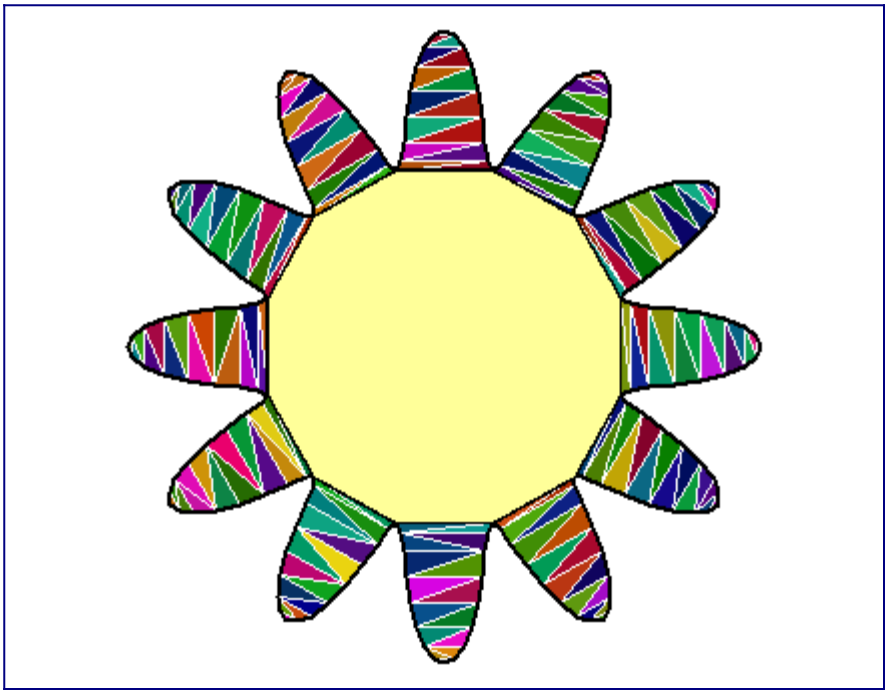
We have really everything in common with machine learning nowadays, except, of course, language.
statmodeling.stat.columbia.edu
I had an interesting exchange with Bob regarding the differences between statistics and machine learning. If it were just differences in jargon, it would be no big deal—you could just translate back …
Shared by Andrew Gelman et al.

Resources for Ordinal Regression Models
fharrell.com
Consider an ordinal outcome variable Y and note that an interval-scaled Y is also ordinal. Such response variables can be analyzed with nonparametric tests and correlation coefficients and wi…
Shared by Frank Harrell

R Workflow
fharrell.com
This article outlines analysis project workflow that I've found to be efficient in making reproducible research reports using R with Rmarkdown and now Quarto.
Shared by Frank Harrell
Directed Acyclic Graphs (DAGs)
lancaster.ac.uk
In my last two blog posts I focused on how to analyse the results of clinical trials through both Meta Analysis and Simultaneous Inference. Here we're going to take a step back and look at we choose …
Shared by Stephen John Senn



Collaborative Author Involved Replication Studies
daniellakens.blogspot.com
Recently a new category of studies have started to appear in the psychological literature that provide the strongest support too date for a ...
Shared by Brian Nosek
March: "Top 40" New CRAN Packages
rviews.rstudio.com
An R community blog edited by RStudio
Shared by RStudio



Sliced And Diced: The Inside Story Of How An Ivy League Food Scientist Turned Shoddy Data Into Viral Studies

buzzfeednews.com

Brian Wansink won fame, funding, and influence for his science-backed advice on healthy eating. Now, emails show how the Cornell professor and his colleagues have hacked and massaged low-quality data…

Shared by Ivan Oransky

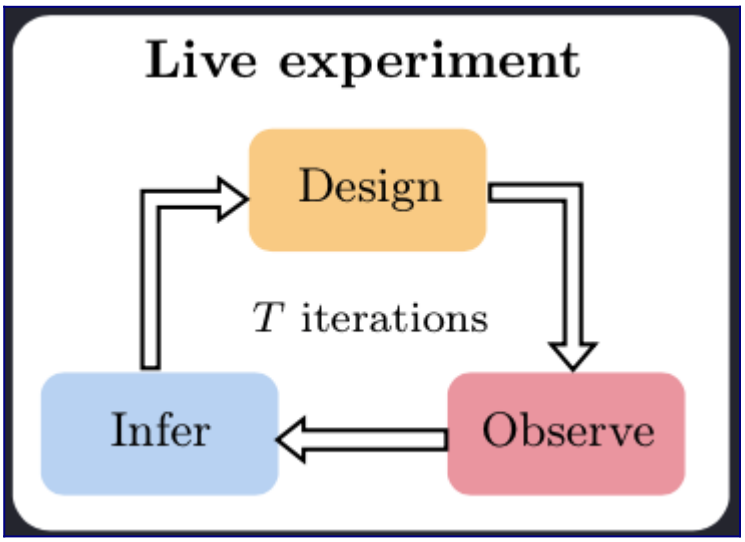
Controlling α

hbiostat.org
How Often Do We Assert an Effect vs. How Often Are We Wrong About an Effect Type I error rate is not a rate and is not the probability of an error. It is the probability of making an assertion of an …
Shared by Frank Harrell
Tutorials
user2022.r-project.org
List of tutorials
Shared by Vanderbilt Biostatistics
Improving clinical trials using Bayesian adaptive designs: a breast cancer example
bmcmedresmethodol.biomedcentral.com
To perform virtual re-executions of a breast cancer clinical trial with a time-to-event outcome to demonstrate what would have happened if the trial had used various Bayesian adaptive designs instead…
Shared by Darren Dahly

Evaluating the impact of prediction models: lessons learned, challenges, and recommendations
diagnprognres.biomedcentral.com
An important aim of clinical prediction models is to positively impact clinical decision making and subsequent patient outcomes. The impact on clinical decision making and patient outcome can be quan…
Shared by Kazuki Yoshida

[Leading Academics: "Stop Epidemiologists From Giving Public Health Advice"](#)
gregmantell.substack.com
[Top researchers at the world's leading universities say most epidemiologists haven't got the proper background to give public health advice](#)
Shared by [David Spiegelhalter](#)

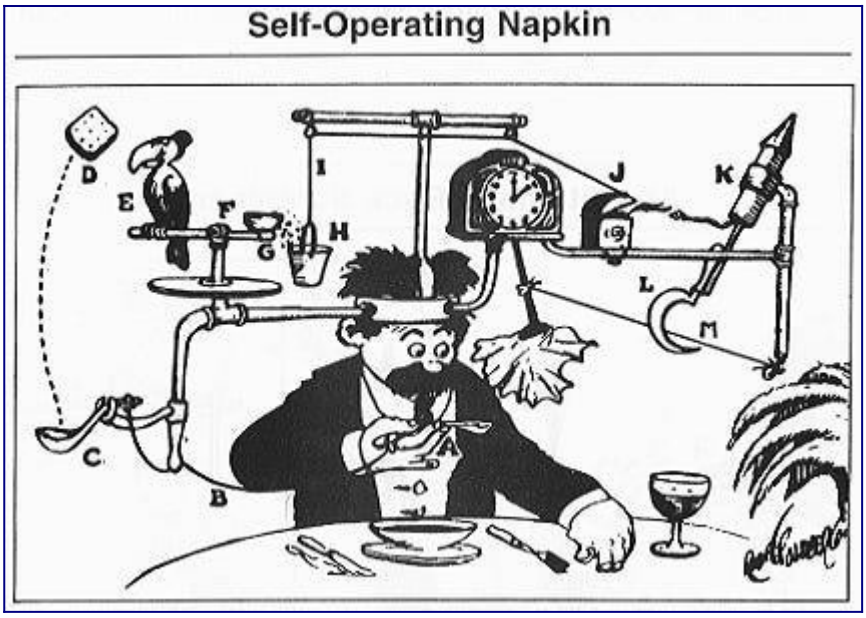
[Introduction to Bayesian Optimal Experimental Design | Desi R. Ivanova](#)
desirivanova.com
[Bayesian Optimal Experimental Design (BOED) is an elegant mathematical framework that enables us to design experiments optimally. This introductory post describes the BOED framework and the computati…](#)
Shared by [Sam Finlayson](#)



[The perils of (apparently) logical thinking | Statistical Modeling, Causal Inference, and Social Science](#)
statmodeling.stat.columbia.edu
[Regular readers know that I've been screaming for years about the problems with binary thinking, for example here and here. Or, for a applied example from medical research, here. This comes up in sta…](#)
Shared by [Andrew Gelman et al.](#)

Preprint: Analysis of Likert Type Data Using Metric Methods
discourse.datamethods.org
Dennis Boos and Judy Chen (2022). Analysis of Likert-Type Data Using Metric Methods. The authors defend the common practice of using parametric models on ordinal data on the basis of pragmatics (ie.“…

No, you can't explain what a p-value is with one sentence (Parts I, II)
statsepi.substack.com
But I'm going to try and do it in as few as possible. #fml
Shared by Darren Dahly



Comparison of Statistical Model & ML in Predicting RFD risk | RMHP

dovepress.com

Evaluating the performance of machine learning algorithms and comparing them with logistic regression for predicting the risk of renal function decline (RFD)

The authors made the common mistake of assuming that logistic regression is logistic regression as it was used upon its invention in 1958, assuming all effects are linear on the logit scale.

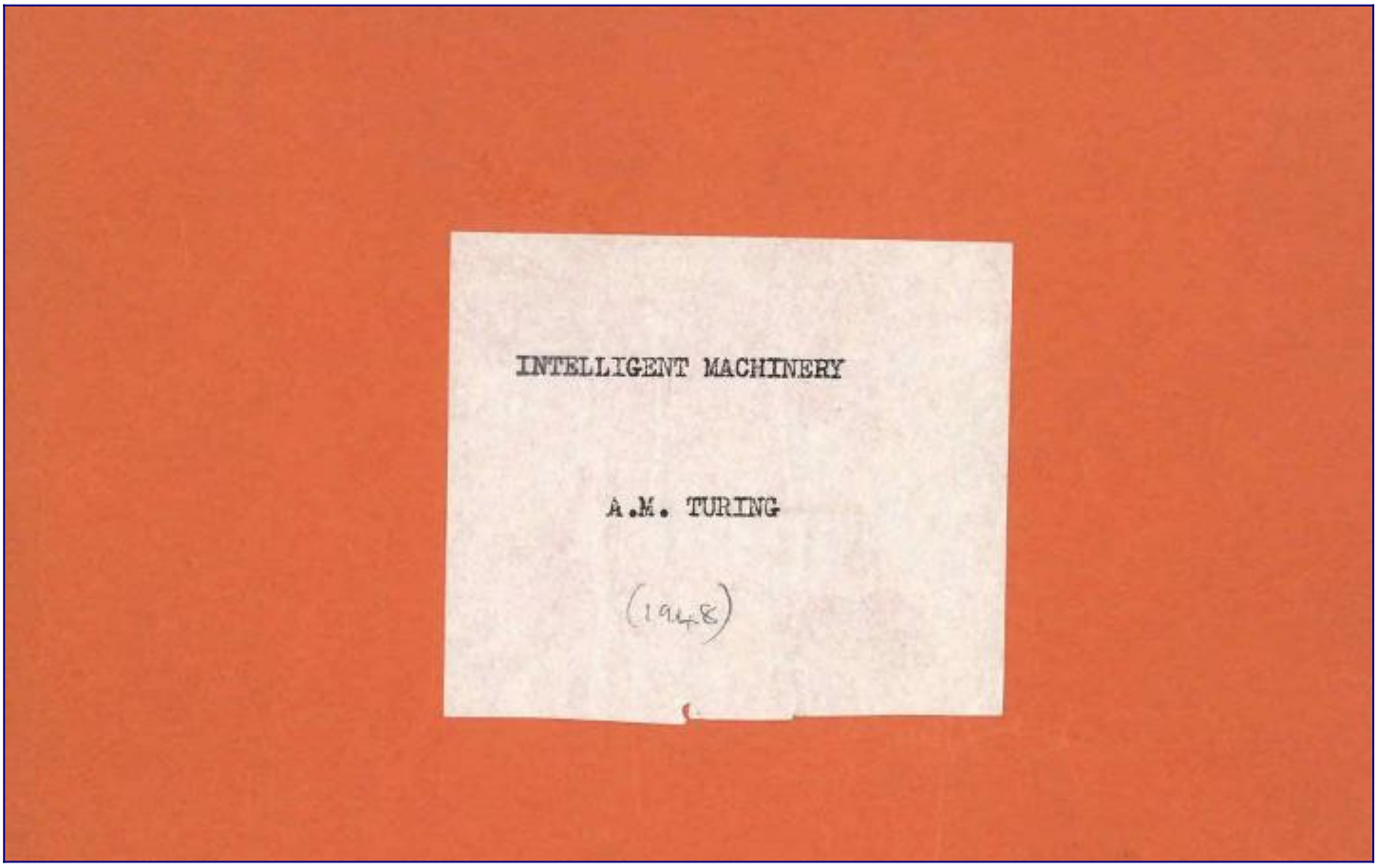
Shared by AI News Feed

Why we should remember Alan Turing as a philosopher | Aeon Essays

aeon.co

Alan Turing was a pioneer of machine learning, whose work continues to shape the crucial question: can machines think?

Shared by Aeon+Psyche



What is a cause?
medium.com
If you’re interested in causal inference, you’ve probably wondered at one time or another what even is a cause? The answer to that is complicated, because it’s not really science — it’s philosophy…
Shared by Dr Ellie Murray, ScD

Stephen Mumford & Rani Lill Anjum

CAUSATION

A Very Short Introduction

OXFORD

Racism as a Leading Cause of Death
medpagetoday.com
'It's time for acknowledgement, accountability, and improvement,' says Harlan Krumholz, MD
Shared by Harlan Krumholz

Unethical studies of ivermectin for covid-19
bmj.com
Flawed research means we still do not know if the drug is safe or effective During disease outbreaks, it can be tempting to sacrifice the scientific rigour of research in favour of speed, given the u…
Shared by Retraction Watch
What do statisticians do that can't be automated?
stats.stackexchange.com
Will software eventually make statisticians obsolete? What is done that can't be programmed into a computer?
Shared by Maarten van Smeden

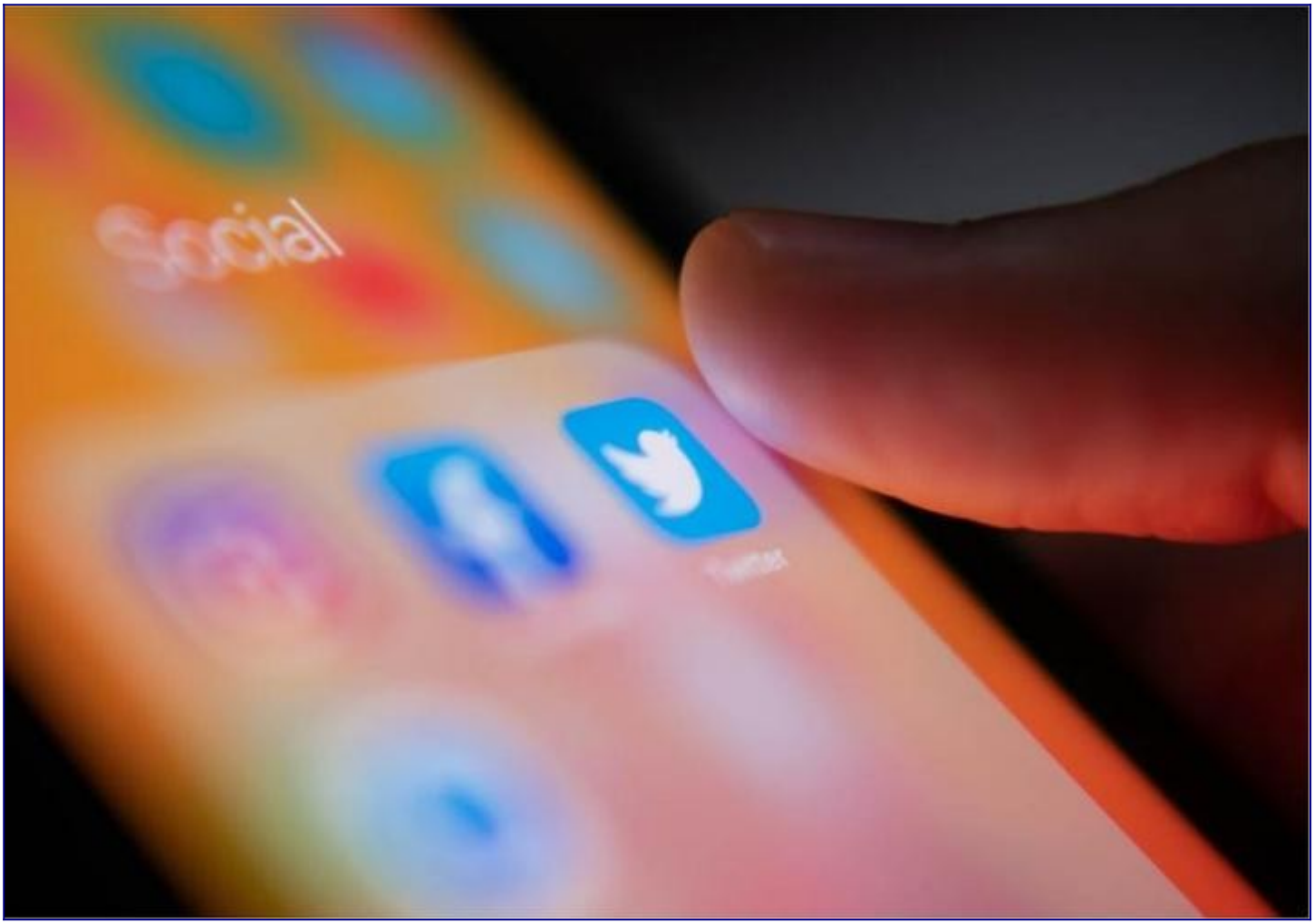
What plants can teach us about human suffering
bigthink.com
Plants are very sensitive to touch, with research showing that touching a plant can change its genome and launch a cascade of plant hormones.
Shared by Haider Warraich

More than 300 at once: Publisher retracts entire conference proceedings

retractionwatch.com
The tip came from the leadership of another scientific conference. Did the Association for Computing Machinery know that they had published the proceedings of a conference with essentially the same n…
Shared by Ivan Oransky
The big idea: should we get rid of the scientific paper?
theguardian.com
As a format it's slow, encourages hype, and is difficult to correct. A radical overhaul of publishing could make science better
Shared by Retraction Watch

COVID-19-vitamin D paper retracted by Springer Nature journal
retractionwatch.com
A journal has retracted a 2021 paper claiming that vitamin D "significantly reduced the inflammatory markers associated with COVID-19 without any side effects" following criticism that led them to "n…
Shared by Ivan Oransky
The CDC Should Admit Its Errors To Regain Trust
reason.com
The CDC thinks a month-long review of COVID policies will be sufficient to redress their errors.
Shared by Vinay Prasad, MD MPH

UK Funding Agency Apologizes for Role in Researchfish Controversy
the-scientist.com
Researchfish is a platform commonly used to track the status of grants and the impacts of research. When academics were critical of the company online, Researchfish shared these comments with the lar…
Shared by Retraction Watch



[Decision curve analysis for quantifying the additional benefit of a new marker](#)
fharrell.com
Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center Decision curve analysis for quantifying the additional benefit of a new marker Decision curve analysis is a method…
[Why autism therapies have an evidence problem | Spectrum | Autism Research News](#)
spectrumnews.org
Early interventions for autism lack solid data. The source of this problem is murky but may stem from ongoing debates about evidence quality and entrenched conflicts of interest within the field.
Shared by Retraction Watch

Methodological conduct of prognostic prediction models developed using machine learning in oncology: a systematic review

bmcmedresmethodol.biomedcentral.com

Describe and evaluate the methodological conduct of prognostic prediction models developed using machine learning methods in oncology. We conducted a systematic review in MEDLINE and Embase between 0…

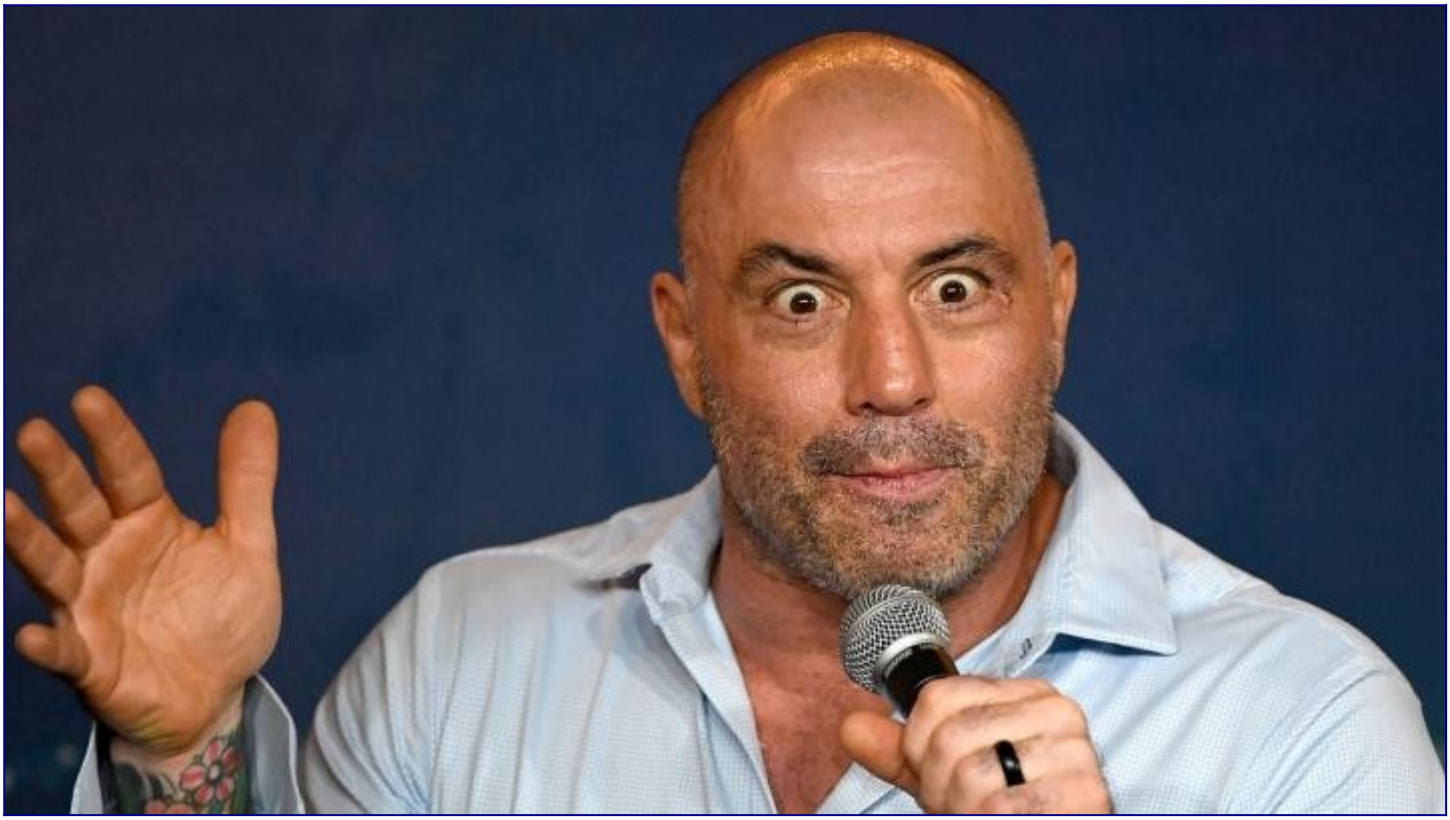
Shared by Gary Collins

The End of the "Syndrome" in Critical Care
discourse.datamethods.org
I have been periodically discussing the state of RCT in critical care. While there have been but few on-line comments, I have been encouraged by the off-line comments and approvals of that which I am…

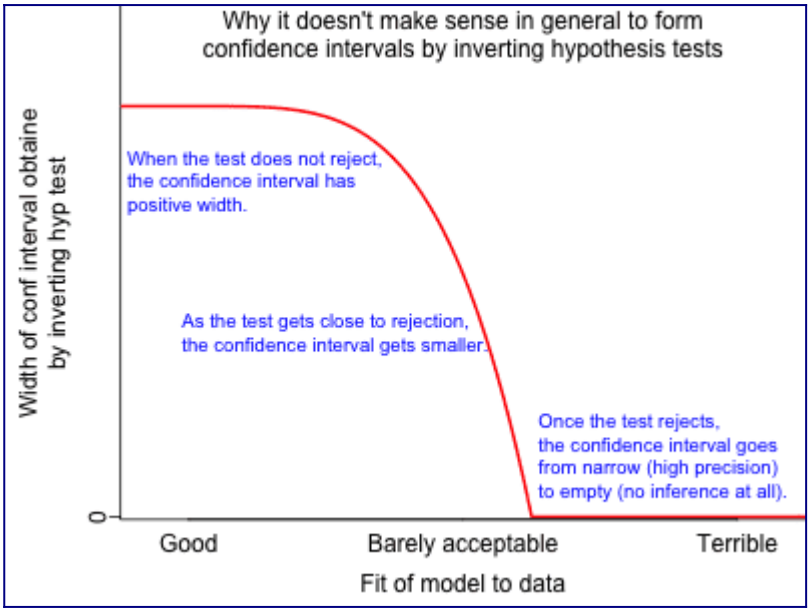
Hundreds of dead rats, sloppy file names: The anatomy of a retraction
retractionwatch.com
via PubPeer It all started – as more and more retractions do – with a post on PubPeer, this one in November 2021. The comment was about a paper titled “Efficient in vivo wound healing using noble met…
Shared by Retraction Watch

Is a “Wall of Shame” a good idea for journals?
retractionwatch.com
Today, the journal Cureus -- which is no stranger to Retraction Watch -- unveiled what they are calling a "Wall of Shame," which "highlights authors and reviewers who have committed egregious ethical…
Shared by Retraction Watch

The Ivermectin Guys’ Whole Thing Has Really Fallen Apart
vice.com
Meanwhile, they’re still pretending the drug is viable as a treatment for COVID.
Shared by Sam Finlayson



Confidence intervals, compatability intervals, uncertainty intervals
statmodeling.stat.columbia.edu
“Communicating uncertainty is not just about recognizing its existence; it is also about placing that uncertainty within a larger web of conditional probability statements. . . . No model can include…
Shared by Andrew Gelman et al.



Classification vs. Prediction | Statistical Thinking
fharrell.com

It is important to distinguish prediction and classification. In many decisionmaking contexts, classification represents a premature decision, because classification combines prediction and decision …

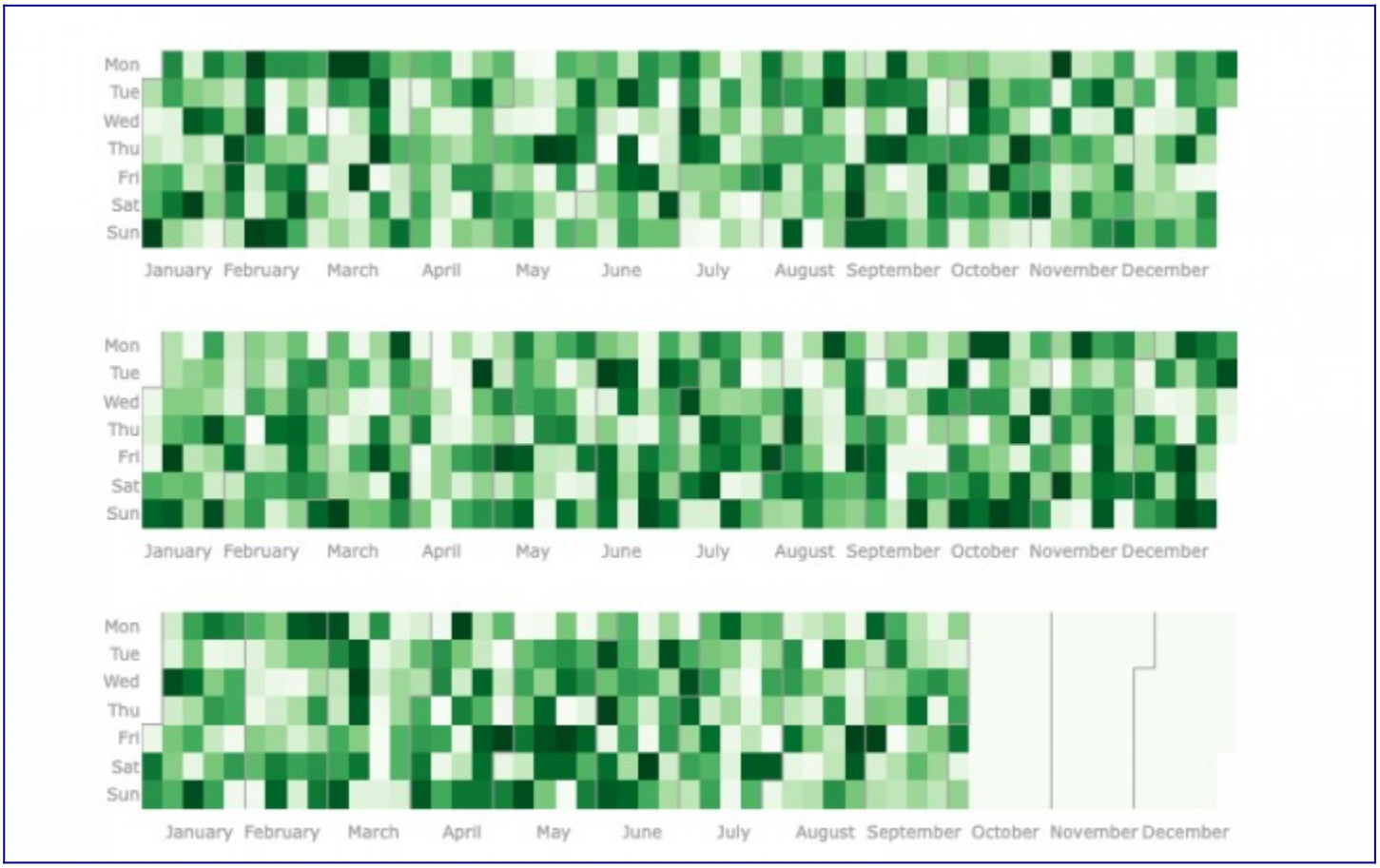
Shared by Frank Harrell

Interactive Calendar Heatmaps with Plotly, the easieast way you'll find

python.plainenglish.io

On my way to developing an application, I had a really hard time with making an interactive Calendar Heatmap (kinda like the one on Github and Gitlab), and thanks to this post on Plotly's community…

Shared by plotly



Collaboration Campfires

contributor.r-project.org

Collaborative sessions for newcomers to learn more about contributing to R.

Shared by Heather Turner

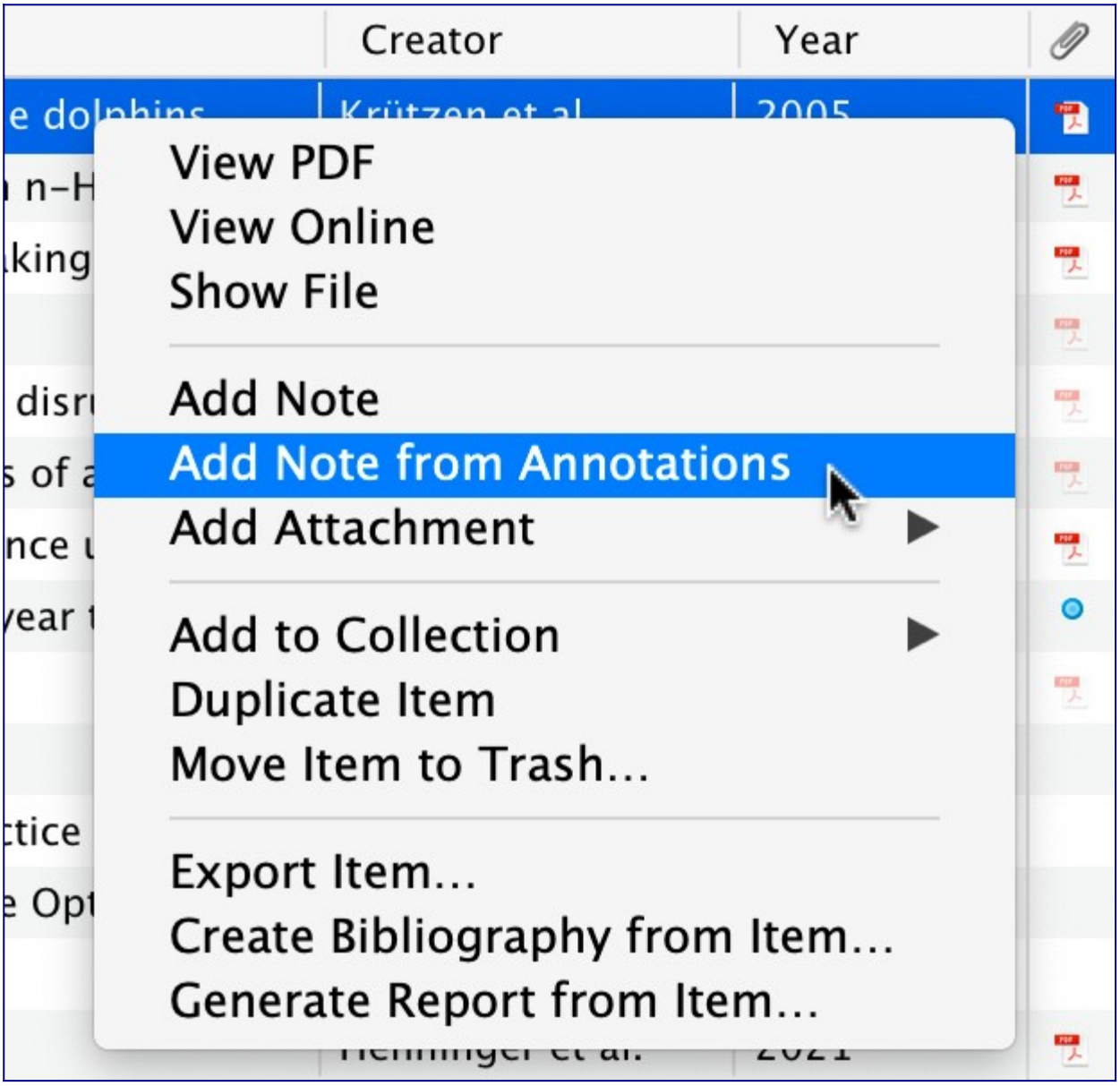
Institute of Mathematical Statistics

imstat.org

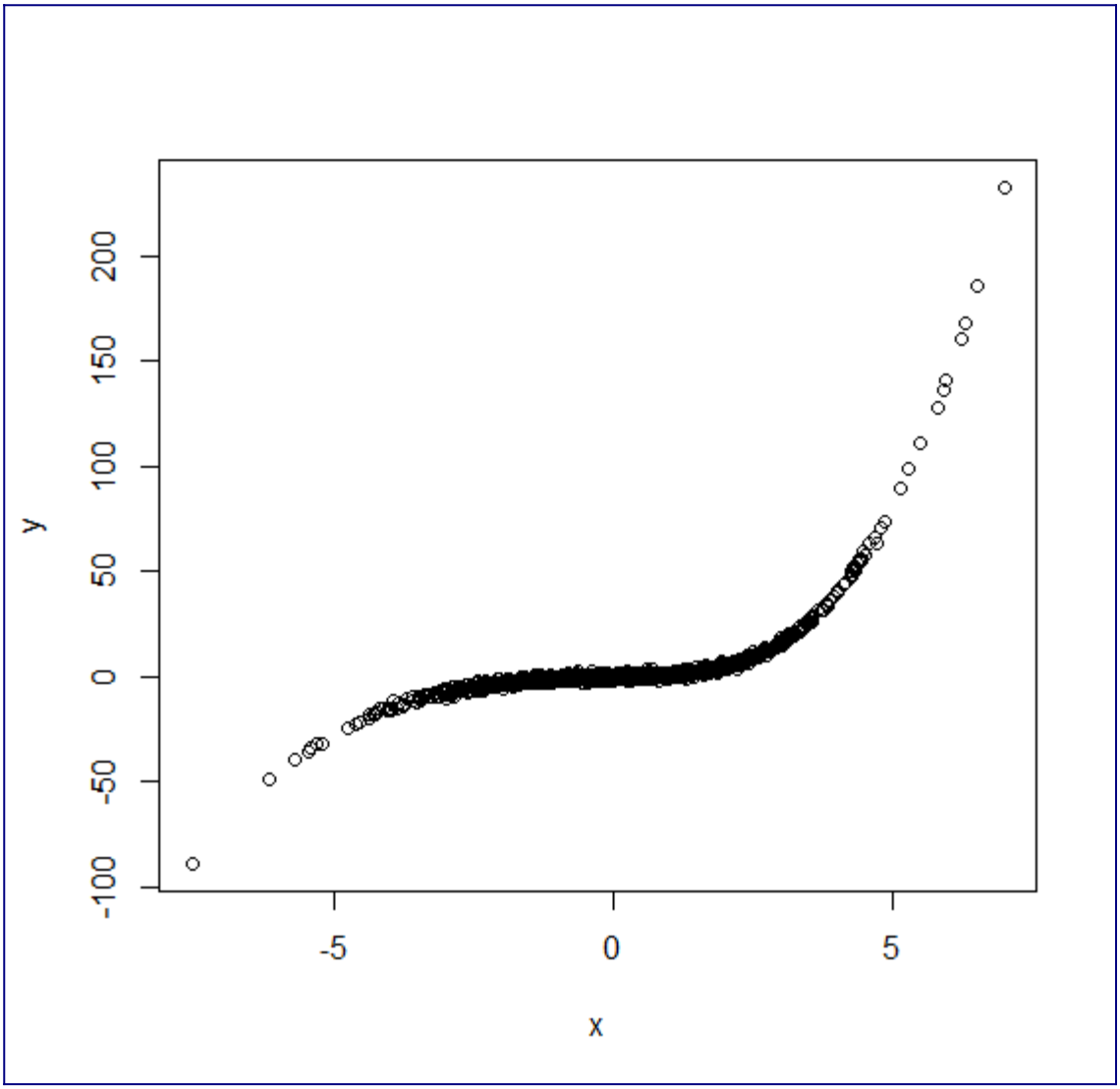
Herman Rubin, Professor of Statistics and Mathematics at Purdue University, passed away in West Lafayette, IN, on April 23, 2018; he was 91. Herman Rubin was among the last remaining great polymaths …

Shared by Stephen John Senn

Longitudinal Data: Think Serial Correlation First, Random Effects Second | Statistical Thinking
fharrell.com
Random effects/mixed effects models shine for multi-level data such as measurements within cities within counties within states. They can also deal with measurements clustered within subjects. There …
Zotero Blog » Blog Archive » Zotero 6: Your research workflow, transformed
zotero.org
Zotero is a free, easy-to-use tool to help you collect, organize, cite, and share research.
Shared by Venk Murthy, MD, PhD



[Multiple imputation with splines in R using smcfcs]
thestatsgeek.com
Tim Morris and I were recently discussing the topic of multiple imputation (MI) of covariates when one wants to assume the covariate affects the outcome via a spline of some kind. We thought that t…
Shared by Jonathan Bartlett

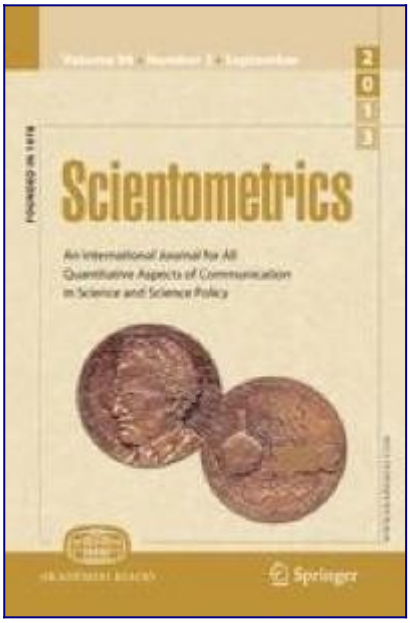


How to think about intellectual diversity in academia? The case of astrology and Dr. Oz
statmodeling.stat.columbia.edu
So, Columbia University has a lot of problems right now. For one thing, it seems we're falsifying or misrepresenting some of the statistics for our college ranking. Also, our most famous professor of…
Shared by Andrew Gelman et al.

The Coronavirus Funding Collapse Is a Disaster
theatlantic.com
All epidemics trigger the same Sisyphean cycle of panic and neglect. Even so, that cycle isn't meant to spin this quickly.
Shared by Karl Broman

You can't make this stuff up: Plagiarism guideline paper retracted for…plagiarism
retractionwatch.com
This could be an April Fools' joke. But it isn't. In what can only be described as an ironic twist, the Indian Journal of Dermatology is retracting a paper that presents guidelines on plagiarism for.…
Shared by Retraction Watch

Citation of retracted research: a case-controlled, ten-year follow-up scientometric analysis of Scott S. Reuben’s malpractice
link.springer.com
A major problem in scientific literature is the citation of retracted research. Until now, no long-term follow-up of the course of citations of such articl
Shared by Retraction Watch



Cancer Drugs Approved on the Basis of a Surrogate End Point and Subsequent Overall Survival: An Analysis of 5 Years of US Food and Drug Administratio…
jamanetwork.com
Most contemporary approvals of new cancer drugs are made on the basis of a surrogate end point, such as response rate or progression-free survival (PFS).1 When the approval is based on a surrogate en…
Shared by Vinay Prasad, MD MPH

Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests
pubmed.ncbi.nlm.nih.gov

Many decisions in medicine involve trade-offs, such as between diagnosing patients with disease versus unnecessary additional testing for those who are healthy. Net benefit is an increasingly reporte…

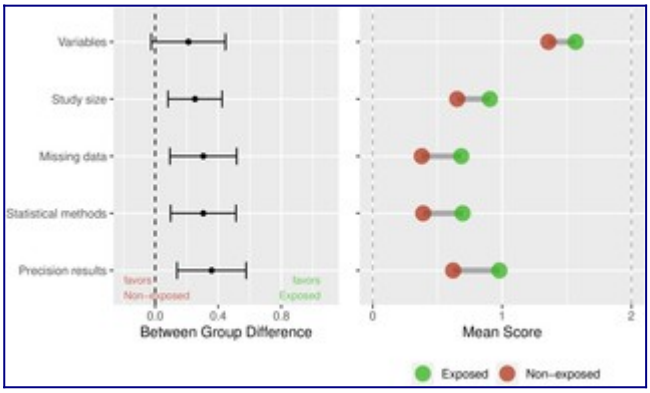
Shared by Ewout Steyerberg

The incremental value of the contribution of a biostatistician to the reporting quality in health research—A retrospective, single center, observatio…

journals.plos.org

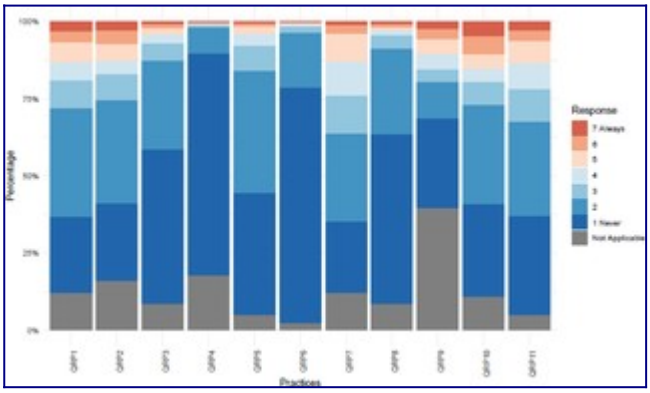
Background The reporting quality in medical research has recently been critically discussed. While reporting guidelines intend to maximize the value from funded research, and initiatives such as the …

Shared by Gary Collins



Prevalence of questionable research practices, research misconduct and their potential explanatory factors: A survey among academic researchers in Th…
journals.plos.org
Prevalence of research misconduct, questionable research practices (QRPs) and their associations with a range of explanatory factors has not been studied sufficiently among academic researchers. The …

Shared by Retraction Watch

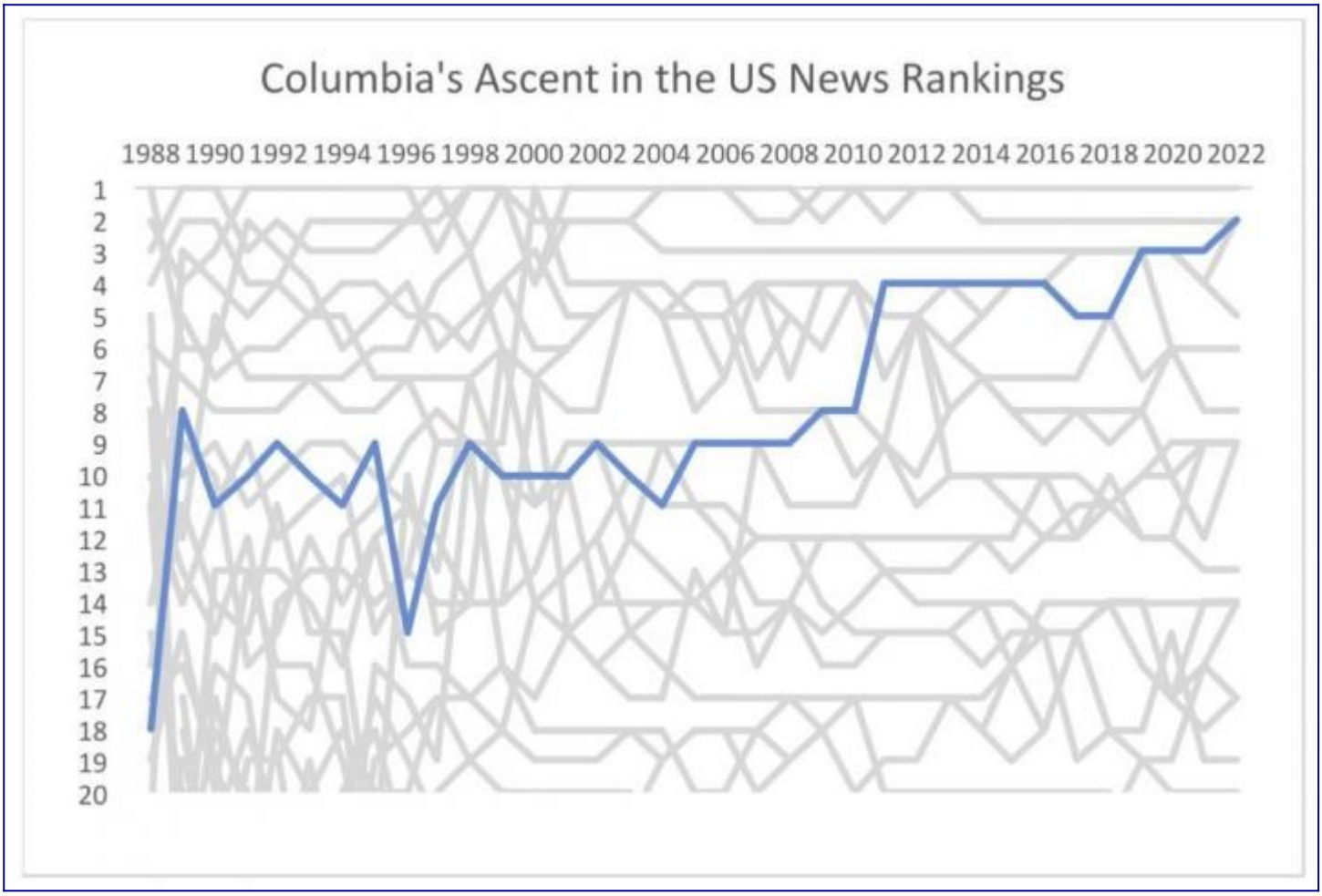


Viewpoints on Heterogeneity of Treatment Effect and Precision Medicine | Statistical Thinking
fharrell.com
This article provides my reflections after the PCORI/PACE Evidence and the Individual Patient meeting on 2018-05-31. The discussion includes a high-level view of heterogeneity of treatment effect in …

Shared by Frank Harrell

“An Investigation of the Facts Behind Columbia’s U.S. News Ranking”
statmodeling.stat.columbia.edu
Nearly forty years after their inception, the U.S. News rankings of colleges and universities continue to fascinate students, parents, and alumni. . . . A selling point of the U.S. News rankings is t…

Shared by Andrew Gelman et al.



Causal Directed Acyclic Graphs
jamanetwork.com
The design and interpretation of clinical studies requires consideration of variables beyond the exposure or treatment of interest and patient outcomes, including decisions about which variables to c…
Shared by Stephen John Senn

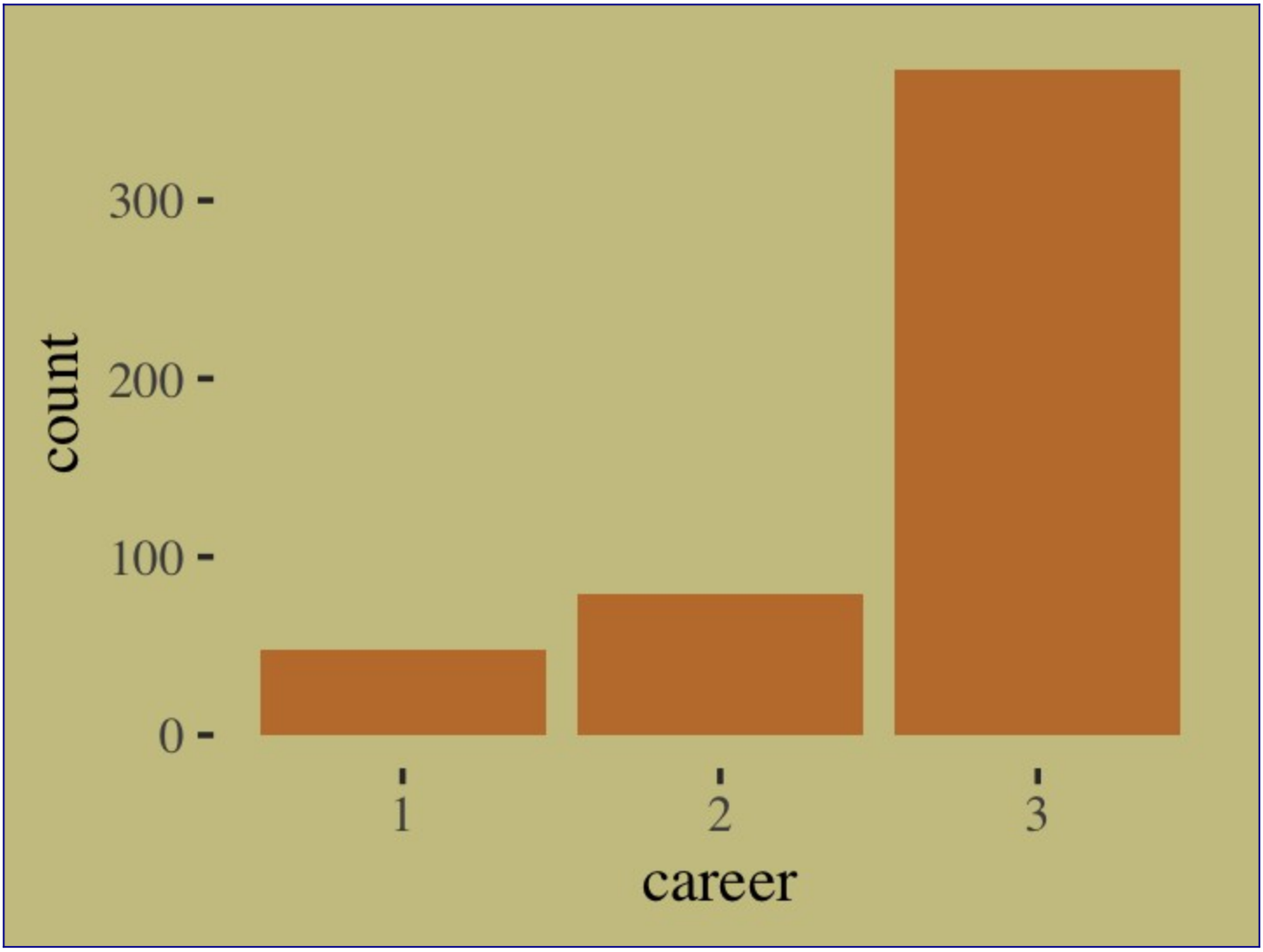
The reproducibility debate is an opportunity, not a crisis
bmcresnotes.biomedcentral.com
There are many factors that contribute to the reproducibility and replicability of scientific research. There is a need to understand the research ecosystem, and improvements will require combined ef…
Shared by Retraction Watch

11 God Spiked the Integers | Statistical rethinking with brms, ggplot2, and the tidyverse: Second edition

bookdown.org

This book is an attempt to re-express the code in the second edition of McElreath's textbook, 'Statistical rethinking.' His models are re-fit in brms, plots are redone with ggplot2, and the general d…

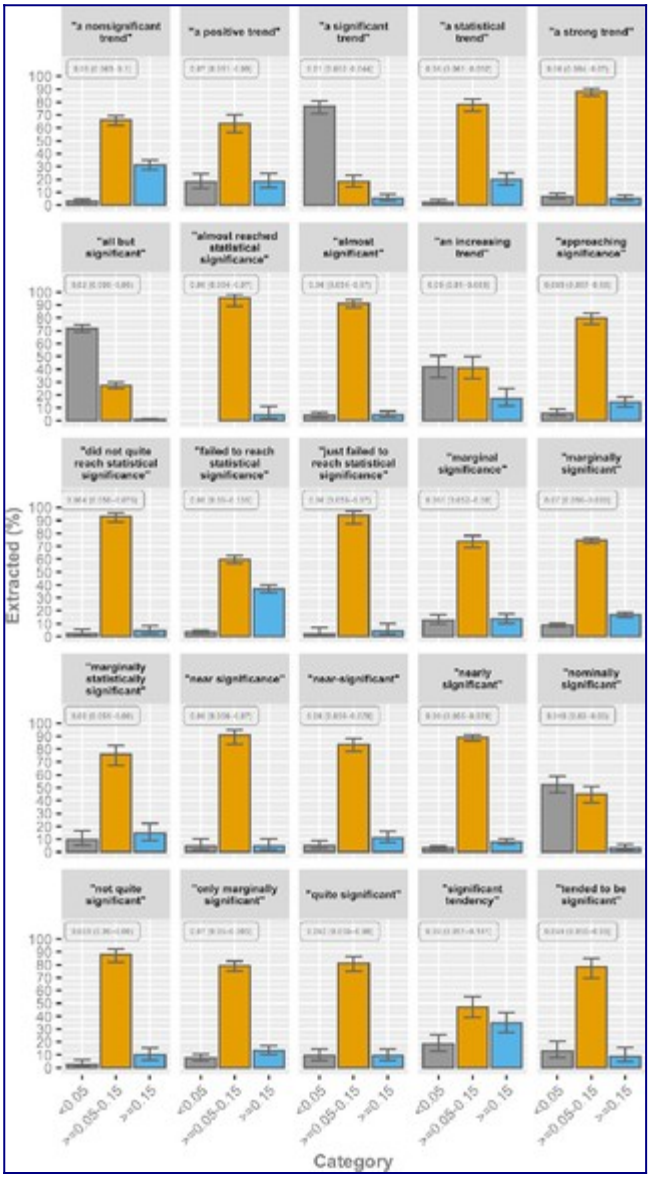
Shared by Darren Dahly



[Opinion | I’m a Parent and a Statistician. There’s a Smarter Way to Think About the Under-5 Vaccine.]
nytimes.com
We need to embrace a Bayesian approach to the child vaccine data.
Shared by Nathan T. James

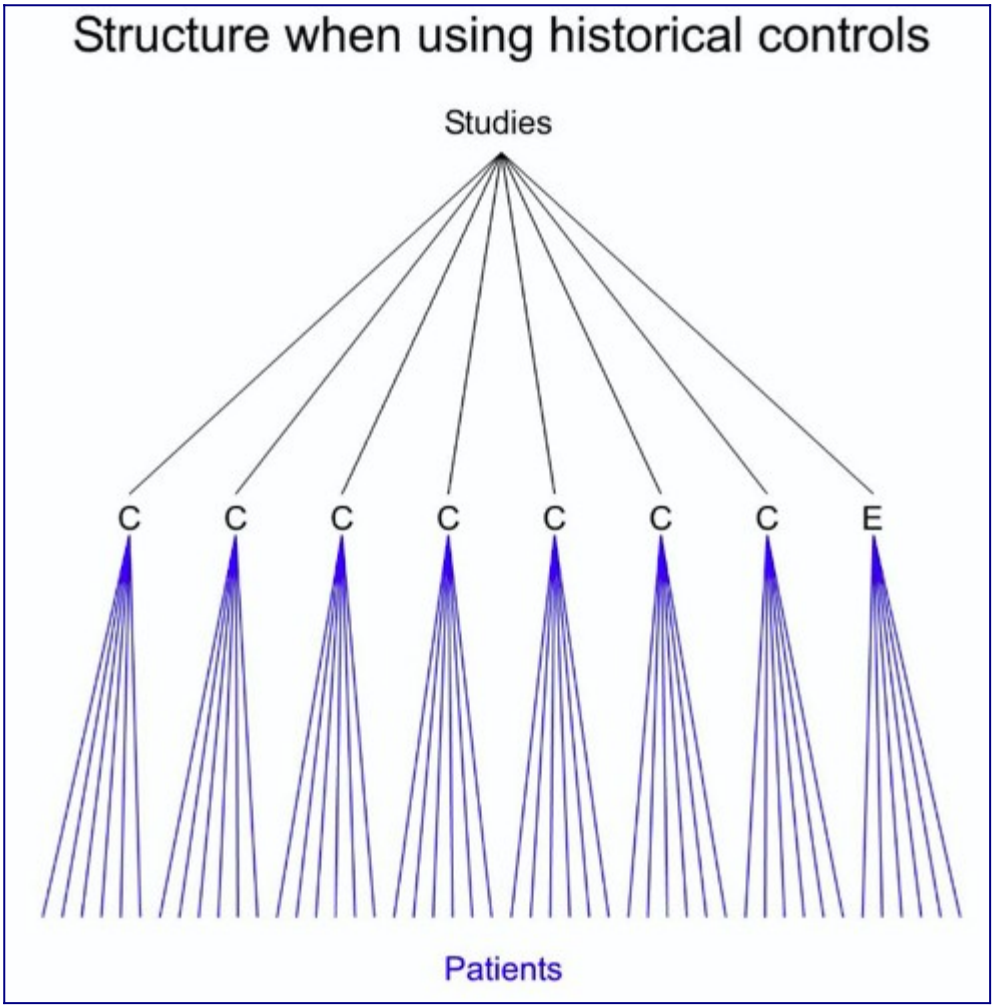
Where did IBM go wrong with Watson Health?
politpost.com
IBM spent more than a decade trying to make a go of Watson Health, its moonshot to apply artificial intelligence
Shared by Edward Tufte
The C.D.C. Isn’t Publishing Large Portions of the Covid Data It Collects
nytimes.com
The agency has withheld critical data on boosters, hospitalizations and, until recently, wastewater analyses.
Shared by Vinay Prasad, MD MPH

Live Free or Dichotomize - The Peril of Power when Prioritizing a Point Estimate
livefreeordichotomize.com
I recently noticed that the Pfizer immunobridging trials, presumably set up to demonstrate that their COVID-19 vaccines elicit the same antibody response in children as was seen in adolescents, for w…
Shared by Lucy D'Agostino McGowan
Analysis of 567,758 randomized controlled trials published over 30 years reveals trends in phrases used to discuss results that do not reach statisti…
journals.plos.org
The power of language to modify the reader's perception of interpreting biomedical results cannot be underestimated. An analysis of more than half a million randomized controlled trials reveals that …
Shared by Gary Collins



[Collaboration Campfires](#)
[contributor.r-project.org](#)
[Collaborative sessions for newcomers to learn more about contributing to R.](#)
Shared by [Heather Turner](#)

[Stephen Senn: Losing Control (guest post)](#)
[errorstatistics.com](#)
[Stephen Senn Consultant Statistician Edinburgh Losing Control Match points The idea of local control is fundamental to the design and analysis of experiments and contributes greatly to a design's e…](#)
Shared by [Stephen John Senn](#)



A U.S. federal science watchdog made just three findings of misconduct in 2021. We asked them why.
retractionwatch.com
Retraction Watch readers are likely familiar with the U.S. Office of Research Integrity (ORI), the agency that oversees institutional investigations into misconduct in research funded by the NIH, as …
Shared by Retraction Watch

Retraction in the online world—Shall we rethink the policy?
academic.oup.com
Abstract. In case of scientific misconduct, the papers can be retracted at the request of the journal's editors and/or authors. However, the retracted articles
Shared by Retraction Watch

Estimation of Study Time Reduction Using Surrogate End Points Rather Than Overall Survival in Oncology Clinical Trials
jamanetwork.com
Question How much time is saved in the 6 to 15 years of drug development by using surrogate end points rather than overall survival as the basis for US Food and Drug Administration approval in contem…
Shared by Vinay Prasad, MD MPH

SBC: A new R package for simulation-based calibration of Bayesian models!
statmodeling.stat.columbia.edu

Shinyoung Kim, Angie Hyunji Moon, Martin Modrák, and Teemu Säilynoja prepared this R package for simulation-based calibration (SBC)! They write: SBC provides tools to validate your Bayesian model and…

Shared by Andrew Gelman et al.

```
 - 35 (70%) fits had some steps rejected. Maximum number of rejections was 4.
Not all diagnostics are OK.
You can learn more by inspecting $default_diagnostics, $backend_diagnostics
and/or investigating $outputs/$messages/$warnings for detailed output from the backend.
```

Fake Histories, Calls by Staff Impersonating Subjects: A Look Inside Two Trial Fraud Cases | JD Supra
jdsupra.com

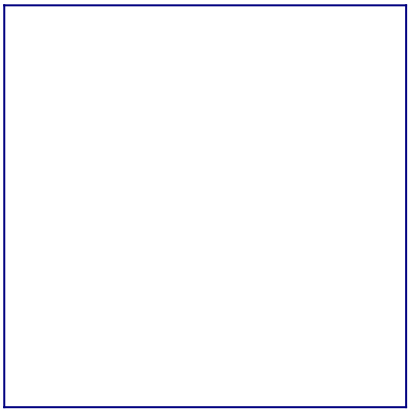
Report on Research Compliance 19, no. 2 (January 27, 2022) - The trial was to be like any other that the clinical research organization (CRO) would oversee. The...

Shared by Retraction Watch

Release of CmdStan 2.29
blog.mc-stan.org

We are very happy to announce that the 2.29.0 release of CmdStan is now available on Github! As usual, the release of CmdStan is accompanied by new releases of Stan Math, core Stan and Stanc3. A PySt…

Shared by Stan



Medical research needs a Sue Gray

bmj.com

Research in medicine needs someone with the power to investigate and uncover misconduct by professionals and organisations who abuse people's trust. Trial participants have every faith in the researc…

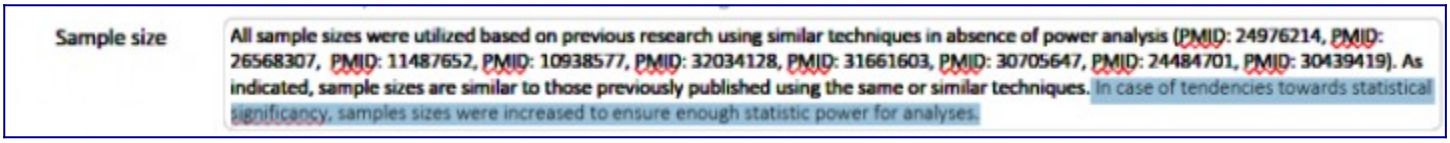
Shared by Retraction Watch

The joke of self-correcting science: The Andero lab and Nature Communications

thenewstatistics.com

In 2017, I (Bob) attended a National Academies of Science conference on reproducibility in science. Among many memorable events, I saw a talk by David Allison about his efforts to correct simple stat…

Shared by Retraction Watch



Guest Post — Research Integrity: Ensuring Trust in Global Research
scholarlykitchen.sspnet.org
A look at developments in research integrity, and the attempt to build a universal culture of ethical and responsible practice in research as well as systems within the overall research ecosystem for…
Shared by Retraction Watch

What happens if the explanatory and response variables are sorted independently before regression?
stats.stackexchange.com
Suppose we have data set (X_i,Y_i) with n points. We want to perform a linear regression, but first we sort the X_i values and the Y_i values independently of each other, forming data set $...
Shared by Maarten van Smeden

[Mask studies reach a new scientific low point](#)
[vinayprasadmdmph.substack.com](#)
[MMWR appears to no longer be in the science business](#)
Shared by [Vinay Prasad, MD MPH](#)

Characteristic	No. (%) Case-participants (SARS-CoV-2–positive) N = 652	Control-participants (SARS-CoV-2–negative) N = 1,176
Reasons for SARS-CoV-2 testing**		
Experiencing symptoms	508 (77.9)	196 (16.7)
Testing required for medical procedure	40 (6.1)	199 (16.9)
Routine screening through work or school	71 (10.9)	507 (43.1)
Pre-travel test	33 (5.1)	120 (10.2)
Just wanted to see if I was infected	65 (10.0)	172 (14.6)
Test required for admission to an event or gathering	3 (0.5)	21 (1.8)

[Using the change in the value of a predictor variable as a parameter in a multivariable prognostic model](#)

discourse.datamethods.org

I have been tasked with developing a prognostic model. This particular area of medicine is fairly unique as many “risk factors” for prognosis have been identified and replicated, but not combined in …

The different lives of fringe and strange scientific ideas | Aeon Essays

aeon.co

Believe in the Loch Ness monster and you’re more likely to believe the Apollo missions were fake. How do weird beliefs work?

Shared by Aeon+Psyche

The media has lost the plot on ivermectin

salon.com

Molnupiravir, Paxlovid, and Sotrovimab all treat COVID-19 very well. Why are we fixating on ivermectin?

Shared by Јаков Минг Дановић

Statistics notes: Absence of evidence is not evidence of absence
bmj.com
The non-equivalence of statistical significance and clinical importance has long been recognised, but this error of interpretation remains common. Although a significant result in a large study may s…
Shared by Juan Senior

Unit of Biostatistics, Epidemiology and Public Health at Padova University on LinkedIn: #DarioGregori #GiuliaLorenzoni #FrankHarrell
linkedin.com
Watch on our YouTube Channel "Controversies in Statistical Learning in Clinical Epidemiology" Day 1- PhD course in Translational Specialistic Medicine...
Shared by Frank Harrell

Bayesian approach to comparing common models of life-course epidemiology
academic.oup.com
AbstractBackground. Life-course epidemiology studies people's health over long periods, treating repeated measures of their experiences (usually risk factors) a
Shared by Adan Z Becerra, PhD

How Margaret Mead became a hate figure for conservatives | Aeon Essays
aeon.co

Mead argued that non-Western cultures offered alternative (often better) ways to be human. Why was she so vilified for it?

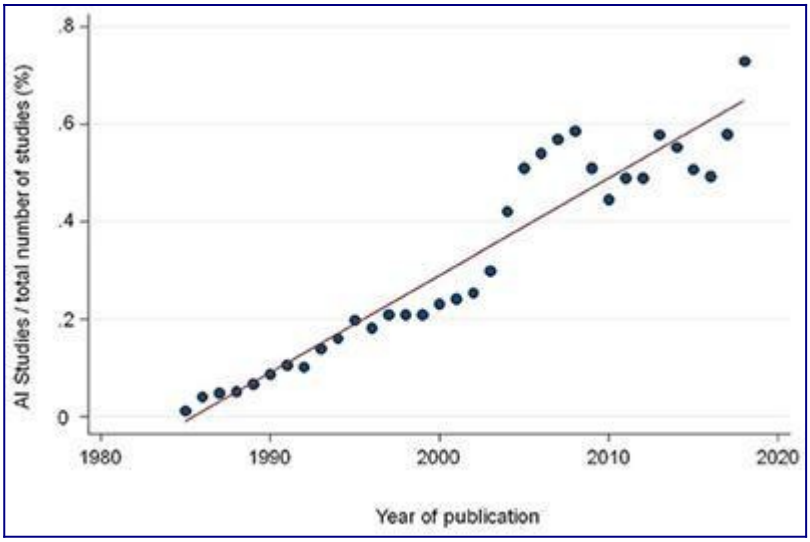
Shared by Aeon+Psyche

Artificial Intelligence and Statistics: Just the Old Wine in New Wineskins?

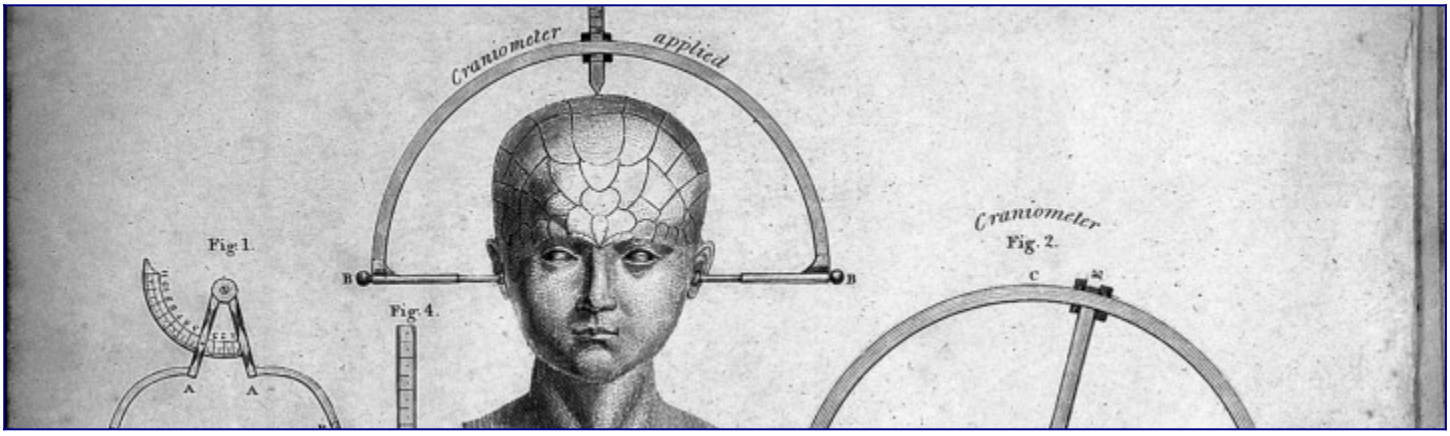
frontiersin.org

We are witnessing a tremendous increase in scientific studies using Artificial Intelligence (AI) and in particular Machine Learning (ML) methods in the medical literature. A recent systematic review …

Shared by Maarten van Smeden

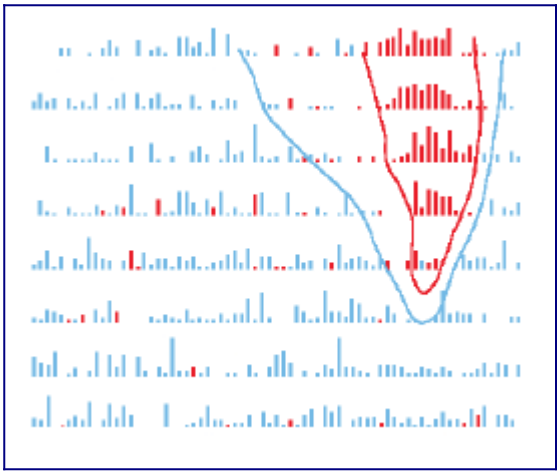


How Eugenics Shaped Statistics
nautil.us
Exposing the damned lies of three science pioneers.
Shared by Frank Harrell



These Doctors Wanted to Get Schools Back to Normal. Their Botched Effort May Backfire.
motherjones.com
The first rule of science: Check your facts before publishing.
Shared by Boback Ziaeian

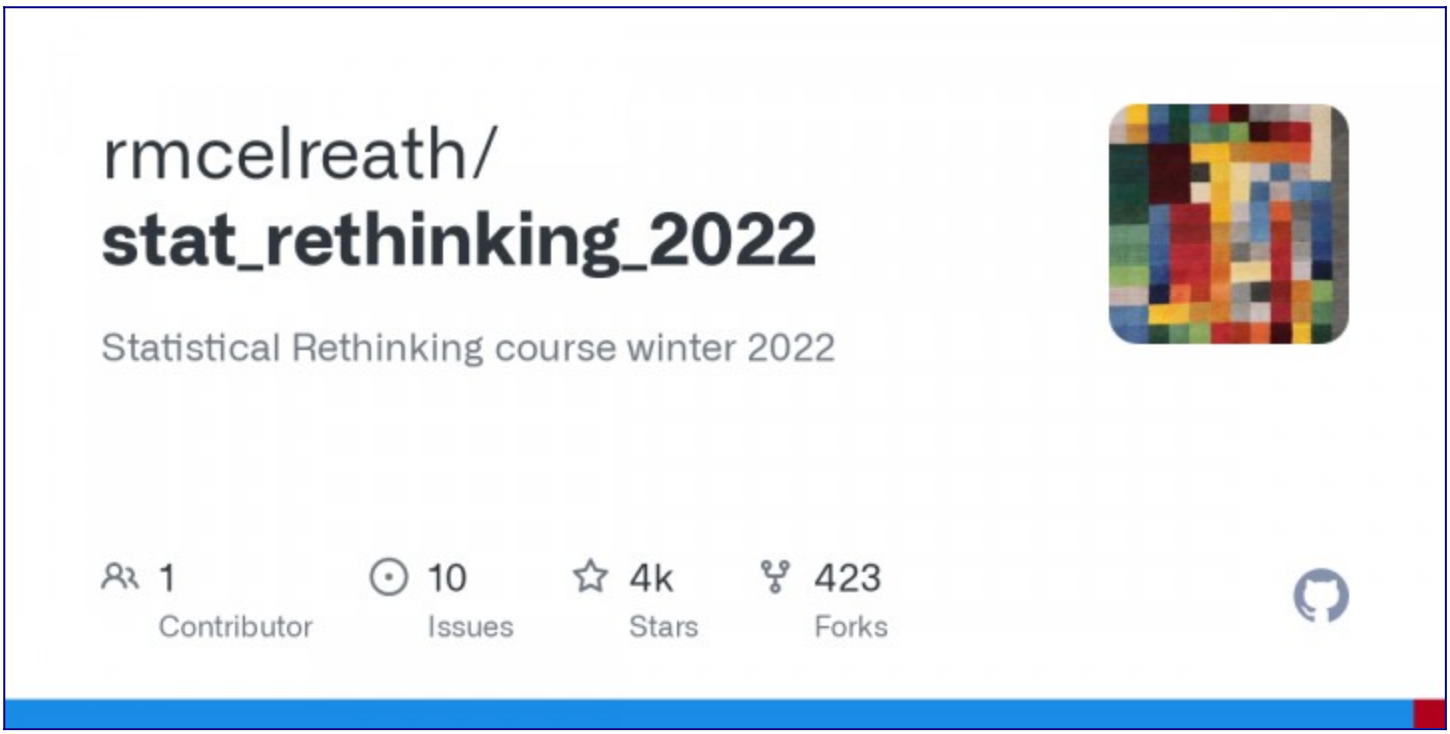
Edward Tufte forum: History of Sparklines and Edward Tufte Contribution
edwardtufte.com
The essay provides a history of my work creating the concept of sparklines, and then a long history of sparkline-like graphics for hundreds of years. Sparklines are small, intense, word-sized graphic…
Shared by Edward Tufte



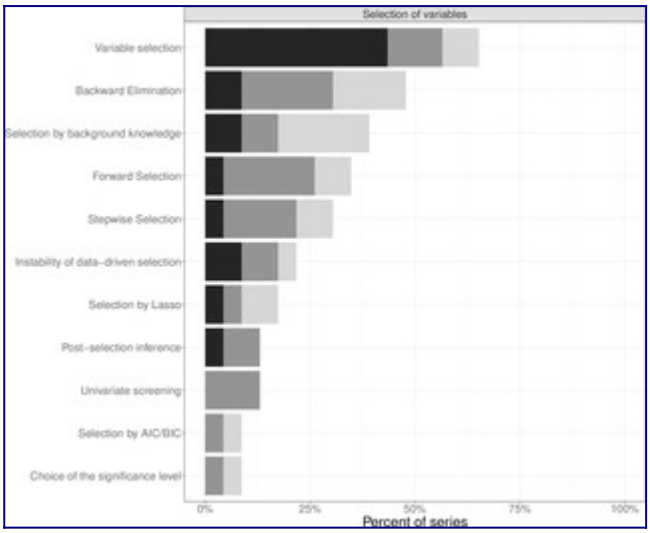
23 - The Difference-in-Differences Saga — Causal Inference for the Brave and True
matheusfacure.github.io
Like I’ve said, panel data is when we have multiple units i over multiple periods of time t. Think about a policy evaluation scenario in the US, where you want to check the effect of cannabis legaliz…
Shared by Darren Dahly

Two stage hierarchical model for multiple outcomes
discourse.datamethods.org
In RCTs in cognition research, the primary outcomes are usually summary measures (e.g., a weighted average of several reaction time tasks) which putatively measure some latent cognitive ability, such…

GitHub - rmcelreath/stat_rethinking_2022: Statistical Rethinking course winter 2022
github.com
Statistical Rethinking course winter 2022. Contribute to rmcelreath/stat_rethinking_2022 development by creating an account on GitHub.
Shared by Richard McElreath



Review of guidance papers on regression modeling in statistical series of medical journals
journals.plos.org
Although regression models play a central role in the analysis of medical research projects, there still exist many misconceptions on various aspects of modeling leading to faulty analyses. Indeed, t…
Shared by Giuseppe Biondi-Zoccai

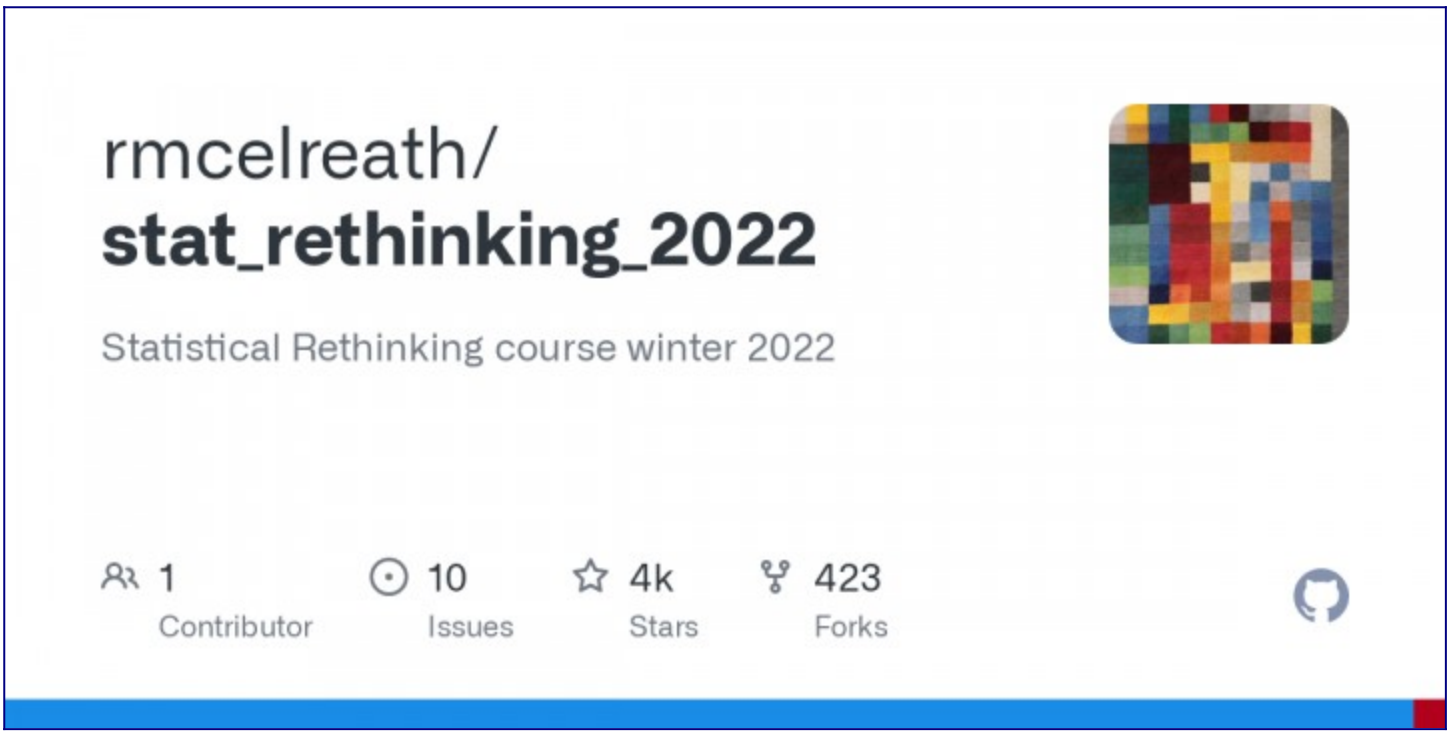


Protecting against researcher bias in secondary data analysis: challenges and potential solutions
link.springer.com
Analysis of secondary data sources (such as cohort studies, survey data, and administrative records) has the potential to provide answers to science and so
Shared by Brian Nosek

[Live Free or Dichotomize - Would seeing Spider-Man: No Way Home decrease COVID-19 Cases?](#)
livefreeordichotomize.com
In SNL's cold open last night, "President Joe Biden" suggested that the COVID-19 surge we are seeing in the US is due to people seeing Spider-Man: No Way Home. If people would just stop seeing this f…
Shared by Lucy D'Agostino McGowan

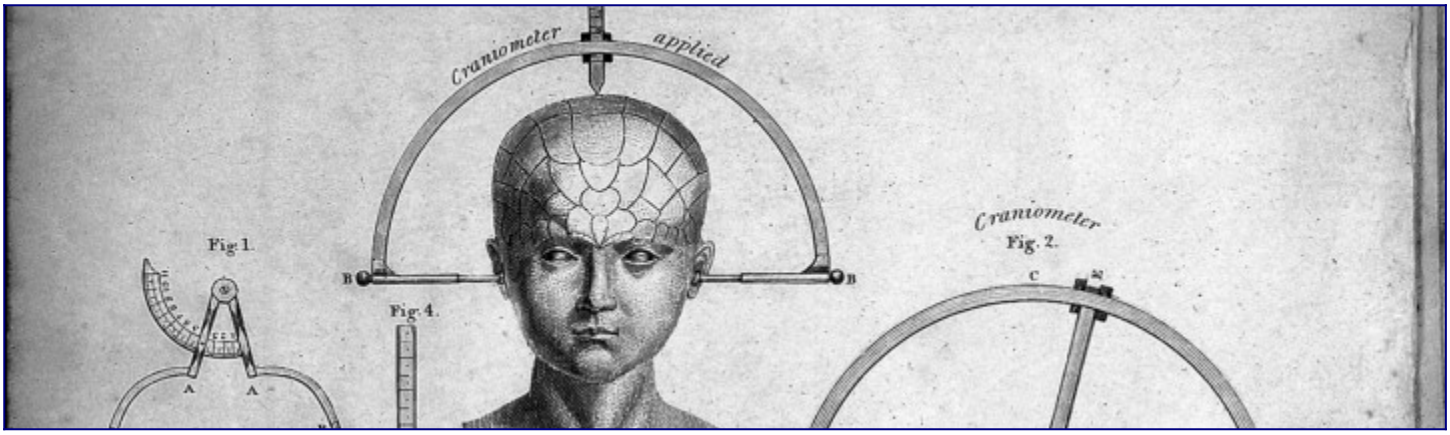
[Paradigm lost: Carl Liebermeister and the development of modern medical statistics. - The James Lind Library](#)
jameslindlibrary.org
Introduction The role of probabilistic reasoning in medicine has been a source of controversy for millennia (Matthews 2020). This reflects its implications for such vexed questions as the extent to ….
Shared by Stephen John Senn

[GitHub - rmcelreath/stat_rethinking_2022: Statistical Rethinking course winter 2022](#)
github.com
Statistical Rethinking course winter 2022. Contribute to rmcelreath/stat_rethinking_2022 development by creating an account on GitHub.
Shared by Richard McElreath



Most medical practices are not parachutes: a citation analysis of practices felt by biomedical authors to be analogous to parachutes
ncbi.nlm.nih.gov
In a 2003 paper in BMJ, the authors made the tongue-in-cheek observation that there are no randomized controlled trials (RCTs) of parachutes. This paper has been widely read, cited and used to argue …
Shared by Vinay Prasad, MD MPH

How Eugenics Shaped Statistics
nautil.us
In early 2018, officials at University College London were shocked to learn that meetings organized by "race scientists" and neo-Nazis,…
Shared by Aubrey Clayton



[Not-so-obviously heuristic-proof reforms to statistical communication](#)
statmodeling.stat.columbia.edu
[This is Jessica. I've subscribed to aspects of the "estimation" movement–the move toward emphasizing magnitude and uncertainty of effects and testing multiple hypotheses rather than NHST–for awhile, …](#)
Shared by [Andrew Gelman et al.](#)

[The pandemic's true death toll: millions more than official counts](#)
nature.com
[Countries have reported some five million COVID-19 deaths in two years, but global excess deaths are estimated at double or even quadruple that figure.](#)
Shared by [C. Michael Gibson MD](#)

[Why are not There More Bayesian Clinical Trials? Perceived Barriers and Educational Preferences Among Medical Researchers Involved in Drug Developmen…](#)
link.springer.com

The clinical trials community has been hesitant to adopt Bayesian statistical methods, which are often more flexible and efficient with more naturally inte

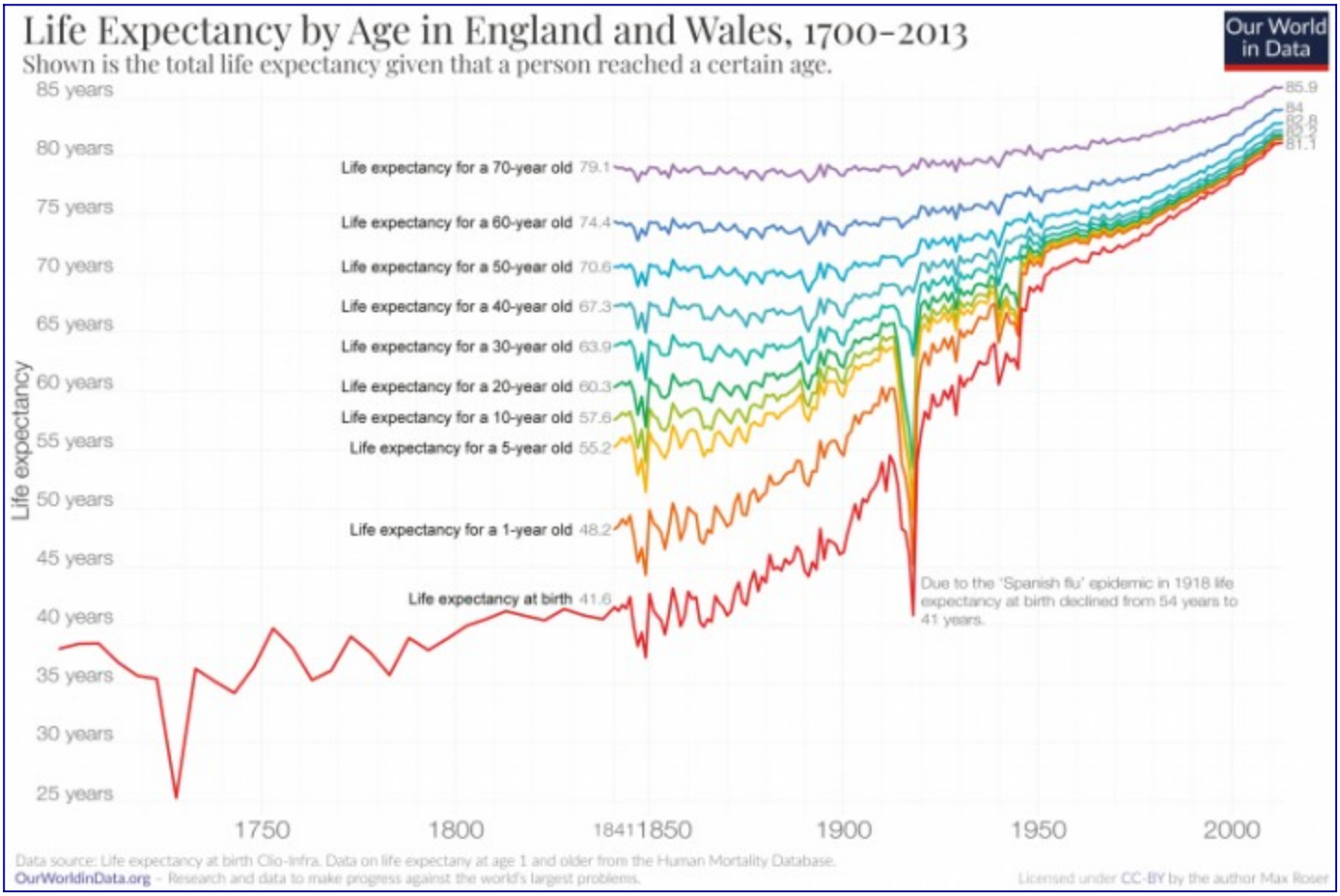
Shared by Frank Harrell

It’s not just about child mortality, life expectancy improved at all ages

ourworldindata.org

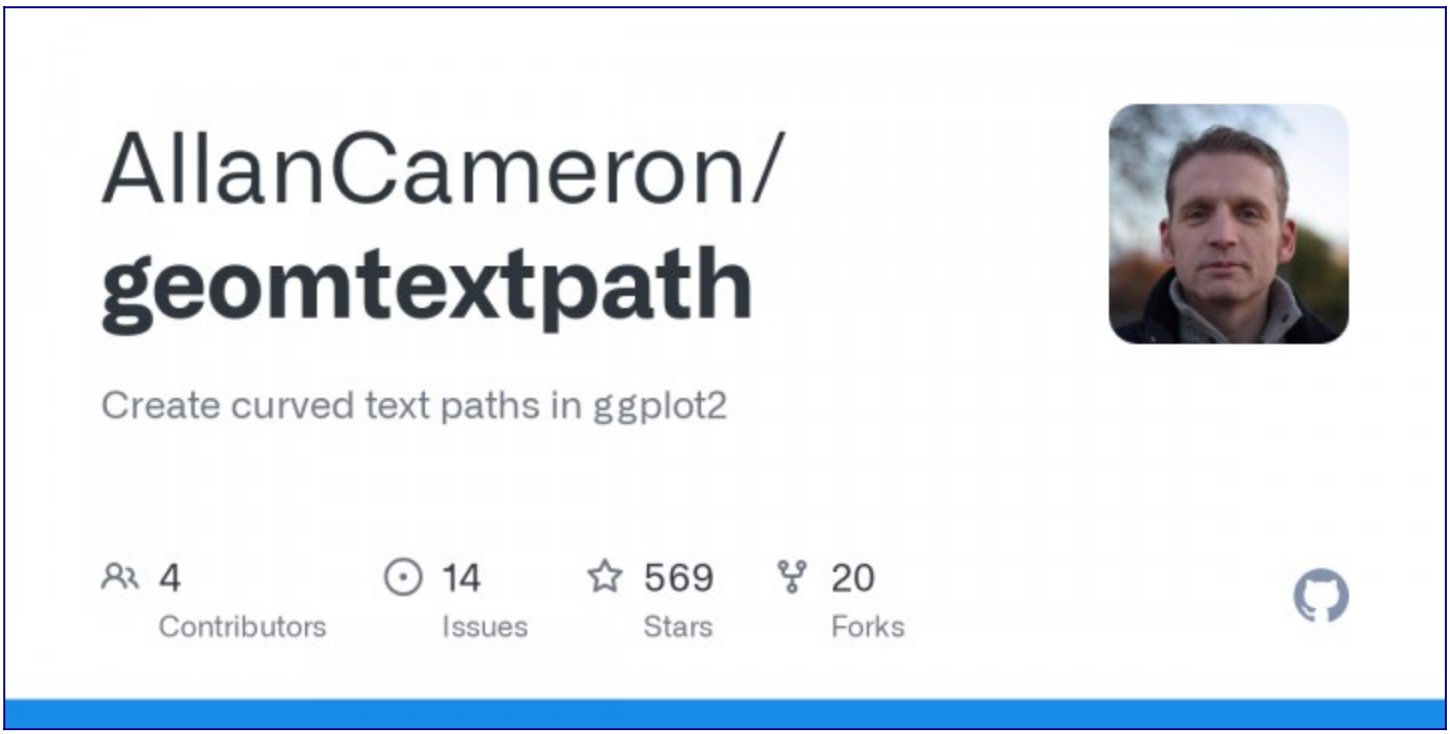
It’s often argued that life expectancy across the world has only increased because child mortality has fallen. But this untrue. The data shows that life expectancy has increased at all ages.

Shared by Sam Finlayson



GitHub - AllanCameron/geomtextpath: Create curved text paths in ggplot2
github.com
Create curved text paths in ggplot2. Contribute to AllanCameron/geomtextpath development by creating an account on GitHub.

Shared by Lucy D'Agostino McGowan



[2021 Shkreli Awards]
[lowninstitute.org]
[A top ten list of the worst examples of profiteering and dysfunction in health care, named for Martin Shkreli, the price-hiking "pharma bro" that everyone loves to hate.]
Shared by [Allen Frances]

[Reflection on modern methods: five myths about measurement error in epidemiological research]
[academic.oup.com]

Abstract. Epidemiologists are often confronted with datasets to analyse which contain measurement error due to, for instance, mistaken data entries, inaccurate

Shared by Maarten van Smeden

Estimands: bringing clarity and focus to research questions in clinical trials

bmjopen.bmj.com

Precise specification of the research question and associated treatment effect of interest is essential in clinical research, yet recent work shows that they are often incompletely specified. The ICH…

Shared by Andrew Althouse

Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review

pubmed.ncbi.nlm.nih.gov

While the opportunities of ML and AI in healthcare are promising, the growth of complex data-driven prediction models requires careful quality and applicability assessment before they are applied and…

Shared by Carl Moons

Efficient design and analysis of randomized controlled trials in rare neurological diseases: An example in Guillain-Barré syndrome

pubmed.ncbi.nlm.nih.gov

Covariate adjustment and proportional odds analysis most efficiently use the available data and ensure balance between the treatment arms to obtain reliable and valid treatment effect estimates. Thes…

Shared by Ewout Steyerberg

Determine power to detect a misclassification event in rule-out pathways for myocardial infarction
discourse.datamethods.org
The question concerns work we have recently conducted, looking at risk prediction using a serum biomarker in a study population presenting with chest pain to emergency departments. In clinical medici…

Performance of binary prediction models in high-correlation low-dimensional settings: a comparison of methods

diagnprognres.biomedcentral.com

Clinical prediction models are developed widely across medical disciplines. When predictors in such models are highly collinear, unexpected or spurious predictor-outcome associations may occur, there…

Shared by Maarten van Smeden

Prediction or causality? A scoping review of their conflation within current observational research

link.springer.com

Etiological research aims to uncover causal effects, whilst prediction research aims to forecast an outcome with the best accuracy. Causal and prediction r

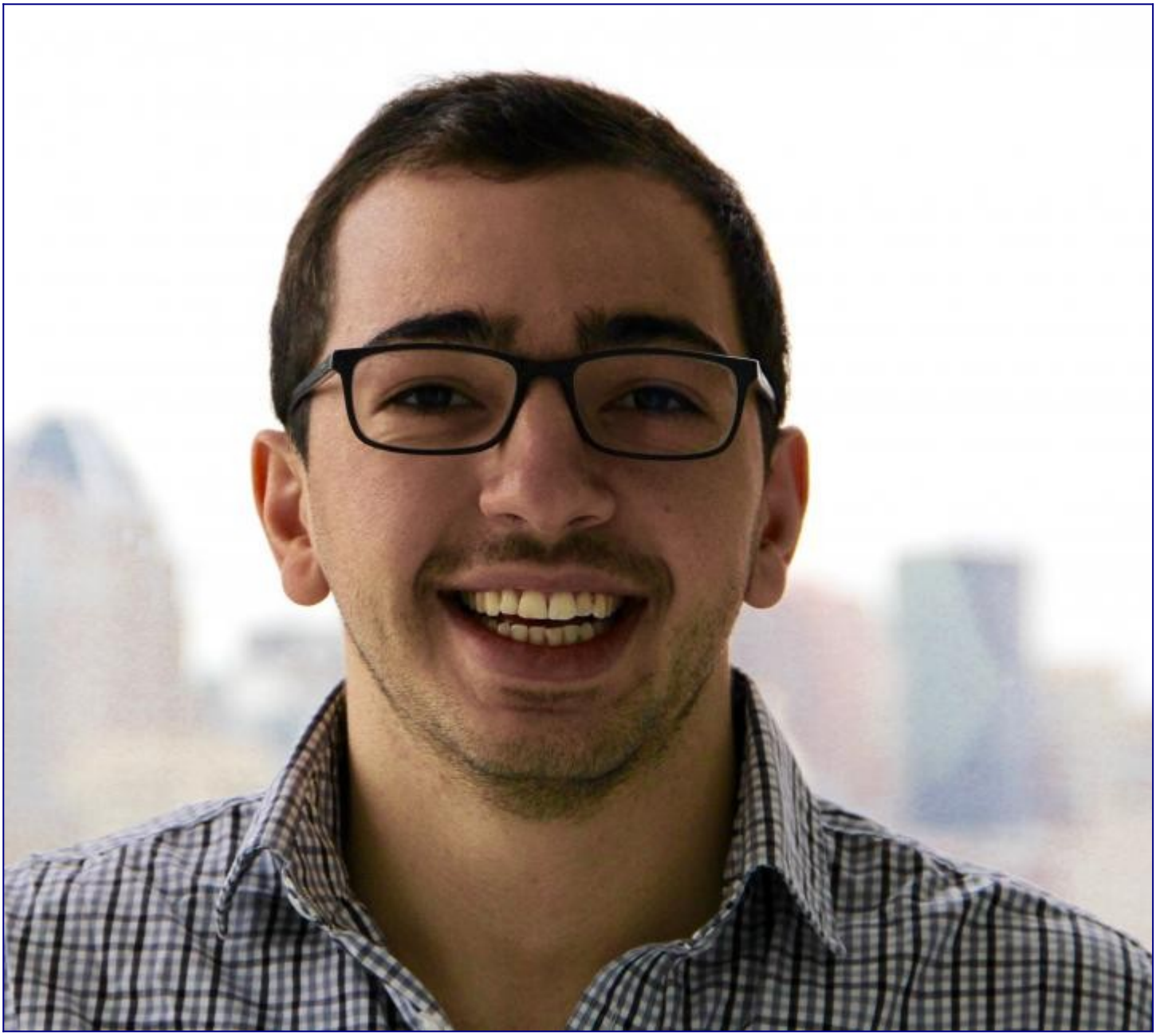
Shared by HN@病院総合診療医

[Paper claiming a lack of evidence COVID-19 lockdowns work is retracted](#)
retractionwatch.com
[This is an evolving story, and we will update as we learn more. A paper in Springer Nature's Scientific Reports claiming there was essentially no evidence that lockdowns prevented COVID-19 deaths has…](#)
Shared by [Retraction Watch](#)
[Can you capture the complex reality of the pandemic with numbers? Well, we tried…](#)
theguardian.com
[Throughout 2021, two leading lights of the Royal Statistical Society Covid-19 Task Force drew on data for a weekly Observer column, and found themselves in the middle of Covid culture wars](#)
Shared by [David Spiegelhalter](#)

[Why I teach my students about scientific failure](#)
science.org

“It’s a lesson I wish I’d learned before starting grad school,” this professor writes
Shared by Retraction Watch
How a largely untested AI algorithm crept into hundreds of hospitals
fastcompany.com
During the pandemic, the electronic health record giant Epic quickly rolled out an algorithm to help doctors decide which patients needed the most immediate care. Doctors believe it will change how t…
Shared by fml.lol

The complexity of dealing with missing data, variable selection, and cross validation for prediction models
jfiksel.github.io
Part of my research is developing prediction models using medical data collected from electronic health records (EHR). Three of the largest issues I have to deal with are: Missing predictor values. T…
Shared by Jacob Fiksel, PhD



At what point does incompetence become fraud?
medium.com
In March of 2021 a paper was published by the journal Plastic and Reconstructive Surgery Global Open, the open access only (meaning pay-to-publish) version of Plastic and Reconstructive Surgery. This…
Shared by Retraction Watch
A Comparison of Decision Curve Analysis with Traditional Decision Analysis | Statistical Thinking
fharrell.com
Andrew Vickers Department of Epidemiology and Biostatistics Memorial Sloan Kettering Cancer Center vickersa@mskcc.org Introduction In a traditional decision analysis, the analyst creates a decision t…
Health Care and Journalism Are Facing the Same Crises

slate.com

And people are dying as a result.

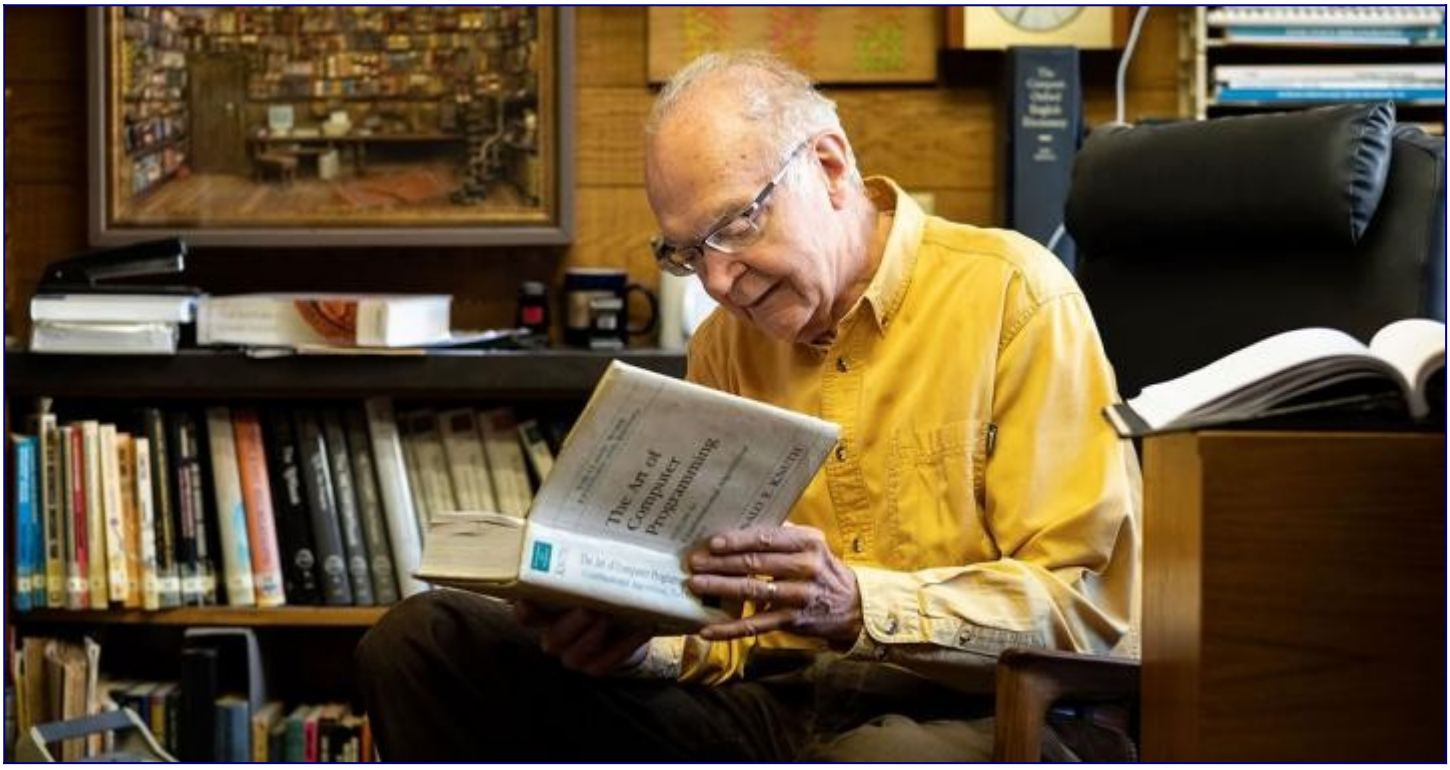
Shared by Retraction Watch

The Computer Scientist Who Can’t Stop Telling Stories

quantamagazine.org

For pioneering computer scientist Donald Knuth, good coding is synonymous with beautiful expression.

Shared by Arthur Charpentier



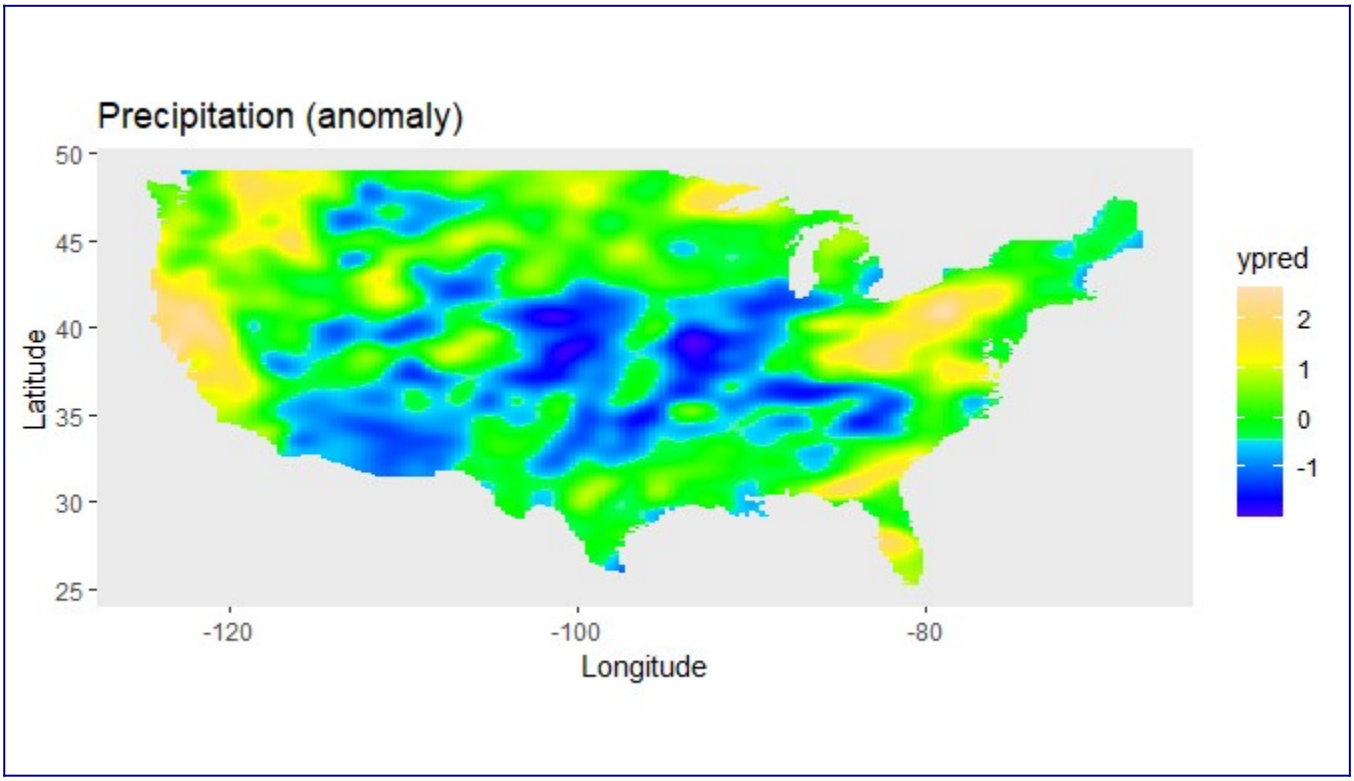
[At a time when the U.S. needed Covid-19 dialogue between scientists, Francis Collins moved to shut it down]

statnews.com

In October 2020, the U.S. needed scientific debate on how best to control Covid-19. NIH director Francis Collins tried to quash it.

Shared by Vinay Prasad, MD MPH

[November 2021: "Top 40" New CRAN Packages]
[rviews.rstudio.com]
[An R community blog edited by RStudio]
Shared by [RStudio]



[Individual response]

discourse.datamethods.org

An intriguing post on Twitter by Judea Pearl linked to a blog written jointly with Scott Mueller Personalized Decision Making . The blog stated “The purpose of this note is to clarify the distinction…

Shared by Frank Harrell

How economists rode maths to become our era’s astrologers | Aeon Essays

aeon.co

By fetishising mathematical models, economists turned economics into a highly paid pseudoscience

Shared by Darren Dahly

Will innumeracy cause this study to be retracted? Don't count on it...

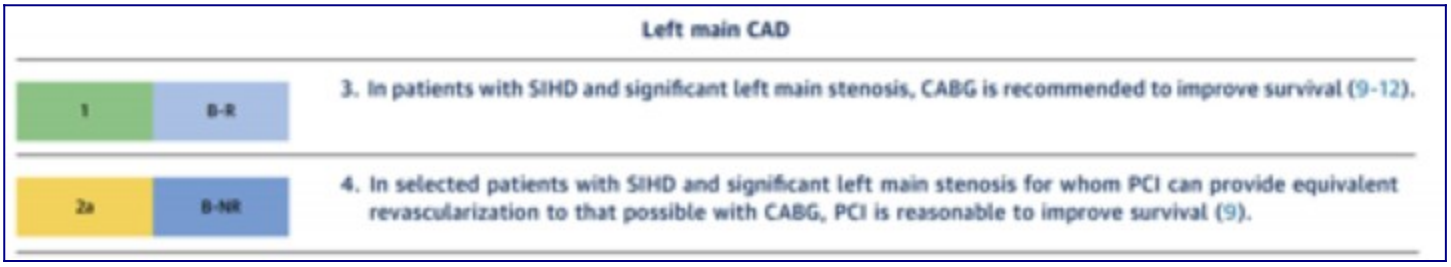
thesciencebit.net

I used to be concerned about bad science. These days, what gets me going is wrong science: blatant error somehow surviving peer-review and ending up published as if it were fact. It seems that is whe…

Shared by Retraction Watch

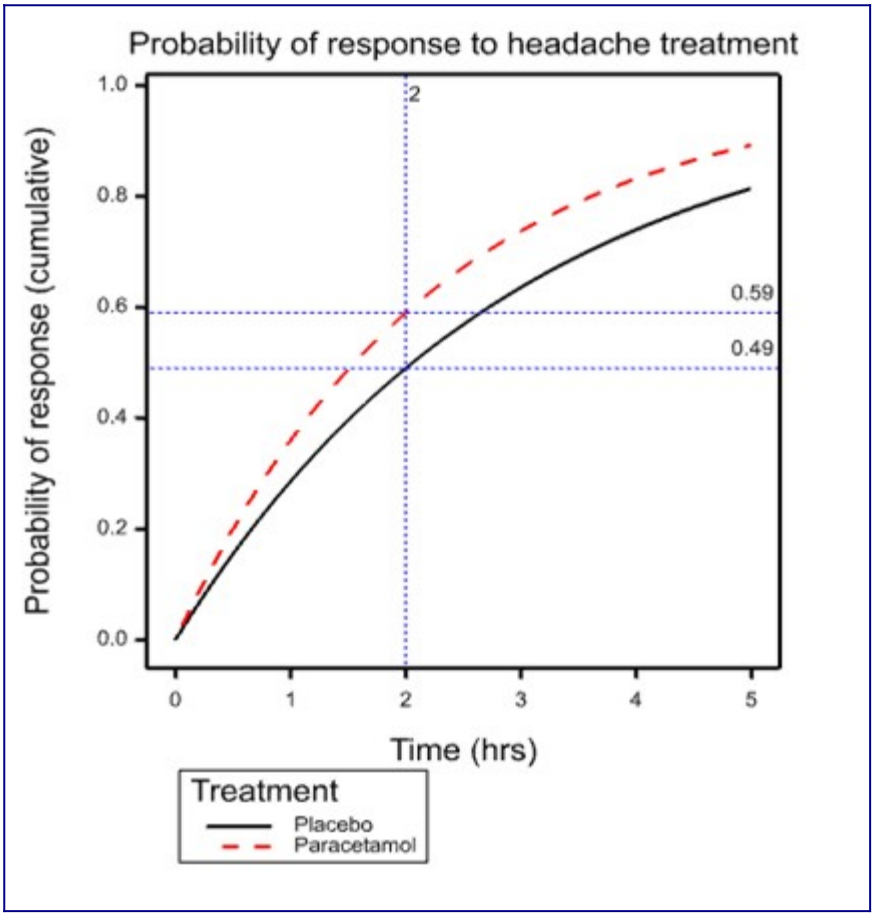
Individual response
discourse.datamethods.org
An intriguing post on Twitter by Judea Pearl linked to a blog written jointly with Scott Mueller Personalized Decision Making . The bog stated “The purpose of this note is to clarify the distinction …
Shared by Stephen John Senn

Inaccurate References Guile Duplicity and Perfidious Publications | Jay Brophy's website
brophyj.com
Good scientific writing is characterized by clarity of thought and often buttressed by appropriate peer reviewed references. This is of extreme importance when the scientific writing is that of conse…
Shared by jay brophy



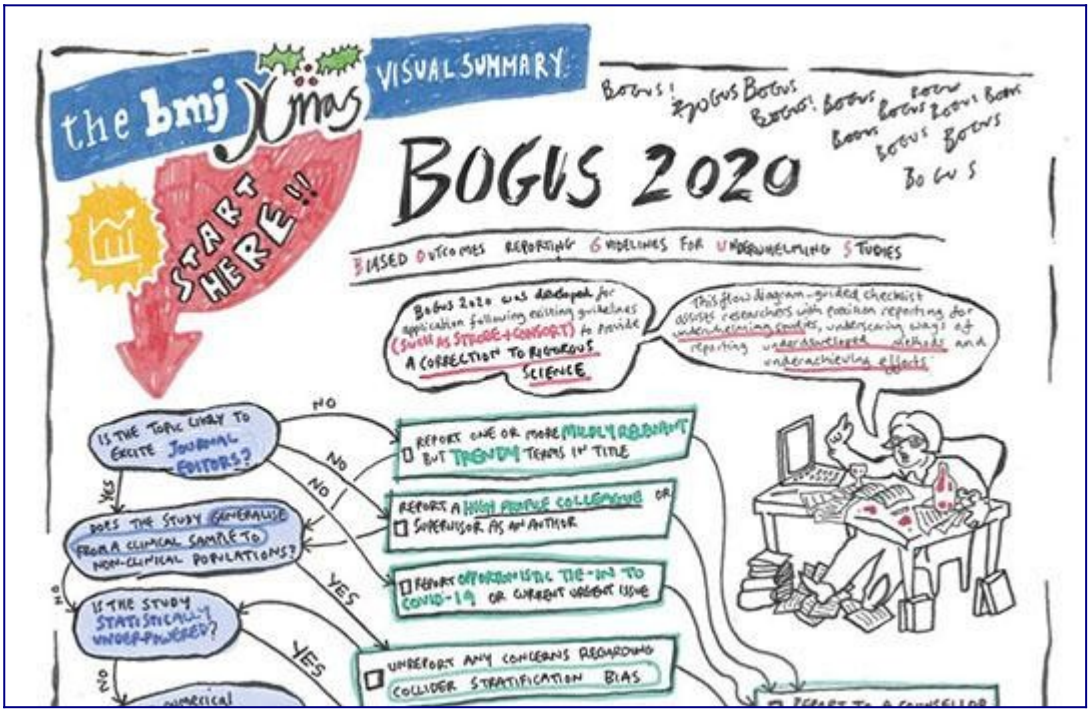
Significance magazine - Our most read articles of 2021
significancemagazine.com
Significance is an official magazine and website of the Royal Statistical Society (RSS) and the American Statistical Association (ASA).
Shared by Stats & Data Science Views

S. Senn: “Painful dichotomies” (Guest Post)
errorstatistics.com
Stephen Senn Head of Competence Center for Methodology and Statistics (CCMS) Luxembourg Institute of Health Twitter @stephensenn Painful dichotomies The tweet read “Featured review: Only 10% people w…
Shared by Stephen John Senn



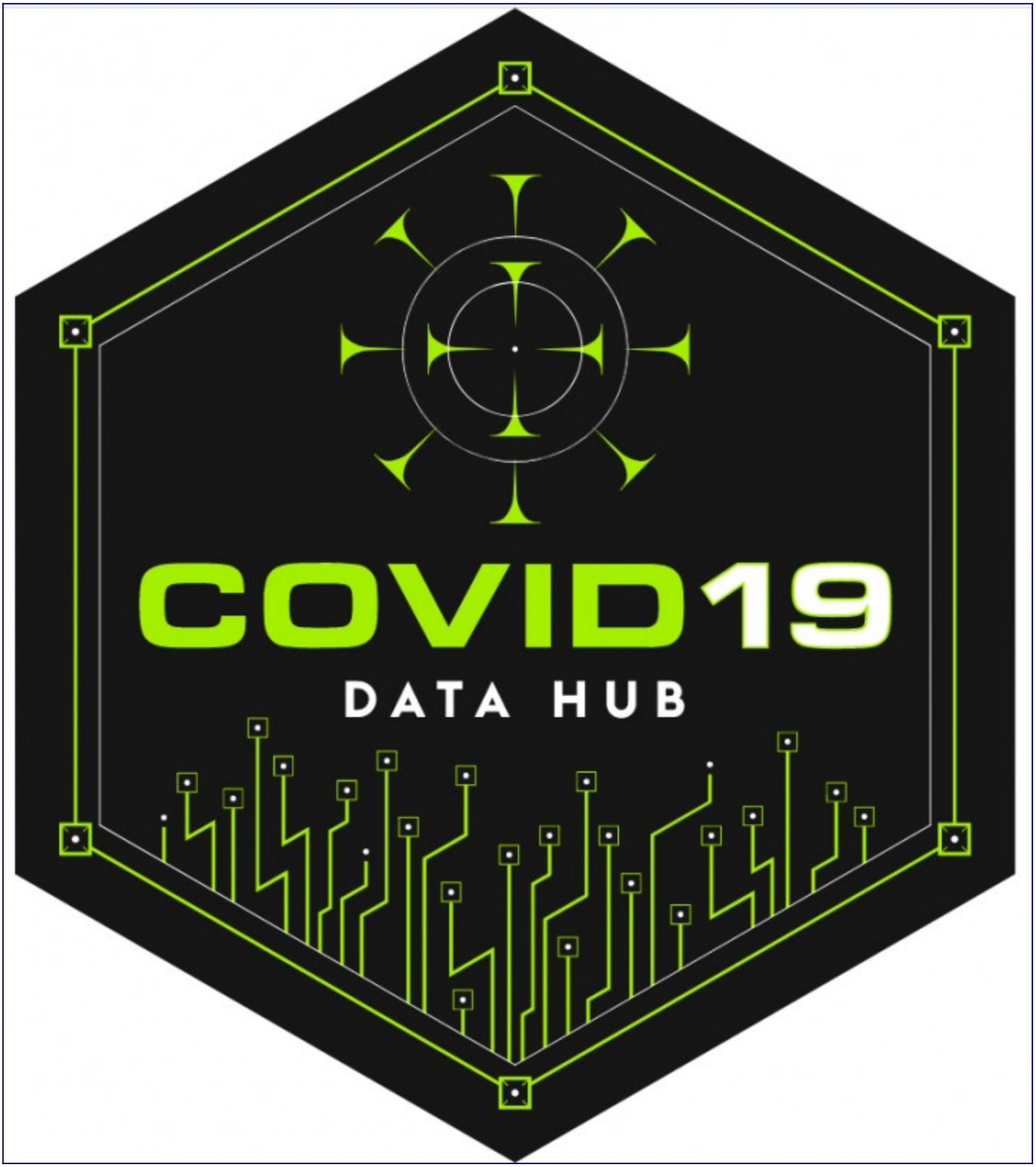
[Eyewitness Accounts Cast Doubt on Mammography Guidelines]
[medscape.com]
[The Canadian National Breast Screening Study trials, which are an anomaly among major mammography trials, have had a significant impact on guidelines for women aged 40 to 49.]
Shared by [Retraction Watch]

Biased Outcome reporting Guidelines for Underwhelming Studies (BOGUS) statement and checklist
bmj.com
Abstract Most biomedical research is flawed. While existing reporting guidelines (eg, STROBE, CONSORT, STARD) highlight the strengths of already strong studies, few resources exist to highlight t…
Shared by Maarten van Smeden



Avoiding One-Number Summaries of Treatment Effects for RCTs with Binary Outcomes | Statistical Thinking
fharrell.com
This article presents an argument that for RCTs with a binary outcome the primary result should be a distribution and not any single number summary. The GUSTO-I study is used to exemplify risk differ…
Shared by Frank Harrell
Problems with NNT
discourse.datamethods.org
In talking to a patient or in making grand policy decisions, when the likely benefit of a treatment is being considered, it is wise to utilize both relative (e.g., hazard ratio) and absolute treatmen…
Shared by Frank Harrell

Recommended Statistics Articles
hbiostat.org
These articles, largely from the last 60 days, are recommended because of their potential applicability. A small minority are recommended because they demonstrate how not to do research.
Shared by Frank Harrell
AHA “regrets any confusion” and is reviewing meeting policies after outcry over Covid-19 vaccine abstract
retractionwatch.com
Days after a leading heart journal issued an expression of concern for a meeting abstract suggesting that vaccines against Covid-19 may cause cardiac damage, its publisher, the American Heart Associa…
The AHA doesn't look good here
Shared by Retraction Watch
General COVID-19 Therapeutics Trial Design
hbiostat.org
The design, execution, and analysis will differ according to whether a classical frequentist1 or Bayesian approach is chosen. Bayesian approaches are generally more flexible and are especially effici…
Shared by Frank Harrell
The COVID19 package, an interface to the COVID-19 Data Hub
rviews.rstudio.com
An R community blog edited by RStudio
Shared by RStudio



Causal Analysis in Theory and Practice
causality.cs.ucla.edu
The purpose of this note is to clarify the distinction between personalized and population-based decision making, the former concerns the behavior of a specific individual while the latter concerns a…

Shared by Judea Pearl

Difference-in-Difference in the Time of Cholera: a Gentle Introduction for Epidemiologists
ncbi.nlm.nih.gov
The goal of this article is to provide an introduction to the intuition behind the difference-in-difference method for epidemiologists. We focus on the theoretical aspects of this tool, including the…

Shared by Dr Ellie Murray, ScD

UF researchers felt pressure to destroy COVID-19 data, not criticize DeSantis, report says
miamiherald.com
UF faculty report says researchers working on COVID-19 felt pressure by state of Florida to destroy COVID data and not criticize Gov. Ron DeSantis, raising academic freedom issues.

Shared by Boback Ziaeian

Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement
acpjournals.org
The TRIPOD Initiative developed a set of recommendations for reporting studies developing, validating, or updating a prediction model. Authors present a checklist of items deemed essential for tran...

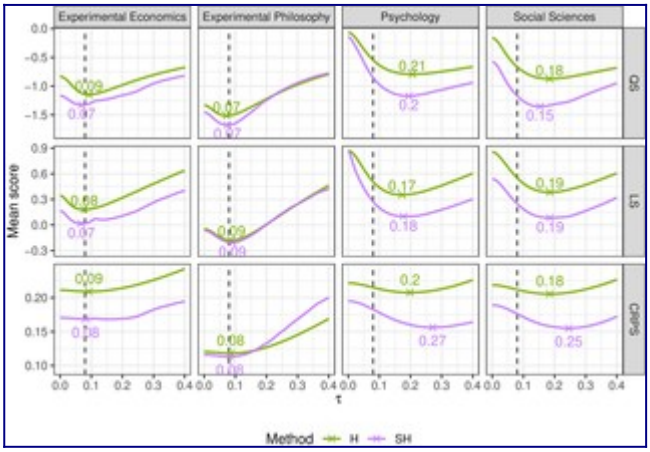
Shared by Gary Collins

A simple, step-by-step guide to interpreting decision curve analysis
diagnprognres.biomedcentral.com
Decision curve analysis is a method to evaluate prediction models and diagnostic tests that was introduced in a 2006 publication. Decision curves are now commonly reported in the literature, but ther…

Shared by Gary Collins

Journals submit article with fake data, demonstrate that some higher education journal you'd never heard of is as bad as PNAS, Psychological Science,…
statmodeling.stat.columbia.edu
I've just had my attention drawn to an article, "Donor money and the academy: Perceptions of undue donor pressure in political science, economics, and philosophy" by Sage Owens and Kal Avers-Lynde II…
Shared by Andrew Gelman et al.

Probabilistic forecasting of replication studies
journals.plos.org
Throughout the last decade, the so-called replication crisis has stimulated many researchers to conduct large-scale replication projects. With data from four of these projects, we computed probabilis…
Shared by Leonhard Held



Misinformation fuelled by 'tsunami' of poor research, says science prize winner
theguardian.com
Dutch microbiologist Elisabeth Bik, winner of prestigious John Maddox prize, says trust in science is being undermined
Shared by Retraction Watch

[When confidence intervals include unreasonable values . . . When confidence intervals include only unreasonable values . . .]
statmodeling.stat.columbia.edu
[Economists' love affair with randomized controlled trials (RCTs) is growing stronger by the day. But what should we make of an RCT that produces a point estimate and confidence interval that largely …]
Shared by Andrew Gelman et al.

[How many lives have Covid-19 vaccines saved?]
qz.com
[Researchers wanted to know how many lives vaccines may have saved due to the states' Covid-19 vaccination campaigns in the US.]
Shared by Virtual Mountain | Fully Vaccinated

[No retractions necessary: An interview with the editor of the Journal of Universal Rejection]
retractionwatch.com
[Earlier today, we were introduced to the Journal of Universal Rejection by our friend Duncan Moore. From its homepage: The founding principle of the Journal of Universal Rejection (JofUR) is rejectio…]
Shared by Retraction Watch

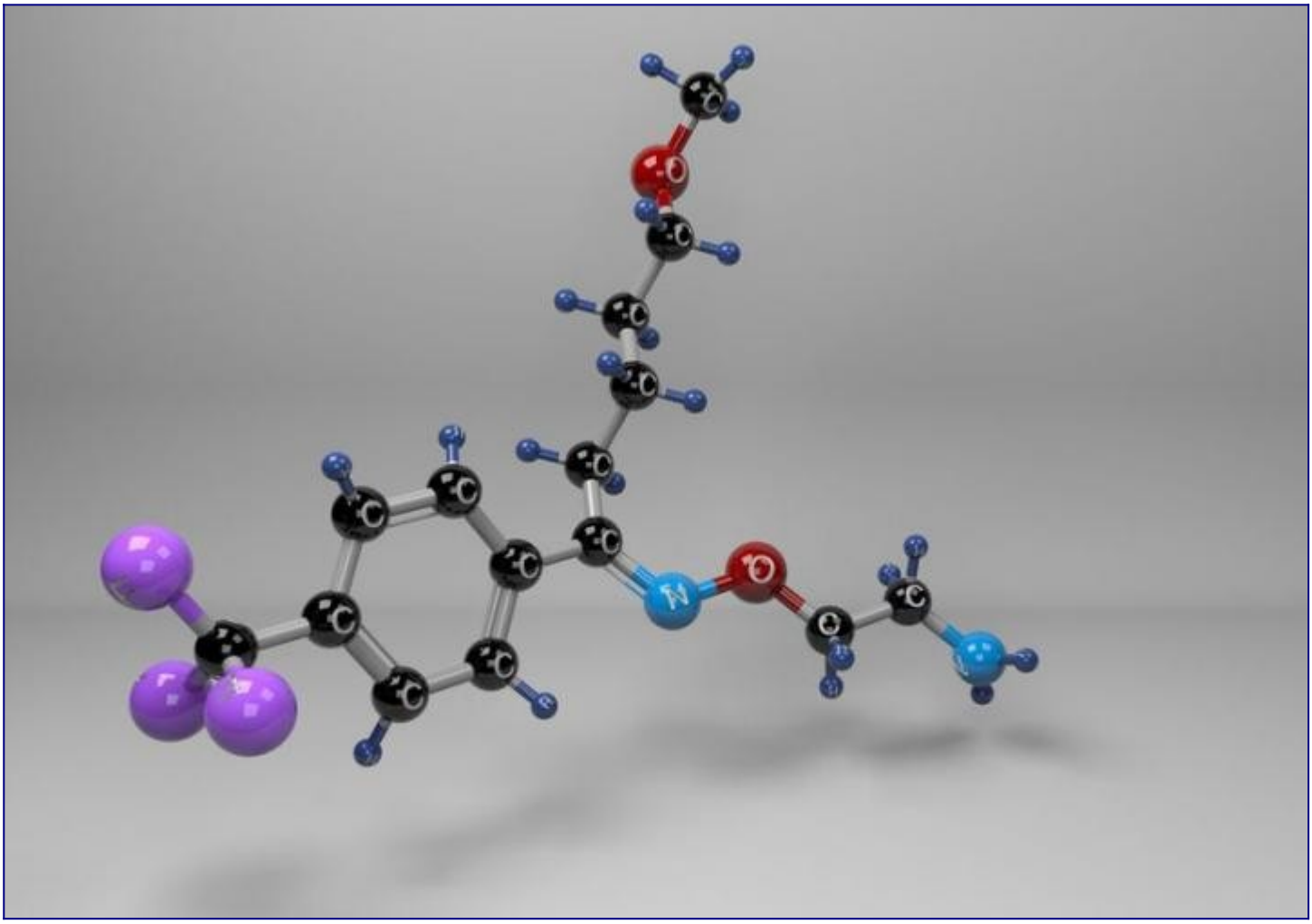
Keeping science reproducible in a world of custom code and data
arstechnica.com
Computer analysis is now a central part of science.
Shared by Retraction Watch

Yale researchers unlock secrets through study of long-term effects of COVID-19
yaledailynews.com
Yale researchers expand their work on long COVID in conjunction with Survivor Corps to study how vaccinations are affecting long COVID symptoms.
Shared by Harlan Krumholz

Retractions in medicine: the tip of the iceberg
academic-oup-com.eres.qnl.qa
In 1983, in the aftermath of what was then considered one of the most significant cases of scientific fraud ever, The New York Times reported that 82 papers by
Shared by Retraction Watch
A Closer Look at the New Fluvoxamine Trial Data
the-scientist.com

Authors of a newly published study on the use of an antidepressant for COVID-19 claim the drug greatly reduces hospitalizations and mortality. But some experts question whether that's really what the…

Shared by Maia Drazhar



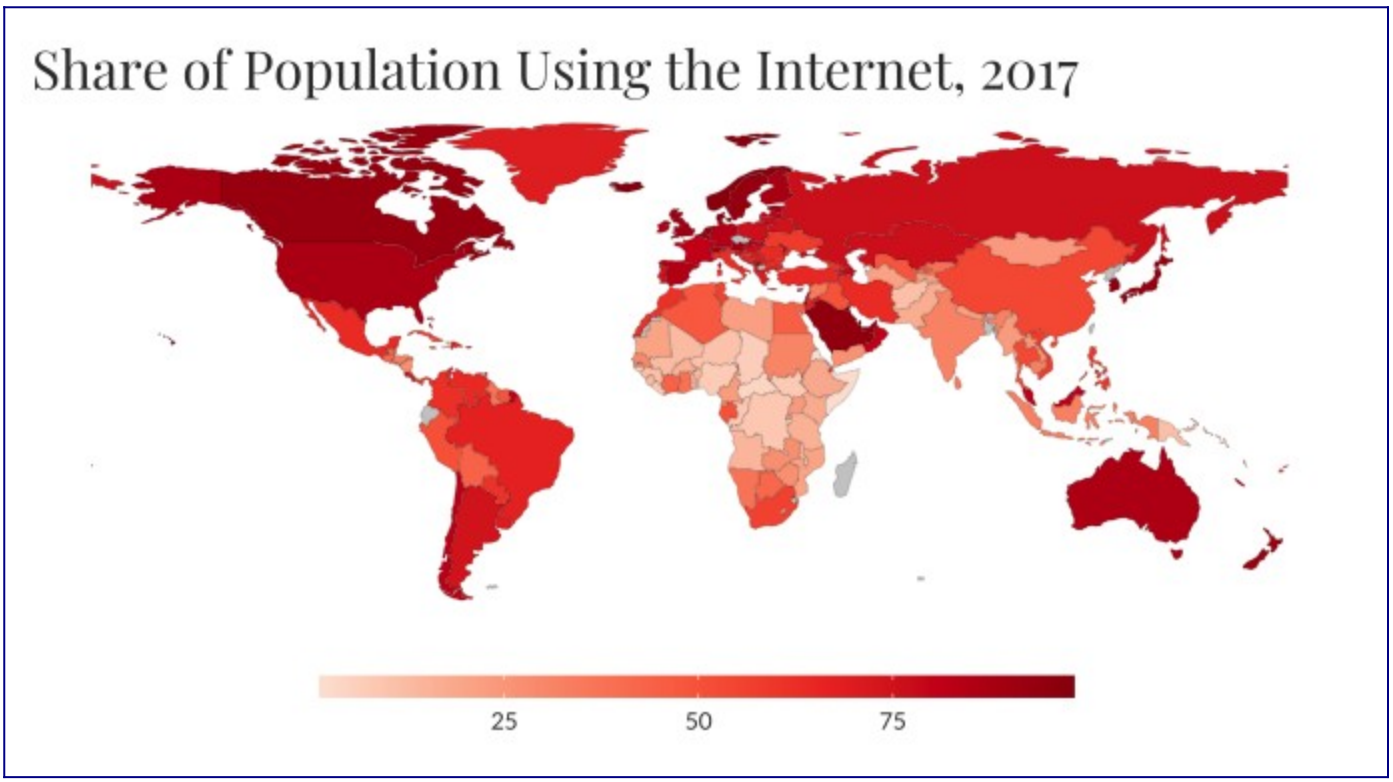
October 2021: "Top 40" New CRAN Packages

rviews.rstudio.com

An R community blog edited by RStudio

See especially scatterPlotMatrix

Shared by RStudio



[Randomized Clinical Trials Do Not Mimic Clinical Practice, Thank Goodness | Statistical Thinking](fharrell.com)
fharrell.com
[What clinicians learn from clinical practice, unless they routinely do n-of-one studies, is based on comparisons of unlikes. Then they criticize like-vs-like comparisons from randomized trials for no…]
Shared by [Stephen John Senn]
[With Covid studies, the quality of the evidence matters]
theguardian.com
[Putting many low-quality studies together cannot provide reliable answers about masks and ivermectin]
Shared by [David Spiegelhalter]

Predicting the COVID-19 Pandemic Impact on Clinical Trial Recruitment
tandfonline.com
Contemporary clinical trials often have complex logistics: they are run across multiple centers/countries and involve a lot of uncertainties about disease prevalence rates, patient characteristics ...
Shared by UW Clinical Research Office

Statistics in Biopharmaceutical Research

A Publication of the American Statistical Association

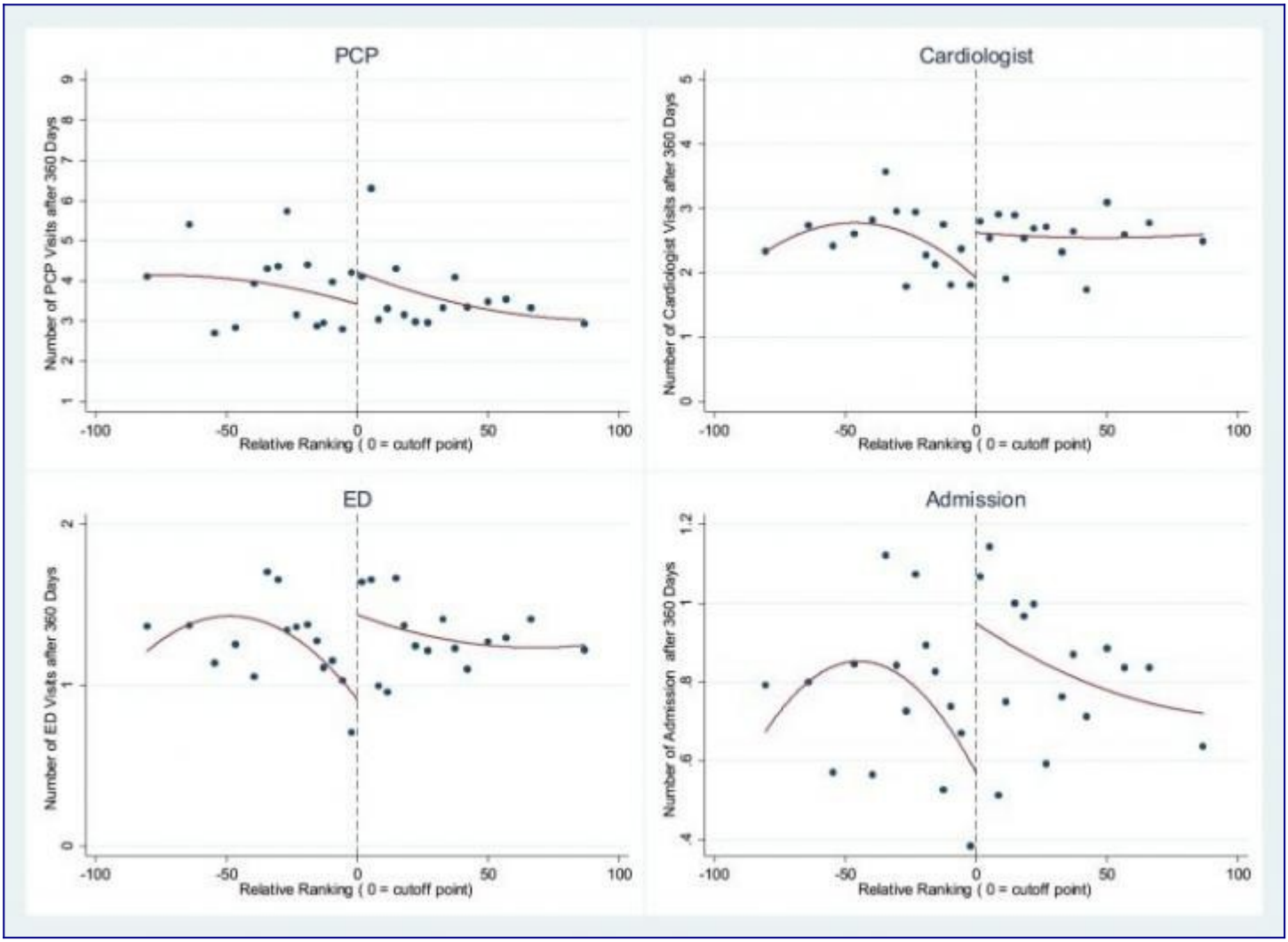
Multiple imputation and bootstrapping for internal validation
discourse.datamethods.org
What is the optimal way to combine multiple imputation and bootstrap to quantify uncertainty in discrimination and calibration metrics for the internal validation of diagnostic or prognostic models? …

3.2.1 Boot MI percentile

Both Schomaker and Heumann[9] and Brand et al.[11] recommended calculating bootstrap percentile intervals to the estimator $\widehat{\theta}_M$. This consists of

1. B bootstrap samples of the observed data are taken $Z_{\text{obs}}^{(b)}$, $b = 1, \ldots, B$.
2. For each $b = 1, \ldots B$, use MI to impute missing data in $Z_{\text{obs}}^{(b)}$ M times and estimate θ in each imputed dataset, giving $\widehat{\theta}_{b,m}$.
3. For point estimation of θ, one can either use $\bar{\theta}_M$ or $\bar{\theta}_{BM} = B^{-1}\sum_{b=1}^{B} \bar{\theta}_b$, where $\bar{\theta}_b = M^{-1}\sum_{m=1}^{M} \widehat{\theta}_{b,m}$.
4. A $(1 - 2\alpha)\%$ percentile confidence interval for θ is then formed by taking the α and $1 - \alpha$ empirical percentiles of the $\bar{\theta}_b$, $b = 1, \ldots, B$ values.

Just another day at the sausage factory . . . It’s just funny how regression discontinuity analyses routinely produce these ridiculous graphs and the…
statmodeling.stat.columbia.edu
Ethan Steinberg sends in these delightful regression discontinuity graphs: He reports that the graphs come from a study from the University of Pennsylvania trying to estimate the effect of a care man…
Shared by Andrew Gelman et al.



Causal Analysis in Theory and Practice
causality.cs.ucla.edu
Many readers have asked for my reaction to Guido Imbens's recent paper, titled, "Potential Outcome and Directed Acyclic Graph Approaches to Causality: Relevance for Empirical Practice in Economics," …

Shared by Judea Pearl

Bayesian Interpretation of the EXCEL Trial and Other Randomized Clinical Trials of Left Main Coronary Artery Revascularization
jamanetwork.com
Question A randomized clinical trial (EXCEL; Evaluation of XIENCE Versus Coronary Artery Bypass Surgery for Effectiveness of Left Main Revascularization) of patients with left main coronary artery di…

Shared by jay brophy

A billion-dollar donation: estimating the cost of researchers' time spent on peer review
researchintegrityjournal.biomedcentral.com
The amount and value of researchers' peer review work is critical for academia and journal publishing. However, this labor is under-recognized, its magnitude is unknown, and alternative ways of organ…

Shared by Gary Collins

Edward Tufte Wants You to See Better
npr.org
Data scientist Edward Tufte (dubbed the "Galileo of graphics" by BusinessWeek) pioneered the field of data visualization. Tufte discusses what he calls "forever knowledge," and his latest projects: s…
Shared by Edward Tufte

From grassroots to global: A blueprint for building a reproducibility network
journals.plos.org

Researchers, institutions, funders, and publishers are considering how to improve research culture and quality, but no single part of the research ecosystem can effect change on its own. The UK Repro…
Shared by Darren Dahly, PhD FML WTF JFC STFU LOL

A brief history of medical statistics and its impact on reproducibility.
statsepi.substack.com
You have to admit that medical research is a strange bird. There are few industries where we take a group of highly-trained, specialized practitioners that already shoulder a great deal of responsibi…
Shared by Darren Dahly, PhD FML WTF JFC STFU LOL

Randomized good, statistical bad
discourse.datamethods.org
The 2021 Friends of Cancer Research Annual Meeting Day 2 featured a panel discussion somehow connected with FDA's #ProjectOptimus, in which the supposed importance of randomized dose-comparison studi…

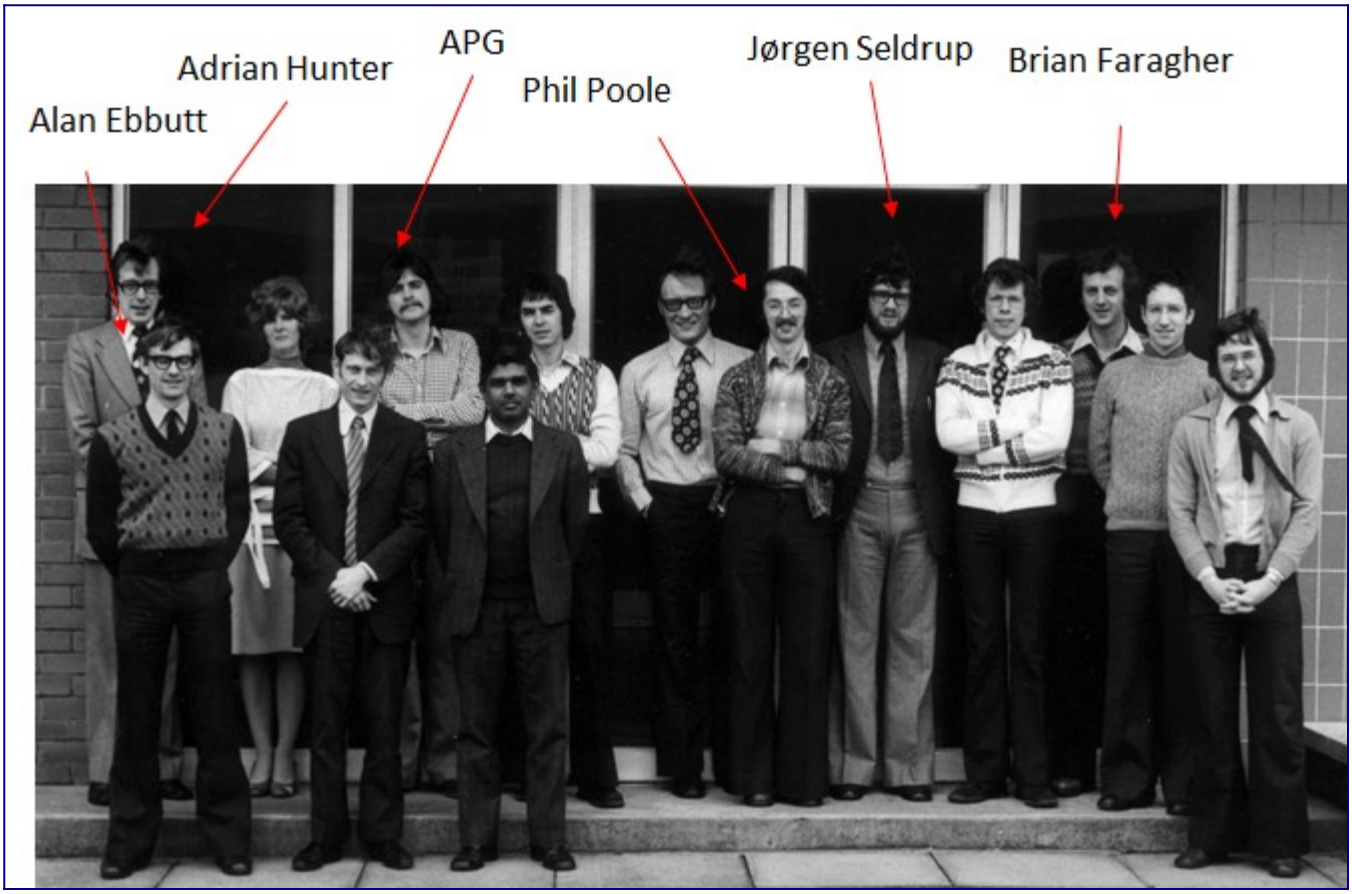
Bad MATH+? Covid treatment paper by Pierre Kory retracted for flawed results
retractionwatch.com
Pierre Kory A Wisconsin physician who has been pushing unproven treatments for Covid-19 has lost a paper on a hospital protocol his group says radically reduced deaths from the infection after one of…
Shared by Retraction Watch

Dexamethasone 12 mg versus 6 mg for patients with COVID-19 and severe hypoxaemia: a pre-planned, secondary Bayesian analysis of the COVID STEROID 2 t…
link.springer.com
We compared dexamethasone 12 versus 6 mg daily for up to 10 days in patients with coronavirus disease 2019 (COVID-19) and severe hypoxaemia in th
Shared by xxxichiroxxx_ID-pharmacist

Are some feasibility studies more feasible than others? A review of the outcomes of feasibility studies on the ISRCTN registry
pilotfeasibilitystudies.biomedcentral.com
Feasibility studies are often conducted before committing to a randomised controlled trial (RCT), yet there is little published evidence to inform how useful feasibility studies are, especially in te…
Shared by Andrew Althouse

America Needs a New Scientific Revolution
theatlantic.com
A repurposed antidepressant might help treat COVID-19, a remarkable study found. The way this research was funded highlights a big problem—and bigger opportunity—in American science.
Shared by Harlan Krumholz
The 60-Year-Old Scientific Screwup That Helped Covid Kill
wired.com
All pandemic long, scientists brawled over how the virus spreads. Droplets! No, aerosols! At the heart of the fight was a teensy error with huge consequences.
Shared by Frank Harrell

[Tales from the Interview Room – Over 45 years Ago.](#)
[bayesicsolutions.wordpress.com](#)
[I'm sure we all believe we are logical, rational beings and I was thinking that while I was dipping into Steven Pinker's new book "Rationality". Pinker appeared on BBC Radio 4's "Start the Week" last…](#)
Shared by [Andy Grieve](#)

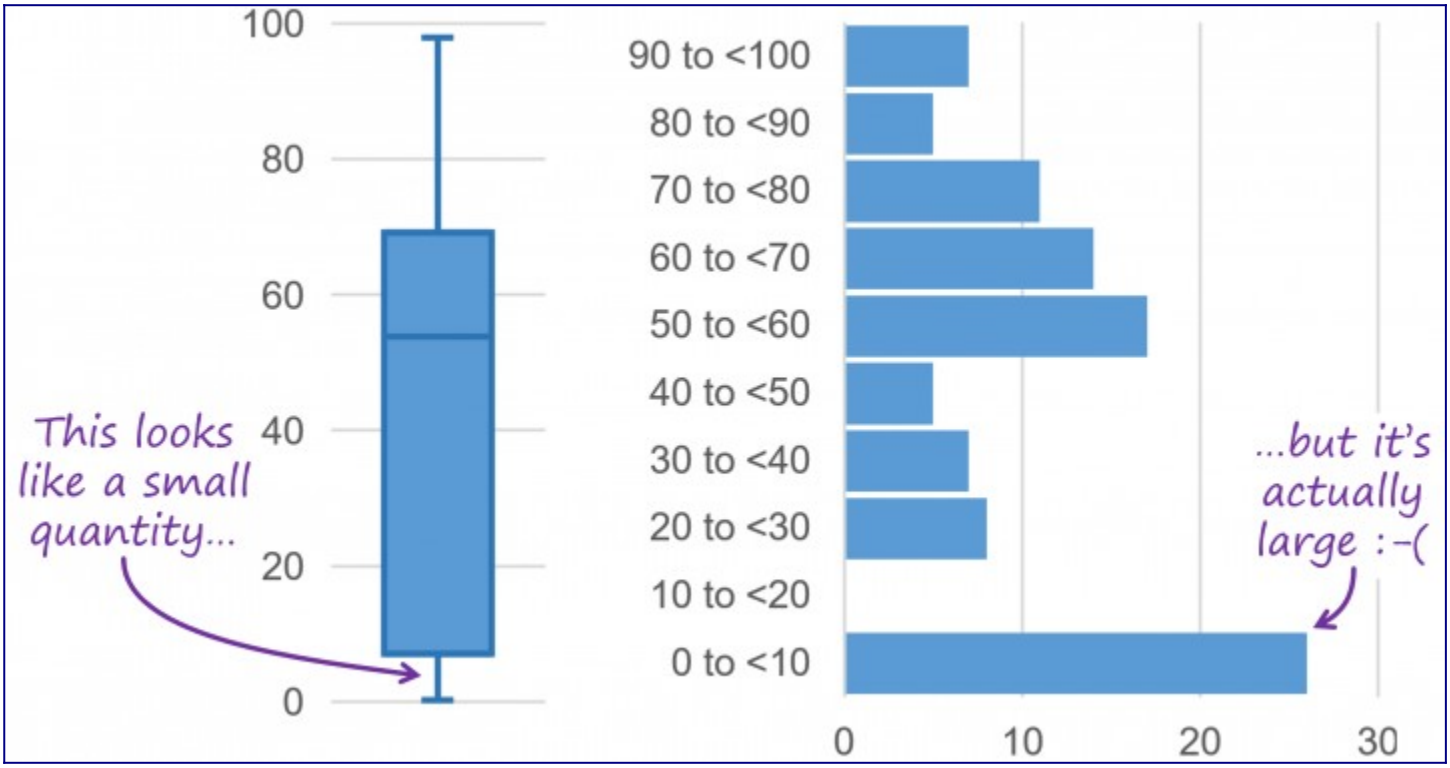


State of the art in selection of variables and functional forms in multivariable analysis—outstanding issues
diagnprognres.biomedcentral.com
How to select variables and identify functional forms for continuous variables is a key concern when creating a multivariable model. Ad hoc ‘traditional’ approaches to variable selection have been in…
Shared by Ben Van Calster

I’ve Stopped Using Box Plots. Should You?
nightingaledvs.com

tl;dr: After having explained how to read box plots to thousands of workshop participants, I now believe that they’re poorly conceived (box plots, not...

Shared by Arthur Charpentier



Why scientific journal authorship practices make no sense et al.

science.org

Some labs give out authorship like Halloween candy, our Experimental Error columnist writes

Shared by Retraction Watch

S. Senn: Guest Post

errorstatistics.com

Stephen Senn Consultant Statistician Edinburgh, Scotland The Many Halls Problem It’s not that paradox but another Generalisation is passing…from the consideration of a restricted set to that of a mor…

Shared by Stephen John Senn

Analysis of variance (adjusted for covariate)

Variate: Final
Covariate: Initial

Source of variation	d.f.	s.s.	m.s.	v.r.	cov.ef.	F pr.
Hall stratum						
Hall diet	1	26.155	26.155	0.33	0.68	0.584
Covariate	1	202.290	202.290	2.55		0.154
Residual	7	554.655	79.236	17.62	1.19	
Hall.Student stratum						
Covariate	1	480.244	480.244	106.79		<.001
Residual	29	130.410	4.497		4.53	
Total	39	1379.937				

Covariate regressions

Variate: Final

Covariate	coefficient	s.e.
Hall stratum		
Initial	0.42	0.265
Hall.Student stratum		
Initial	0.957	0.0926
Combined estimates		
Initial	0.912	0.0888

¶

Polygenic Risk, Fitness, and Obesity in the Coronary Artery Risk Development in Young Adults (CARDIA) Study
ncbi.nlm.nih.gov
What is the added value of polygenic risk in predicting body mass index (BMI) over time beyond young adulthood BMI, parental history of overweight, fitness, and activity?Among 1608 white individuals …
Shared by Venk Murthy

Covid-19: Researcher blows the whistle on data integrity issues in Pfizer’s vaccine trial
bmj.com
Revelations of poor practices at a contract research company helping to carry out Pfizer’s pivotal covid-19 vaccine trial raise questions about data integrity and regulatory oversight. Paul D Thacker…
Shared by Retraction Watch

Risk of bias in studies on prediction models developed using supervised machine learning techniques: systematic review
bmj.com
Objective To assess the methodological quality of studies on prediction models developed using machine learning techniques across all medical specialties. Design Systematic review. Data sources PubMe…
Shared by Gary Collins

Transformations

hbiostat.org

Many authors and pharmaceutical clinical trialists make the mistake of analyzing change from baseline instead of making the raw follow-up measurements the primary outcomes, covariate-adjusted for bas…

Shared by Frank Harrell

[Sep 29] TI Methods Speaker Series: What is your question? Causal and predictive purposes of clinical trials and the implications for analysis

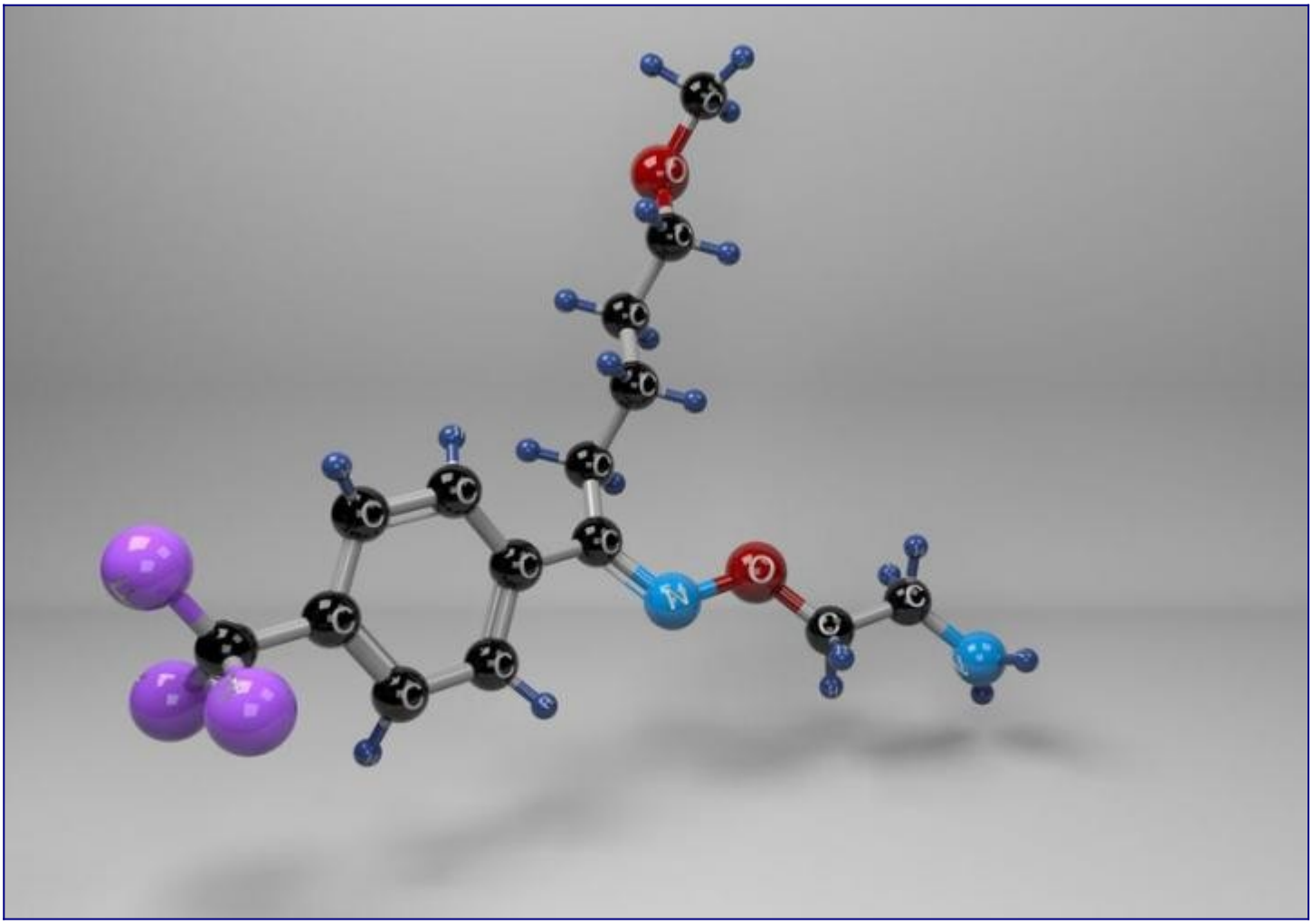
ti.ubc.ca

The TI Methods Speaker Series are offered free of charge and everyone is welcome. The event is held at noon on the last Wednesday of each month and used to take place in Rudy North Lecture Theatre (C…

Shared by Stephen John Senn

A Closer Look at the New Fluvoxamine Trial Data

the-scientist.com
Authors of a newly published study on the use of an antidepressant for COVID-19 claim the drug greatly reduces hospitalizations and mortality. But some experts question whether that's really what the…
Shared by Zyite.news



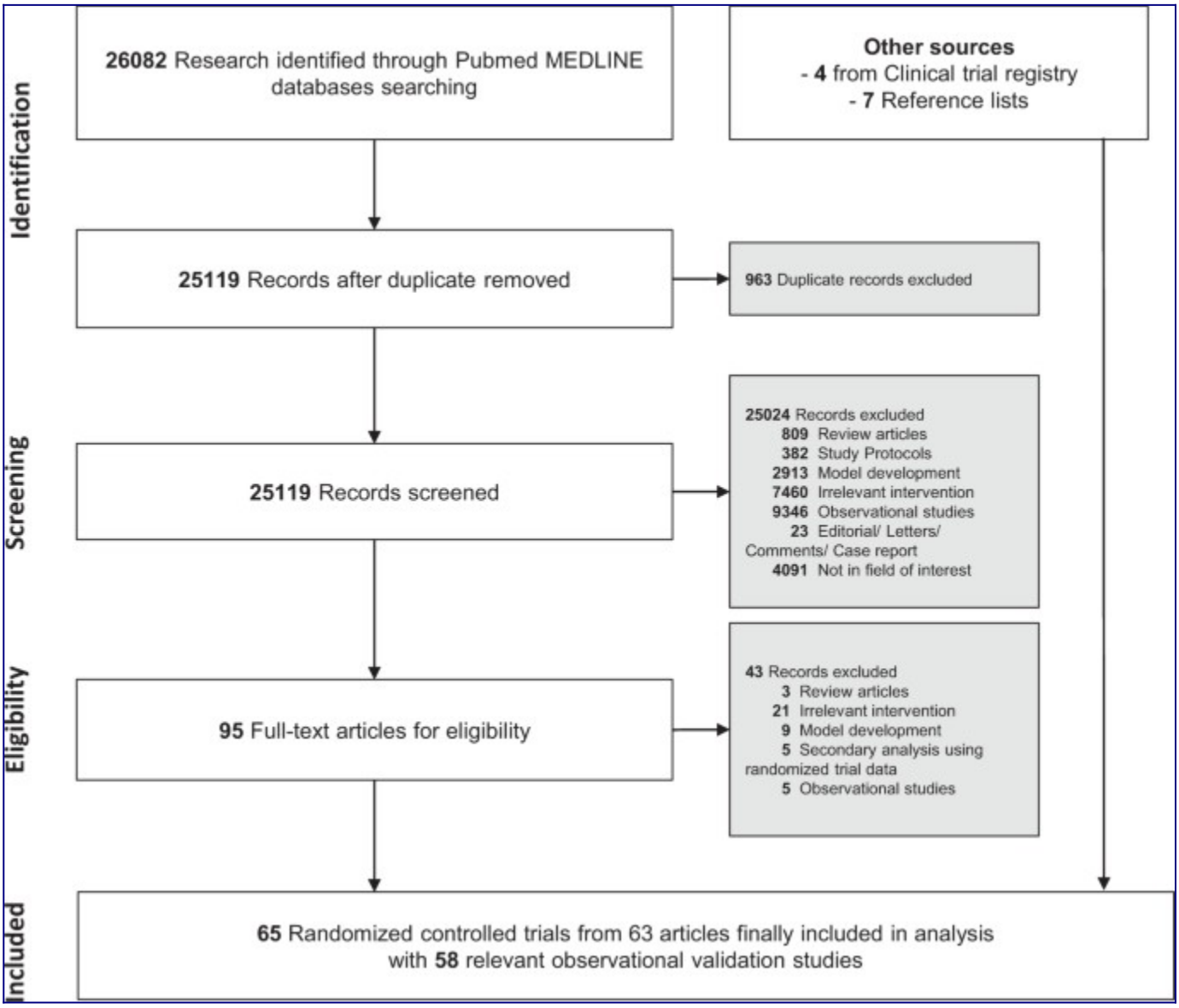
Ivermectin-COVID-19 study retracted; authors blame file mixup
retractionwatch.com
The authors of a study purportedly showing that ivermectin could treat patients with SARS-CoV-2 have retracted their paper after acknowledging that their data were garbled. The paper, “Effects of a S…
Shared by Retraction Watch
In Machine Learning Predictions for Health Care the Confusion Matrix is a Matrix of Confusion | Statistical Thinking
fharrell.com
The performance metrics chosen for prediction tools, and for Machine Learning in particular, have significant implications for health care and a penetrating understanding of the AUROC will lead to be…
Shared by Gary Collins

Guidance for biostatisticians on their essential contributions to clinical and translational research protocol review
cambridge.org
Guidance for biostatisticians on their essential contributions to clinical and translational research protocol review - Volume 5 Issue 1
Shared by BioRankings

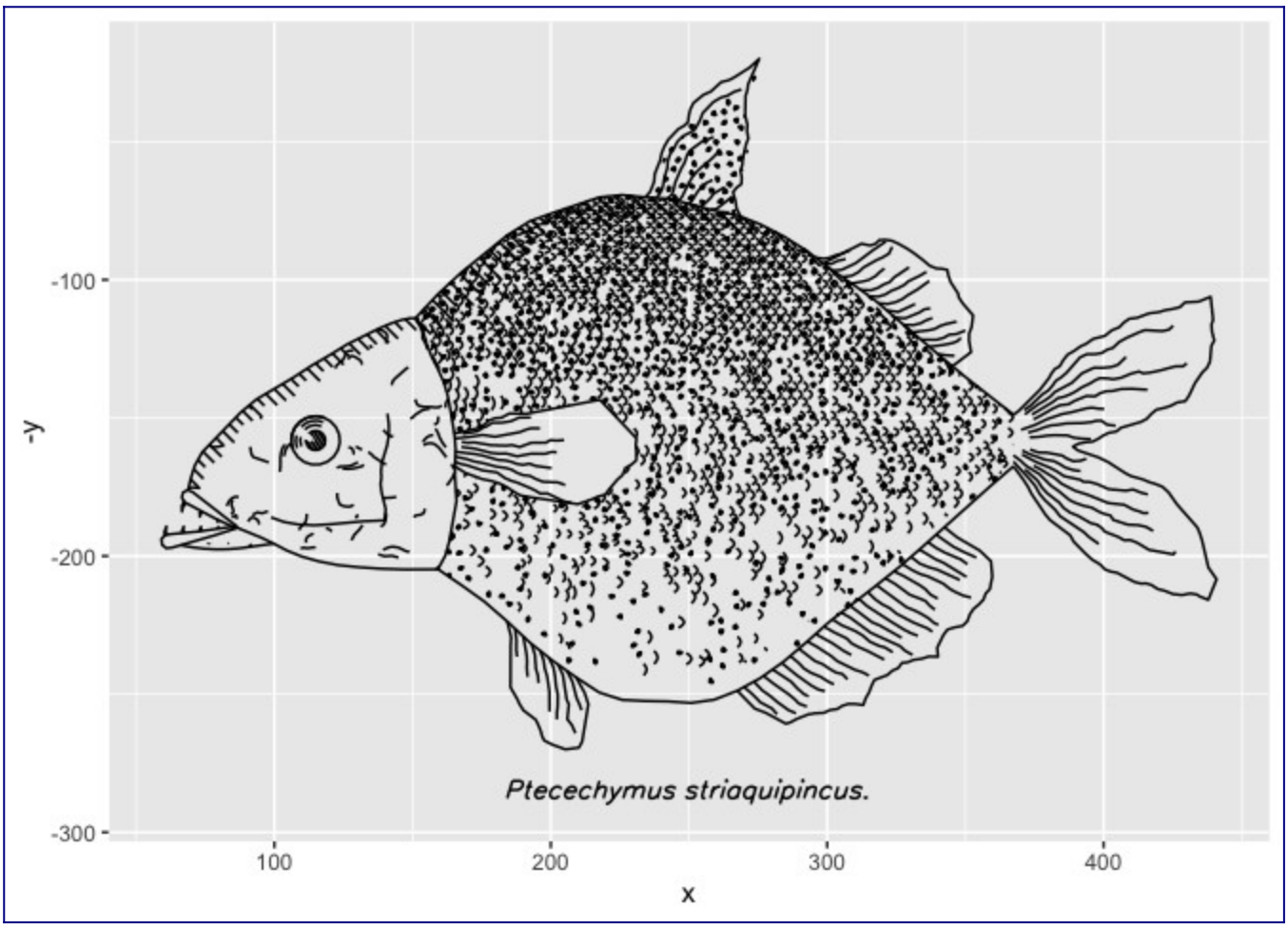
Journal of Clinical and Translational Science

the Official Journal of ACTS and CR Forum

Clinical impact and quality of randomized controlled trials involving interventions evaluating artificial intelligence prediction tools: a systematic…
nature.com
The evidence of the impact of traditional statistical (TS) and artificial intelligence (AI) tool interventions in clinical practice was limited. This study aimed to investigate the clinical impact an…
Shared by Muin J. Khoury



September 2021: "Top 40" New CRAN Packages
rviews.rstudio.com
An R community blog edited by RStudio
Shared by RStudio



[How many lives have coronavirus vaccines saved? We used state data on deaths and vaccination rates to find out]
theconversation.com
[Using a robust statistical model, researchers estimate that coronavirus vaccines had prevented 140,000 deaths by May 9, 2021.]
Shared by [David Lindwall]

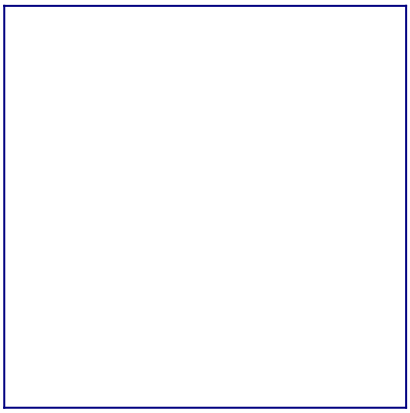
Covid vaccine myocarditis | Jay Brophy's Website
brophyj.com
On Sept 16 2021, 12 researchers from the Ottawa Heart Institute published a pre-print claiming that the COVID 19 mRNA vaccination myocarditis rate was 1 / 1000. On Sept 24 2021, the authors reported …
Shared by Retraction Watch

One of the Most Egregious Ripoffs in the History of Science
nautil.us
James Watson once said his road to the 1962 Nobel Prize began in Naples, Italy. At a conference in 1951, he met Maurice Wilkins, the…
Shared by Harlan Krumholz

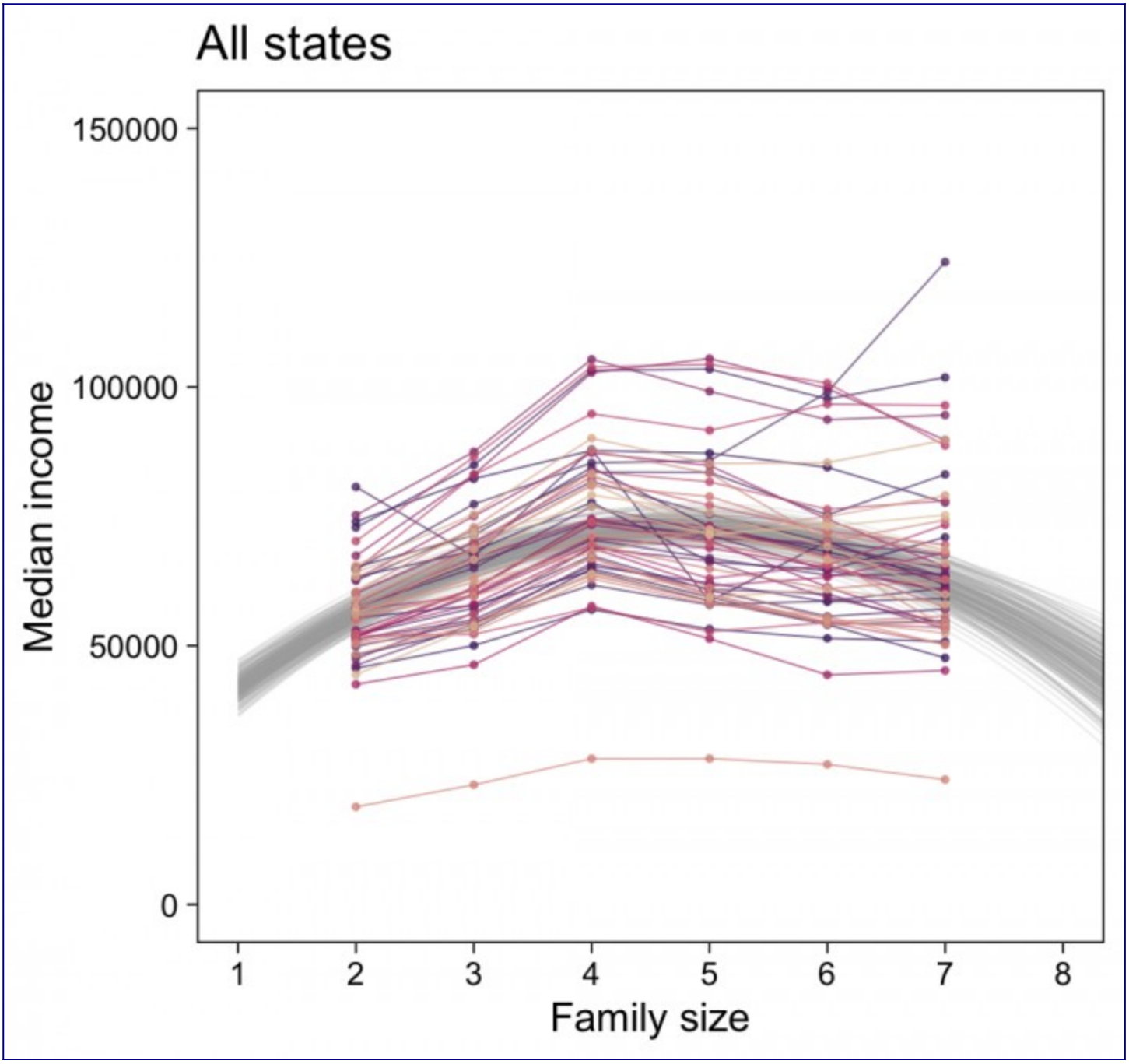
‘Follow the Science’ Doesn't Work When the Science Is This Bad
theatlantic.com
I spent months looking closely at the research on the use of ivermectin. It was not a pretty sight.
Shared by Frank Harrell

Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studi…
bmjopen.bmj.com
Introduction The Transparent Reporting of a multivariable prediction model of Individual Prognosis Or Diagnosis (TRIPOD) statement and the Prediction model Risk Of Bias ASsessment Tool (PROBAST) were…
Shared by TRIPODStatement

Release of CmdStan 2.28.1
blog.mc-stan.org
We are very happy to announce that the 2.28.1 release of CmdStan is now available on Github! As usual, the release of CmdStan 2.28.1 is accompanied with the new releases of Stan Math, core Stan and S…
Shared by Stan



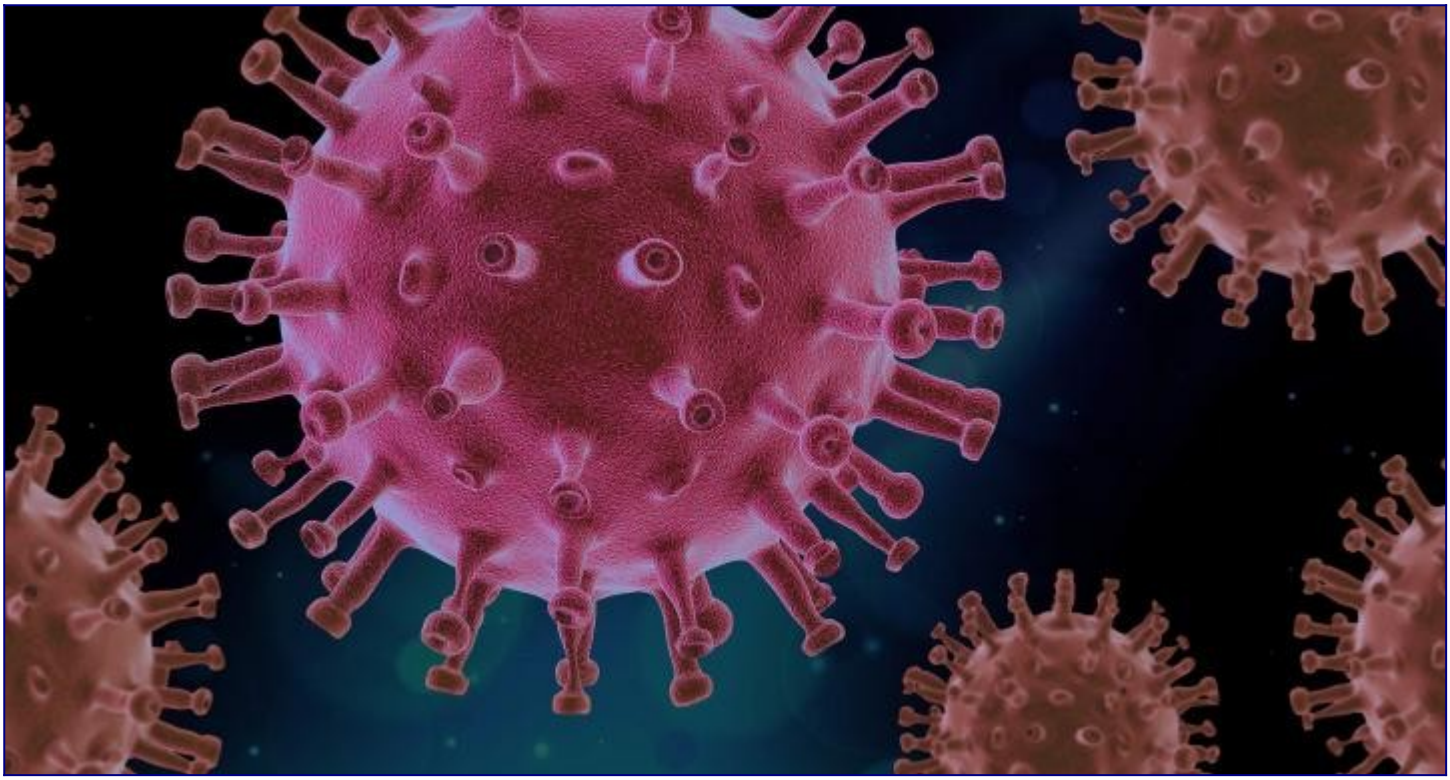
17 Metric Predicted Variable with One Metric Predictor | Doing Bayesian Data Analysis in brms and the tidyverse
bookdown.org
This project is an attempt to re-express the code in Kruschke's (2015) textbook. His models are re-fit in brms, plots are redone with ggplot2, and the general data wrangling code predominantly follow…
Shared by Darren Dahly, PhD



[Florida School Revises Covid Guidelines To Reflect Latest Misinformation](#)
theonion.com
[MIAMI—Promising to remain current on any recommendations that might gain traction on social media or receive a platform on right-wing news outlets, officials at Miami's Centner Academy stated Tuesday…](#)
Shared by [Boback Ziaeian](#)

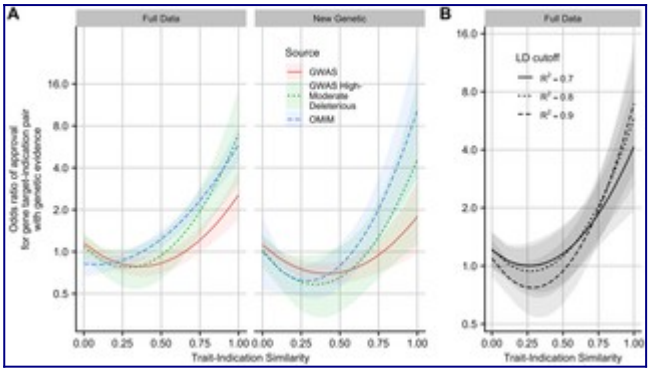
Statistical Errors in the Medical Literature | Statistical Thinking
fharrell.com
Misinterpretation of P-values and Main Study Results Dichotomania Problems With Change Scores Improper Subgrouping Serial Data and Response Trajectories Cluster Analysis As Doug Altman famously wrote…
Shared by Boback Ziaeian

Explained: The Various Sources of Uncertainty in COVID-19 Studies
science.thewire.in
Just as war makes every citizen into an amateur geographer and tactician, a pandemic makes epidemiologists of us all.
Shared by Linda Campbell



Misrepresentation and distortion of research in biomedical literature
pnas.org
Publication in peer-reviewed journals is an essential step in the scientific process. However, publication is not simply the reporting of facts arising from a straightforward analysis thereof. Author…
Shared by Maarten van Smeden

Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the p…
journals.plos.org
Author summary The growth of human genetics resources has the potential to help us develop better drugs. By looking at whether and how historical drug approvals could have been predicted from our cur…
Shared by Veera M. Rajagopal



Slowed canonical progress in large fields of science
pnas.org

The size of scientific fields may impede the rise of new ideas. Examining 1.8 billion citations among 90 million papers across 241 subjects, we find a deluge of papers does not lead to turnover of ce…

Shared by Retraction Watch

Interpreting the random-effect solution in a mixed model

discourse.datamethods.org

I am using a generalized mixed linear model to account for repeated measurement (of counts of things representing performance in multiple matches of different sports teams, but that is not particular…

Comments on a Nobel prize in economics for causal inference

statmodeling.stat.columbia.edu

A reporter contacted me to ask my thoughts on the recent Nobel prize in economics. I didn't know that this had happened so I googled *nobel prize economics* and found the heading, "David Card, Joshua…

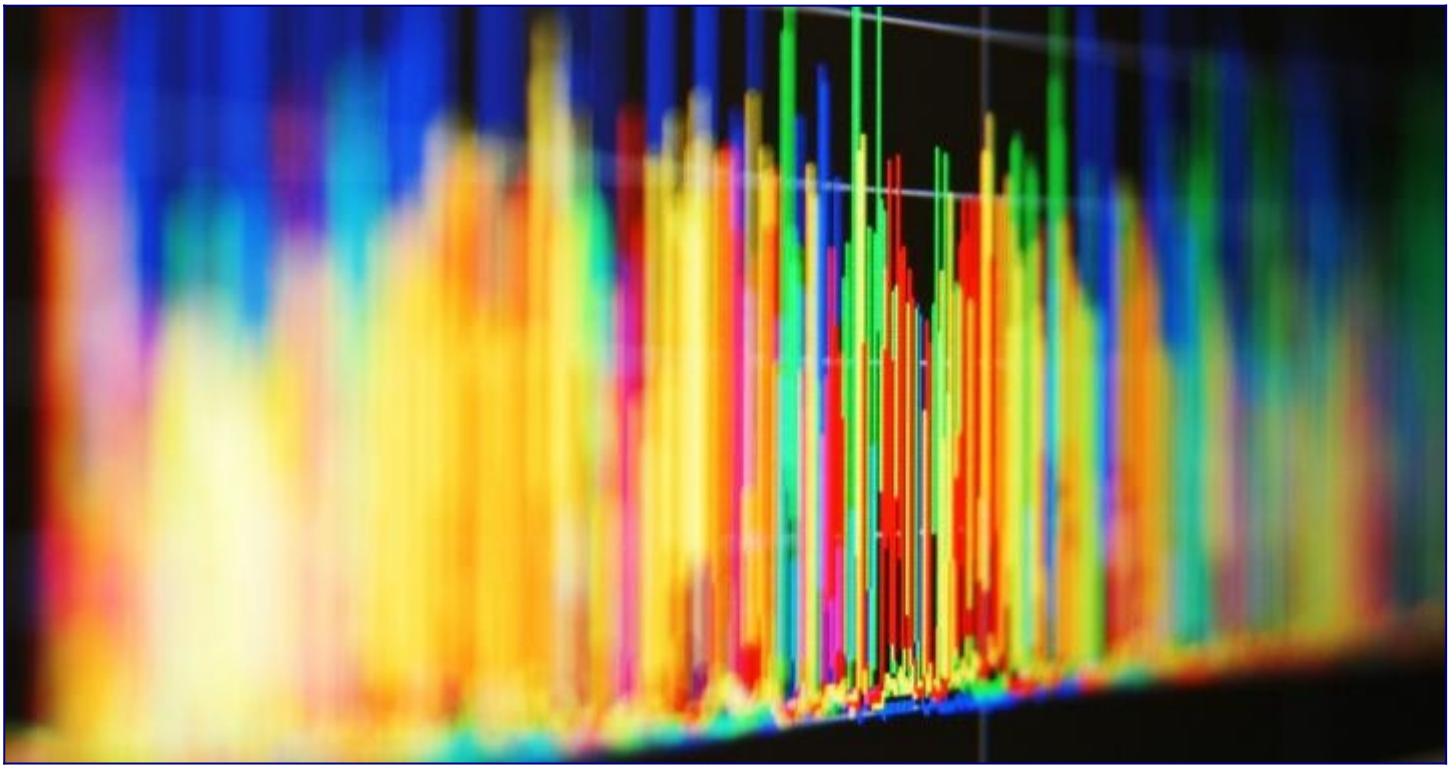
Shared by Andrew Gelman et al.

How media reporting of modelling has shaped our understanding of the pandemic

fullfact.org

Scientific modelling has been widely criticised throughout the pandemic. We spoke to some of the leading experts behind the numbers about how they felt their work had been reported .

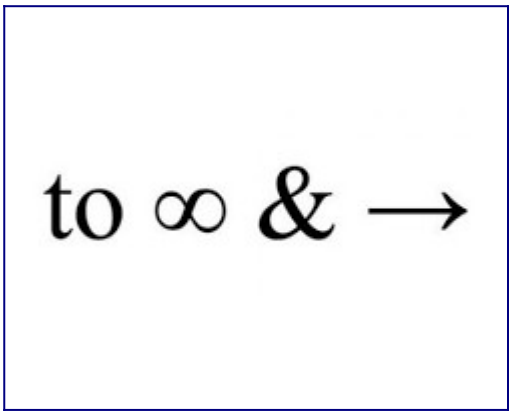
Shared by Scotsgait



Negative Controls: A Tool for Detecting Confounding and Bias in Observational Studies
ncbi.nlm.nih.gov
Non-causal associations between exposures and outcomes are a threat to validity of causal inference in observational studies. Many techniques have been developed for study design and analysis to iden…
Shared by Dr Ellie Murray, ScD

How one US organization hopes to make retractions more visible
retractionwatch.com
Todd Carpenter As Retraction Watch readers likely know, there's ample evidence that retracted papers — 2,500 per year and growing — continue to attract citations that do not mention the fact the pape…
Shared by Retraction Watch
Causal Analysis in Theory and Practice
causality.cs.ucla.edu
I recently attended a seminar presentation by Professor Tom Belin, (AQM RAC seminar, UCLA, November 30, 2012) who spoke on the relationships between the potential outcome model of Neyman, Rubin and H…
Shared by Judea Pearl
Developing clinical prediction models when adhering to minimum sample size recommendations: The importance of quantifying bootstrap variability in tu…
journals.sagepub.com
Recent minimum sample size formula (Riley et al.) for developing clinical prediction models help ensure that development datasets are of sufficient size to minimise overfitting. While these criteri...
Shared by Glen Martin
To tune or not to tune, a case study of ridge logistic regression in small or sparse datasets
bmcmedresmethodol.biomedcentral.com
For finite samples with binary outcomes penalized logistic regression such as ridge logistic regression has the potential of achieving smaller mean squared errors (MSE) of coefficients and prediction…
Shared by Georg Heinze

S. Senn: To infinity and beyond: how big are your data, really? (guest post)
errorstatistics.com
Stephen Senn Consultant Statistician Edinburgh What is this n you boast about? Failure to understand components of variation is the source of much mischief. It can lead researchers to overlook that t…
Shared by Stephen John Senn



12 ideas to refresh your teaching in less than 10 minutes
educationalist.substack.com
The Educationalist. By Alexandra Mihai
Shared by Darren Dahly, PhD

What would be the most truthful way to draw inferences in a multilevel cell biology dataset?
discourse.datamethods.org
Hi, I’m a PhD candidate in cell biology that (in retrospect) has never received extensive statistical training on how to best analyse the most common experimental design in my field. Most of what I a…

EHRs and RCTs: Outcome Prediction vs. Optimal Treatment Selection | Statistical Thinking
fharrell.com
Frank Harrell Professor of Biostatistics Vanderbilt University School of Medicine Laura Lazzeroni Professor of Psychiatry and, by courtesy, of Medicine (Cardiovascular Medicine) and of Biomedical Dat…
Shared by Frank Harrell
Detection of Fraud in a Clinical Trial Using Unsupervised Statistical Monitoring
link.springer.com

A central statistical assessment of the quality of data collected in clinical trials can improve the quality and efficiency of sponsor oversight of clinica

Shared by Retraction Watch

Recommended Statistics Articles

hbiostat.org

Most of these articles were published in the past 60 days, and are recommended for their applicability to biomedical research. A very few serve as examples of how not to approach a problem.

Shared by Frank Harrell

Why do we do matching for causal inference vs regressing on confounders?

stats.stackexchange.com

I'm new to the area of causal inference. From what I understand, one of the main concerns that causal inference tries to address is the effect of confounders! For the sake of reference, let's denot...

Shared by Darren Dahly, PhD

Interim Analyses During Group Sequential Clinical Trials

jamanetwork.com

In the March 2, 2021, issue of JAMA, Locatelli et al1 reported a randomized clinical trial comparing a targeted immunotherapy agent, blinatumomab, with standard consolidation chemotherapy in children…

Editor's note: An article on group sequential analysis that pretends that Bayesian methods do not exist is problematic

Shared by Andrew Althouse

Covid vaccine myocarditis | Jay Brophy's Website

brophyj.com

On Sept 16 2021, 12 researchers from the Ottawa Heart Institute published a pre-print claiming that the COVID 19 mRNA vaccination myocarditis rate was 1 / 1000. On Sept 24 2021, the authors reported …

Shared by jay brophy

Site updates log | Prognosis Research

prognosisresearch.com

30/9/2021: Big update across the whole site, with updated information about courses, a new IPD meta-analysis textbook, new videos (see sample size, categorisation of continuous variables, controversi…

Shared by Richard Riley (R^2)

The problem with p-hacking is not the “hacking,” it’s the “p” (or, Fisher is just fine on this one) « Statistical Modeling, Causal Inference, and Soc…

statmodeling.stat.columbia.edu

On CrossValidated, a discussion came up I thought you may be interested in. The quick summary of it is that a poster posed the question that isn’t Fisher’s advice to go get more data when results are…

Shared by Andrew Gelman et al.

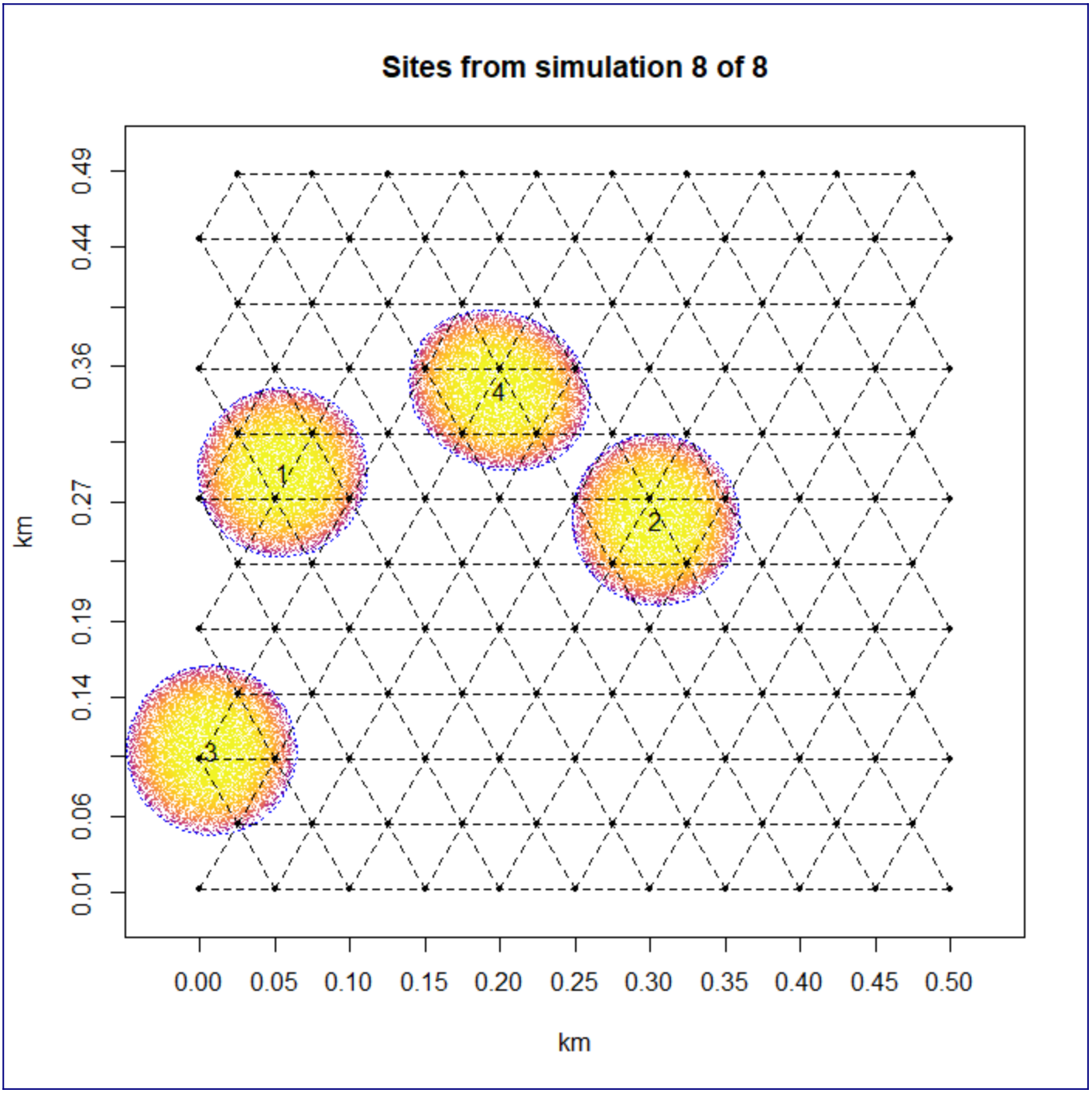
A teenager on TikTok disrupted thousands of scientific studies with a single video

theverge.com

After a TikTok sent a flood of new users to Prolific.co, a survey platform, researchers scrambled to salvage their research.

Shared by Retraction Watch

[August 2021: "Top 40" New CRAN Packages]
[rviews.rstudio.com]
[An R community blog edited by RStudio]
Shared by [RStudio]



Scale whose copyright owner defends zealously falls under scrutiny — and journal takes two years to publish a critique

retractionwatch.com

Donald Morisky As long-time readers of this blog know, we've spilled more than a few pixels on the work of Donald Morisky. His Morisky Medication Adherence Scale (MMAS) has been a financial boon to h…

Shared by Retraction Watch

Causality and causal inference in epidemiology: the need for a pluralistic approach

academic.oup.com

Abstract. Causal inference based on a restricted version of the potential outcomes approach reasoning is assuming an increasingly prominent place in the teachin

Shared by Jiří Hudeček

Students who grew up with search engines might change STEM education forever

theverge.com

Modern college students aren’t organizing their files into folders and directories, forcing some professors to rethink the way they teach programming.

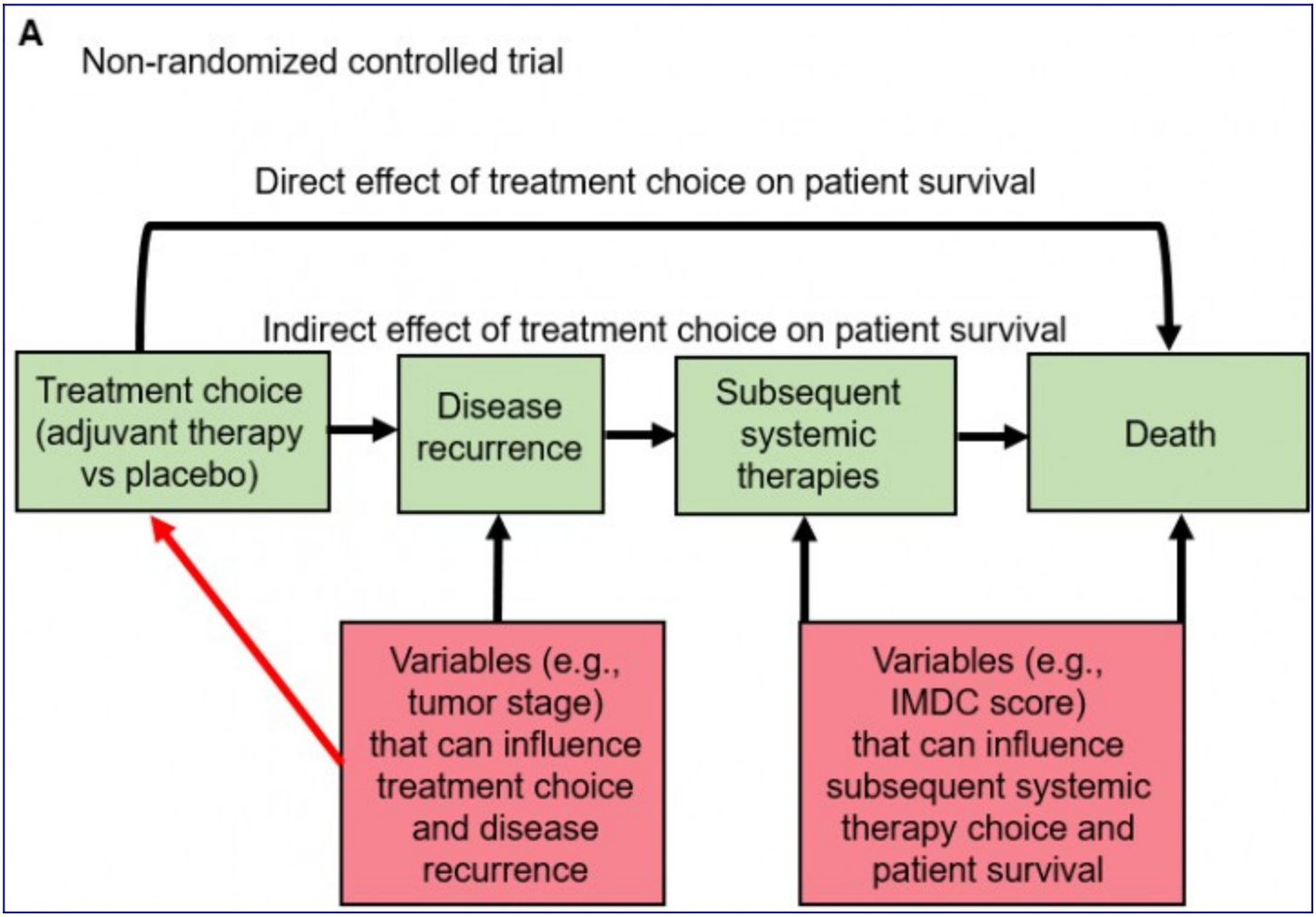
Editor's note: Once students are taught the importance of reproducible research we can lose all patience when they can't understand directory structures

Shared by Jeep Wrangler Owner

Overall survival vs disease-free survival and other surrogate endpoints

discourse.datamethods.org

The vast majority of studies evaluating the reliability of surrogate endpoints vs overall survival (OS) in interpreting RCTs simply make some kind of statistical correlation between the surrogate end…



A Guideline for Reporting Mediation Analyses of Randomized Trials and Observational Studies: The AGReMA Statement
jamanetwork.com
Question What information should be reported in studies that include mediation analyses of randomized trials and observational studies? Findings An international Delphi and consensus process (using t…
Shared by Gary Collins

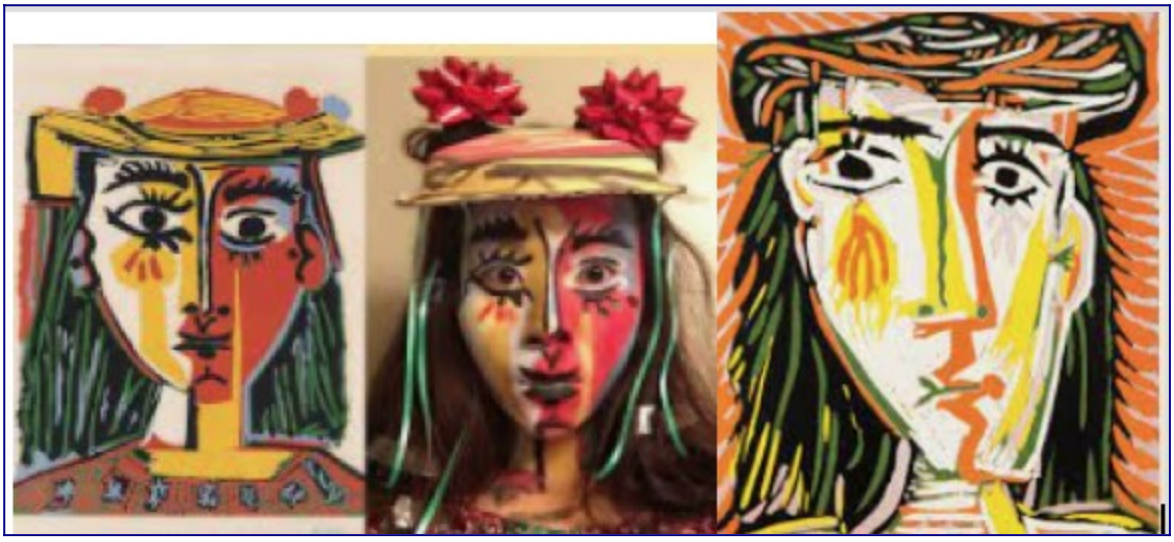
Significance tests, p-values, and falsificationism
discourse.datamethods.org
This thread takes its inspiration from the recent discussions in social science and statistics about significance tests, what they're good for, whether p-values should be banned, and what all of that…
Shared by Frank Harrell

[More on the epidemiologists who other epidemiologists don’t trust « Statistical Modeling, Causal Inference, and Social Science]
statmodeling.stat.columbia.edu
You know how they’ll describe someone as a musician’s musician? Or a chef’s chef? There’s also the opposite: those people who to outsiders represent a certain profession, but who are not respected by…
Shared by Andrew Gelman et al.

[“Fabulous document”, “very helpful guidance”: Sleuths react to recommendations for handling image integrity issues]
retractionwatch.com
Retraction Watch readers are likely familiar with the varied -- and often unsatisfying -- responses of journals to scientific sleuthing that uncovers potential problems with published images. Some ed…
Shared by Retraction Watch

[Reproducible data analysis workflows with R and RStudio]
statsepi.substack.com
Some tips on preventing, identifying and correcting errors in your data analyses
Shared by Darren Dahly, PhD

[Should Bayesian Clinical Trialists Wear Error Statistical Hats? (i)]
errorstatistics.com
I. A principled disagreement The other day I was in a practice (zoom) for a panel I’m in on how different approaches and philosophies (Frequentist, Bayesian, machine learning) might explain "why we d…
Editor's note: This is an interesting discussion but is based on a an unclear notion of "error" that distorts some of the points made.
Shared by Frank Harrell



Is Any Method So Obsolete It Shouldn't Be Used?
discourse.datamethods.org
Are there any regression modeling strategies that should be deemed (authoritatively) to be now (2021) "obsolete?" Obsolete means—"there is always a better alternative/do not do this because the strat…

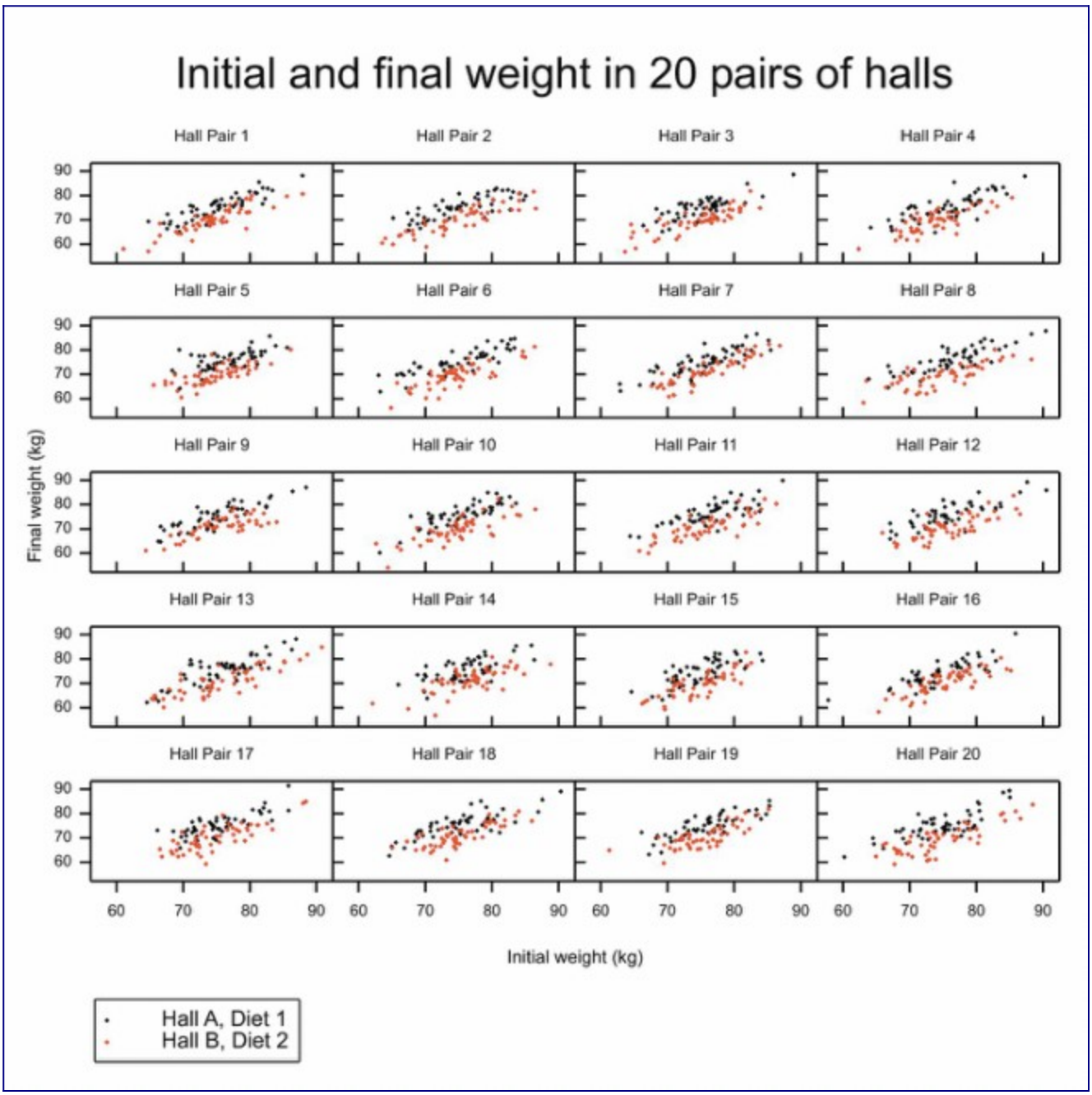
"For a research assistant, do you think there is an ethical responsibility to inform your supervisor/principal investigator if they change their anal…
statmodeling.stat.columbia.edu
A reader who works as a research assistant and who wishes to remain anonymous writes: I have a hypothetical question about ethics in statistics. For a research assistant, do you think there is an eth…
Shared by Andrew Gelman et al.

How the Pandemic Is Changing the Norms of Science
tabletmag.com
Imperatives like skepticism and disinterestedness are being junked to fuel political warfare that has nothing in common with scientific methodology
Shared by Vinay Prasad, MD MPH

C. R. Rao
en.wikipedia.org
Calyampudi Radhakrishna Rao, FRS known as C R Rao (born 10 September 1920) is an Indian-American mathematician and statistician. He is currently professor emeritus at Pennsylvania State University an…
Shared by Stephen John Senn

Statistical Reform
statsepi.substack.com
We see the impact of poor research methods everywhere around us. Reviews focusing on the appropriate use and reporting of research methods are almost universally depressing. Chalmers and Glasziou hav…
Shared by Maarten van Smeden

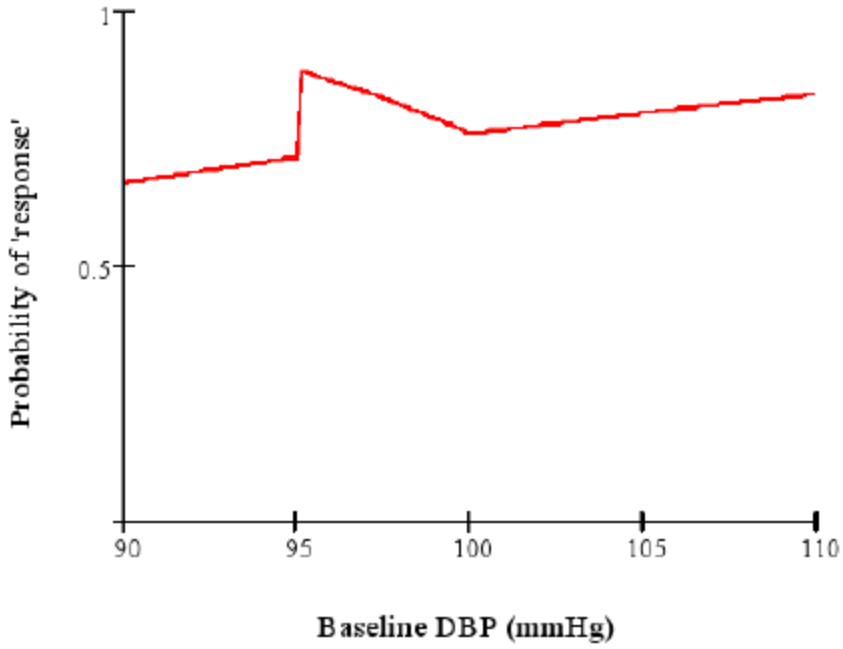
S. Senn: Red herrings and the art of cause fishing: Lord’s Paradox revisited (Guest post)
errorstatistics.com
Stephen Senn Consultant Statistician Edinburgh Background Previous posts[a],[b],[c] of mine have considered Lord’s Paradox. To recap, this was considered in the form described by Wainer and …
Shared by Stephen John Senn



[Responder Analysis: Loser x 4]
discourse.datamethods.org
[Clinician-scientists involved in clinical trials who have inadequate statistical input often design clinical studies to use a binary responder analysis. They seem to believe that because a binary tre…]
Shared by [Frank Harrell]

- Example: "The treatment is called successful if either the patient has gone down from a baseline diastolic blood pressure of ≥ 95 mmHg to ≤ 90 mmHg or has achieved a 10% reduction in blood pressure from baseline."

- Senn derived the response probabililty function for this discontinuous concocted endpoint



Author Checklist
discourse.datamethods.org
Statistical Problems to Document and to Avoid Checklist for Authors References | Ten Simple Rules | Checklist to Avoid p-Hacking | Research Reliability and Publishing | Guidelines for Reporting of St…
Shared by Frank Harrell

The scandal of poor medical research
bmj.com
We need less research, better research, and research done for the right reasons What should we think about a doctor who uses the wrong treatment, either wilfully or through ignorance, or who uses the…
Shared by Darren Dahly, PhD
A Guide to Binge Watching R / Medicine 2021
rviews.rstudio.com
An R community blog edited by RStudio

Shared by RStudio

Significance tests, p-values, and falsificationism
discourse.datamethods.org
This thread takes its inspiration from the recent discussions in social science and statistics about significance tests, what they're good for, whether p-values should be banned, and what all of that…
Shared by Frank Harrell

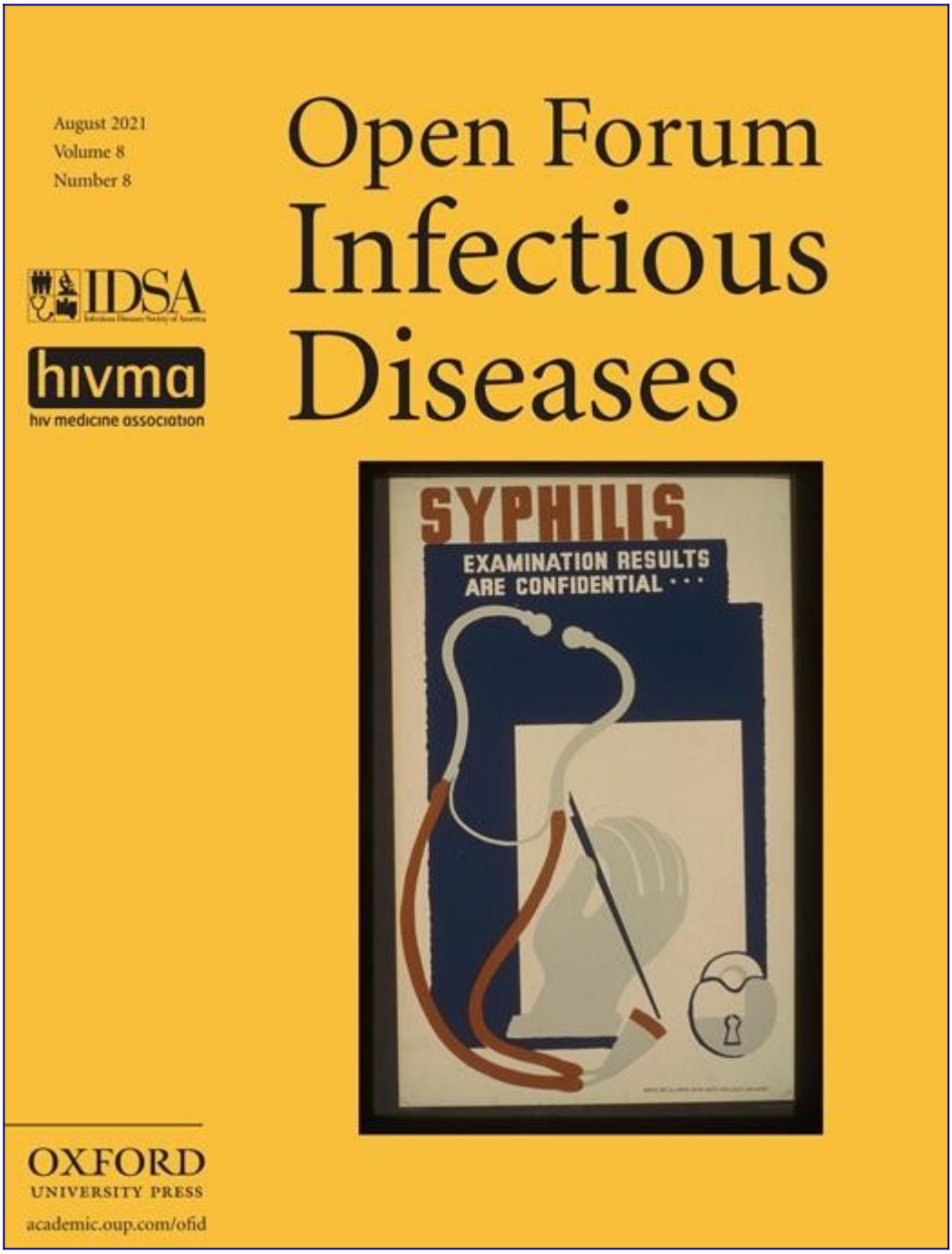
Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests
bmj.com
Many decisions in medicine involve trade-offs, such as between diagnosing patients with disease versus unnecessary additional testing for those who are healthy. Net benefit is an increasingly reporte…
Shared by Richard D Riley

$$\text{Net benefit} = \frac{\text{True positives}}{N} - \frac{\text{False positives}}{N} \times \frac{p_t}{1-p_t}$$

When Researchers Sound the Alarm on Problematic Papers
the-scientist.com

Finding and reporting an irregularity in a published study can lead people down an unexpected path.
Shared by Retraction Watch

Overlooked Shortcomings of Observational Studies of Interventions in Coronavirus Disease 2019: An Illustrated Review for the Clinician
academic.oup.com
An illustrated review of shortcomings of observational studies of interventions in COVID-19 with focus on treatment selection bias due to “goals of care” and me
Shared by Eric Wenzler



[Baseline Balance and Valid Statistical Analyses: Common Misunderstandings](appliedclinicaltrialsonline.com)
[A simple game of chance clears up the confusion over what randomization can and cannot do.]
Shared by [Stephen John Senn]

Table 1. Sample space for a game of chance involving two dice

	Red Die Score					
Black Die Score	**1**	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	(*10*)
5	6	7	8	9	(*10*)	11
6	7	8	9	(*10*)	11	12

Should we ignore covariate imbalance and stop presenting a stratified 'table one' for randomized trials?
discourse.datamethods.org
Suppose that the effect of an intervention versus control on an outcome (denoted ‘Y’) in a randomized trial will be assessed in an unadjusted fashion, e.g., using a t-test. We know that the operating…
Shared by Frank Harrell

Most medical practices are not parachutes: a citation analysis of practices felt by biomedical authors to be analogous to parachutes
pubmed.ncbi.nlm.nih.gov
Although there is widespread interest regarding the BMJ paper arguing that randomized trials are not necessary for practices of clear benefit, there are few analogies in medicine. Most parachu…
Shared by Vinay Prasad, MD MPH

Navigating Tough Conversations in Statistical Collaboration | Amstat News

magazine.amstat.org

Uncomfortable conversations are inevitable in many research and professional relationships, including statistical collaborations. Topics can vary widely due to the different backgrounds and expectati…

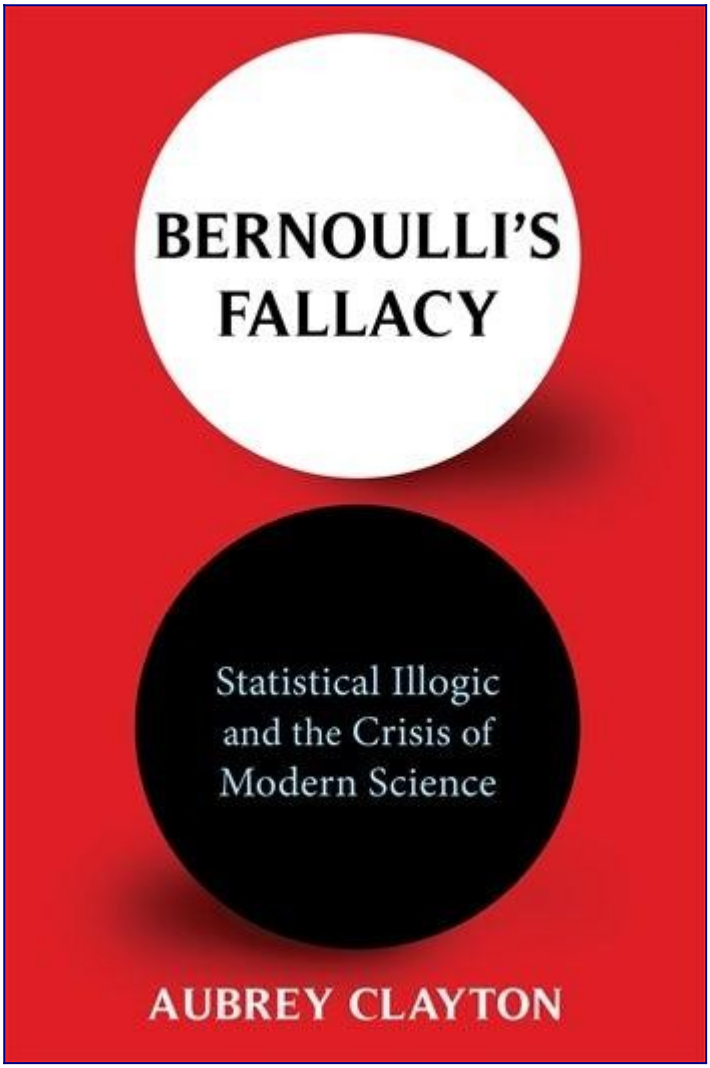
Shared by ASA

AMSTAT NEWS

Significance tests, p-values, and falsificationism
discourse.datamethods.org
This thread takes its inspiration from the recent discussions in social science and statistics about significance tests, what they're good for, whether p-values should be banned, and what all of that…

Table 2 Fallacy: Presenting and Interpreting Confounder and Modifier Coefficients
academic.oup.com
Abstract. It is common to present multiple adjusted effect estimates from a single model in a single table. For example, a table might show odds ratios for one
Shared by Richard McElreath

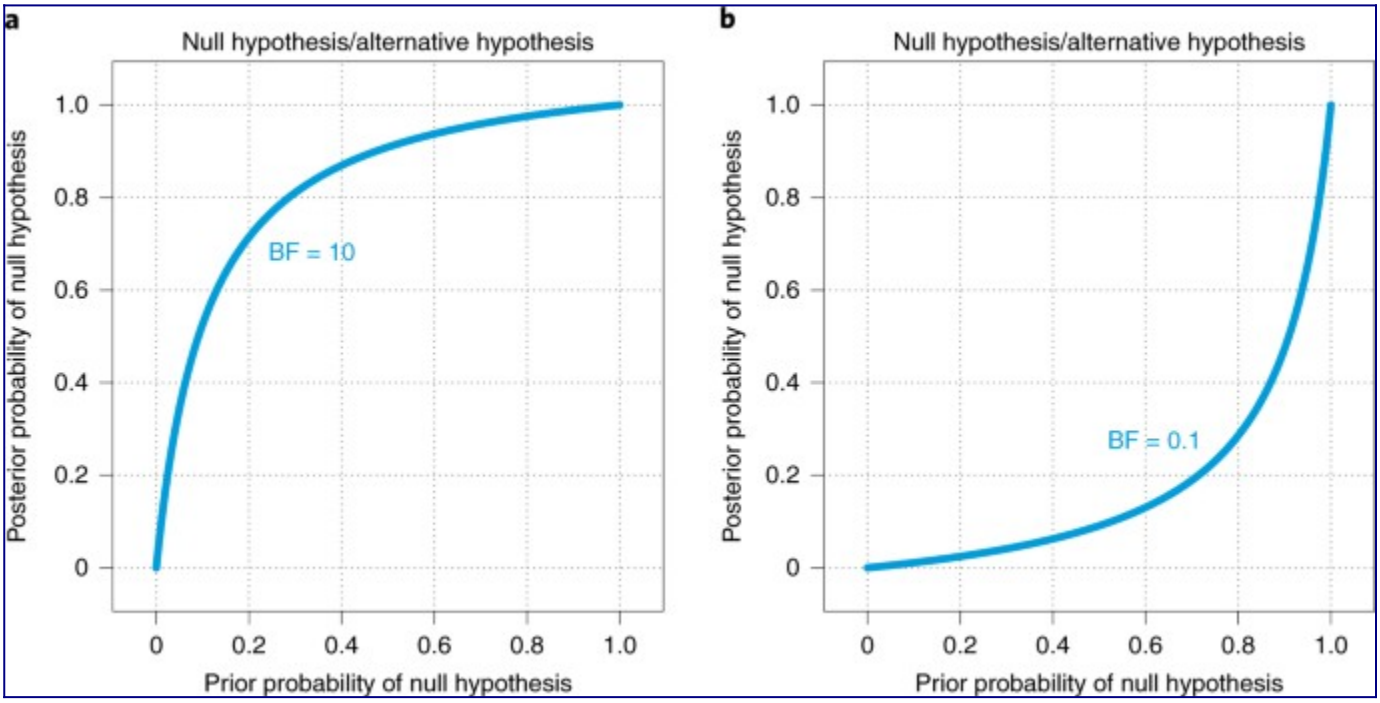
Bernoulli's Fallacy
cup.columbia.edu
There is a logical flaw in the statistical methods used across experimental science. This fault is not a minor academic quibble: it underlies a reproducibili... | CUP
Shared by Arthur Charpentier



An Analysis of 5 Years of Randomized Trials in Gastroenterology and Hepatology Reveals 52 Medical Reversals
link.springer.com
One manifestation of low-value medical practice is the medical reversal, a practice in widespread use that, once subjected to a randomized controlled trial
Shared by Vinay Prasad, MD MPH

A cluster randomised trial of cloth masks compared with medical masks in healthcare workers
bmjopen.bmj.com
Editor's Note The authors of this article, published in 2015, have written a response to their work in light of the COVID-19 pandemic. We urge our readers to consider the response when reading the ar…
Shared by Vinay Prasad, MD MPH

Bayesian Analysis Reporting Guidelines
nature.com
Previous surveys of the literature have shown that reports of statistical analyses often lack important information, causing lack of transparency and failure of reproducibility. Editors and authors a…
Shared by EJ Wagenmakers



‘On the Fringe’ explores the thin line between science and pseudoscience
sciencenews.org

In his latest book, historian Michael Gordin surveys astrology, eugenics and other fringe movements to show how challenging it is to define pseudoscience.

Shared by Frank Harrell

I am asymptomatic with a negative rapid test. What is the probability I have coronavirus?

discourse.datamethods.org

I did this for myself this weekend. I could not find anything online that provided the plot I do here. I wanted to post this somewhere in case someone else was in a similar situation, but I hoped to …

Peer review is not the best way to promote major breakthroughs

timeshighereducation.com

Identifying and seed-funding scientists with ground-breaking ideas is a low-risk, high-reward alternative to traditional grant funding, says Donald Braben

Shared by Retraction Watch

Edward Tufte on Data Visualizations and Art
insights.dice.com
Statistician and data-visualization pioneer Edward Tufte explains how his data visualization work informs his conceptual artwork, and vice versa.
Shared by Edward Tufte

Concerns About the Special Article on Hydroxychloroquine and Azithromycin in High-Risk Outpatients With COVID-19

academic.oup.com

Abstract. In May 2020, the Journal published an opinion piece by a member of the Editorial Board, in which the author reviewed several papers and argued that us

Shared by Dr Ellie Murray, ScD

Does the “Table 1 fallacy” apply if it is Table S1 instead? « Statistical Modeling, Causal Inference, and Social Science

statmodeling.stat.columbia.edu

In a randomized experiment (i.e. RCT, A/B test, etc.) units are randomly assigned to treatments (i.e. conditions, variants, etc.). Let’s focus on Bernoulli randomized experiments for now, where each …

Shared by Andrew Gelman et al.

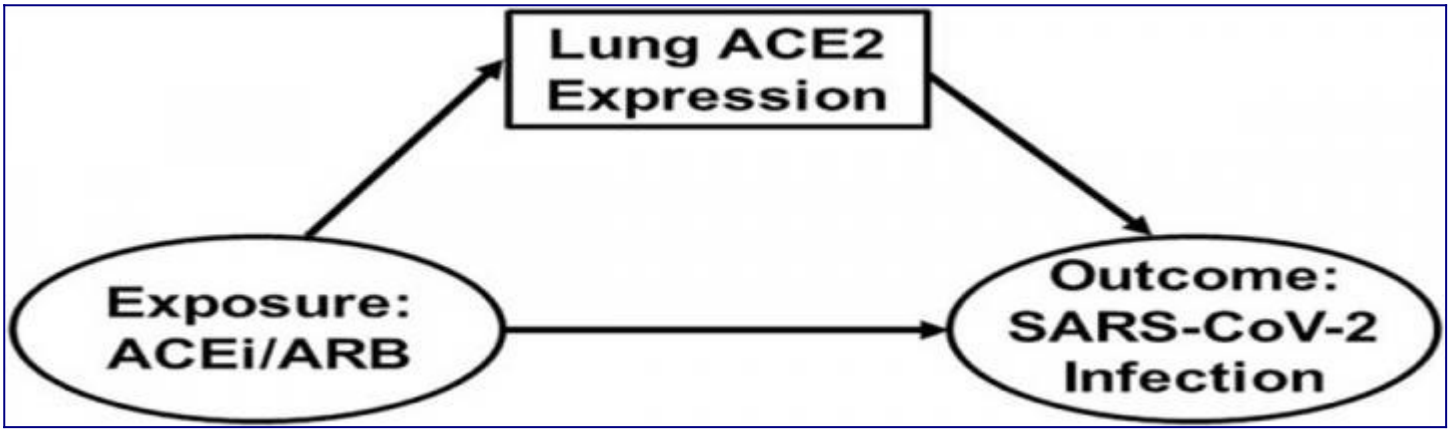
The Swiss cheese infographic that went viral

virologydownunder.com

Some behind-the-scenes facts about how Swiss cheese proved useful to explain complex risk reduction during the COVID-19 pandemic.

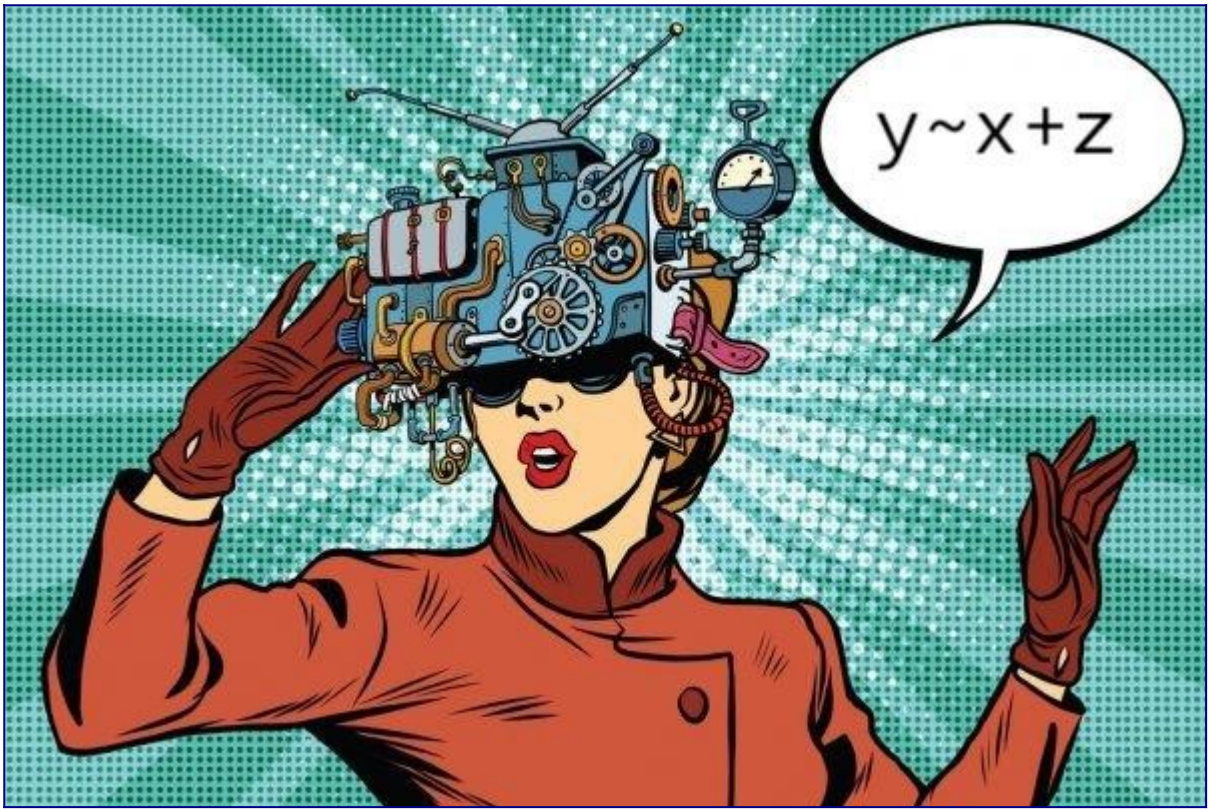
Shared by WesElyMD

The horrible convoluted statistical procedures that are used to make decisions about something as important as Alzheimer's treatments « Statistical M…
statmodeling.stat.columbia.edu
Richard Juster points to this article, "Evaluation of Aducanumab for Alzheimer Disease: Scientific Evidence and Regulatory Review Involving Efficacy, Safety, and Futility." The article begins: On Nov…
Shared by Andrew Gelman et al.
Evaluating sources of bias in observational studies of... : Journal of Hypertension
journals.lww.com
Concerns over ACE inhibitor or ARB use to treat hypertension during COVID-19 remain unresolved. Alth
Shared by Andrew Michael South, MD, MS



Top Honesty Researcher, Dan Ariely, Has Paper Retracted Due to Fraudulent Data
fraudbytes.blogspot.com
Thoughts on fraud prevention and detection by Mark and Aaron Zimbelman.
Shared by Boback Ziaeian

Regression, Fire, and Dangerous Things (1/3)
elevanth.org
It isn't my job to disappoint people, but I'm good at it.
Shared by Richard McElreath

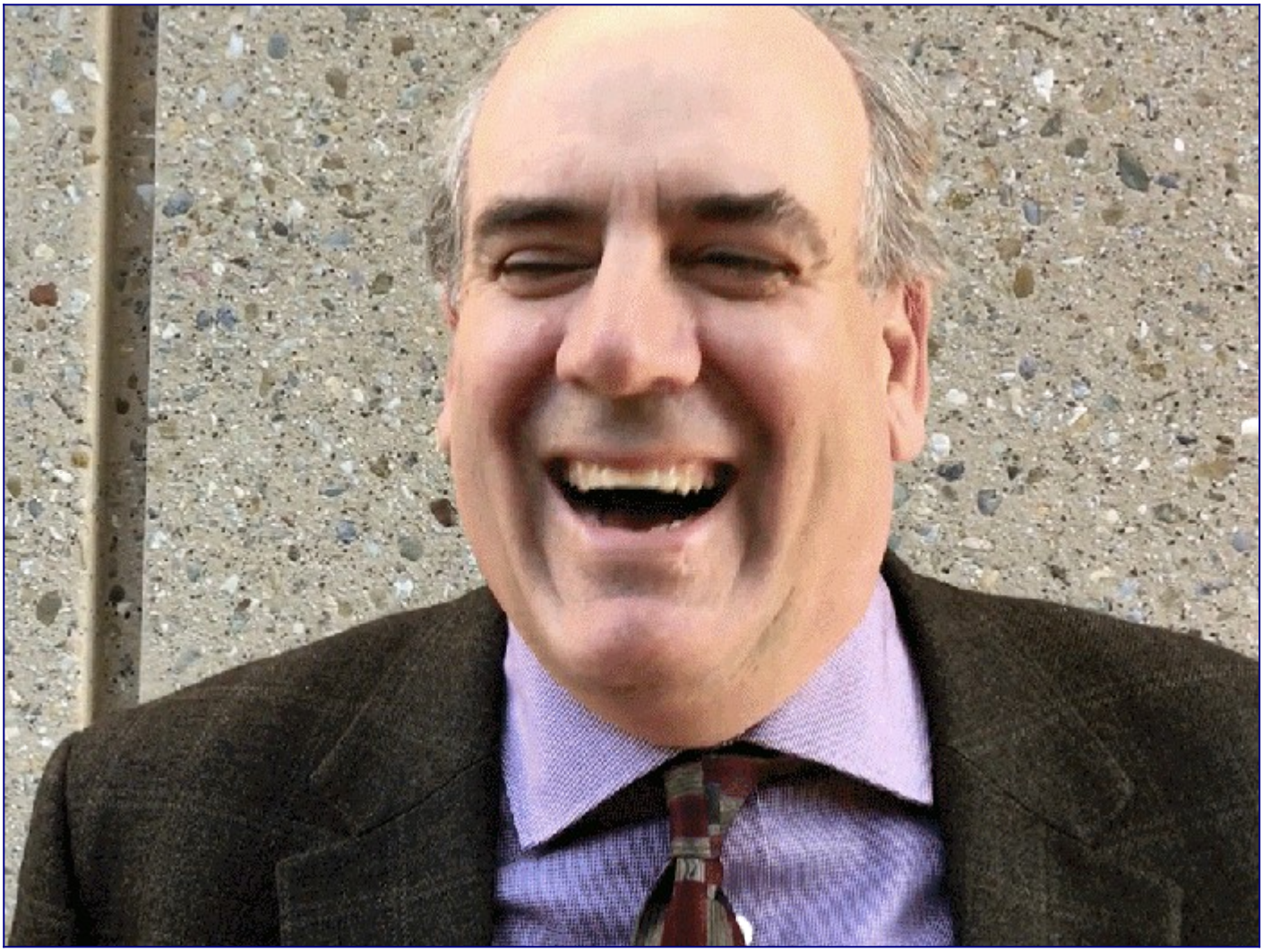


Not Even Scientists Can Easily Explain P-values

fivethirtyeight.com

P-values have taken quite a beating lately. These widely used and commonly misapplied statistics have been blamed for giving a veneer of legitimacy to dodgy stu…

Shared by Frank Harrell



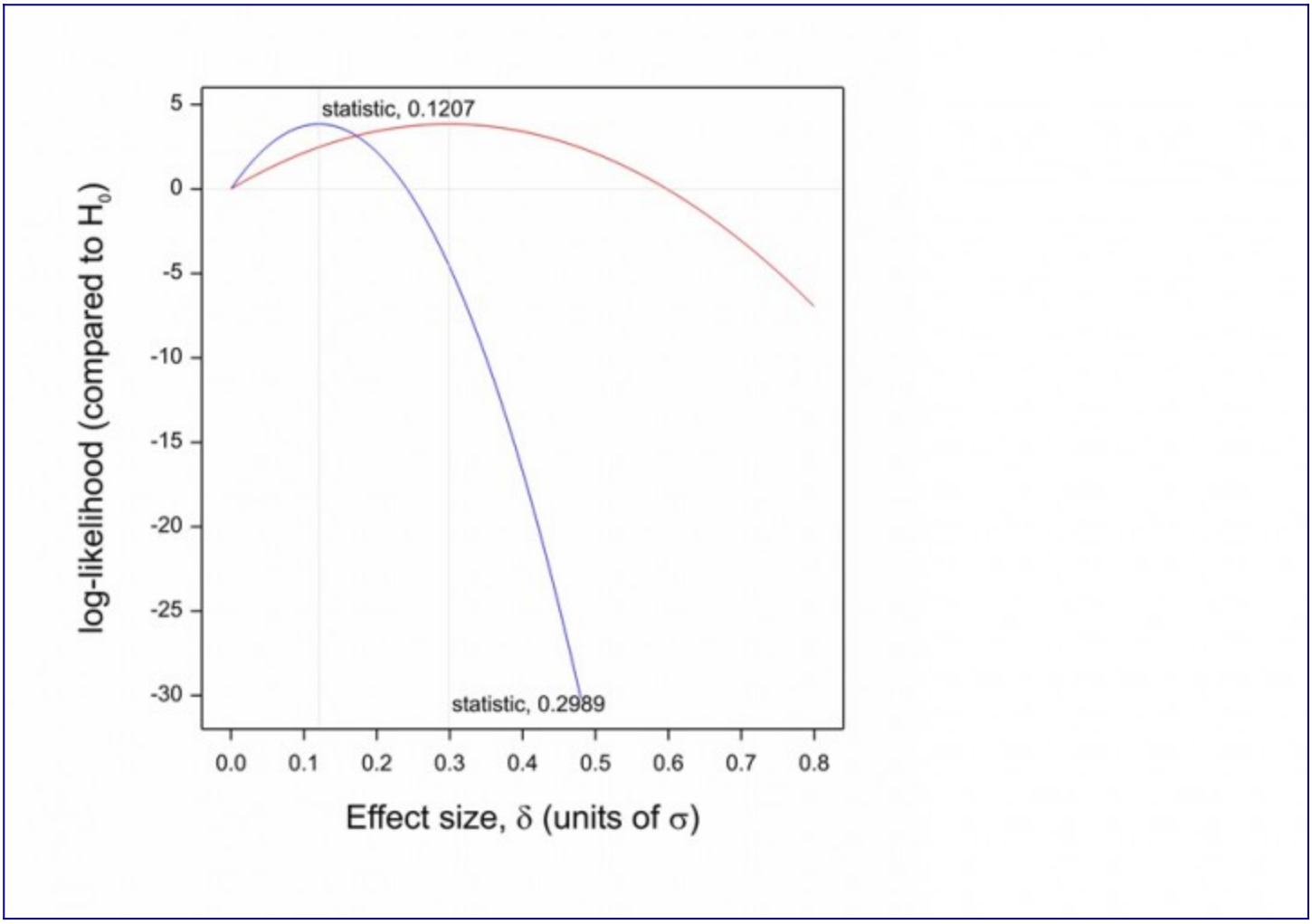
My Journey From Frequentist to Bayesian Statistics | Statistical Thinking
fharrell.com
The difference between Bayesian and frequentist inference in a nutshell: With Bayes you start with a prior distribution for θ and given your data make an inference about the θ-driven process generati…
Shared by Frank Harrell

Autocorrect errors in Excel still creating genomics headache
nature.com
Despite geneticists being warned about spreadsheet problems, 30% of published papers contain mangled gene names in supplementary data.
Shared by Brian Nosek

Six tips for effective e-learning
cebm.ox.ac.uk
I'm writing this blog as the second iteration of my module on Course Design, Assessment and Evaluation is now running. The module is part of a postgraduate certificate course in teaching evidence-bas…
Shared by Darren Dahly, PhD

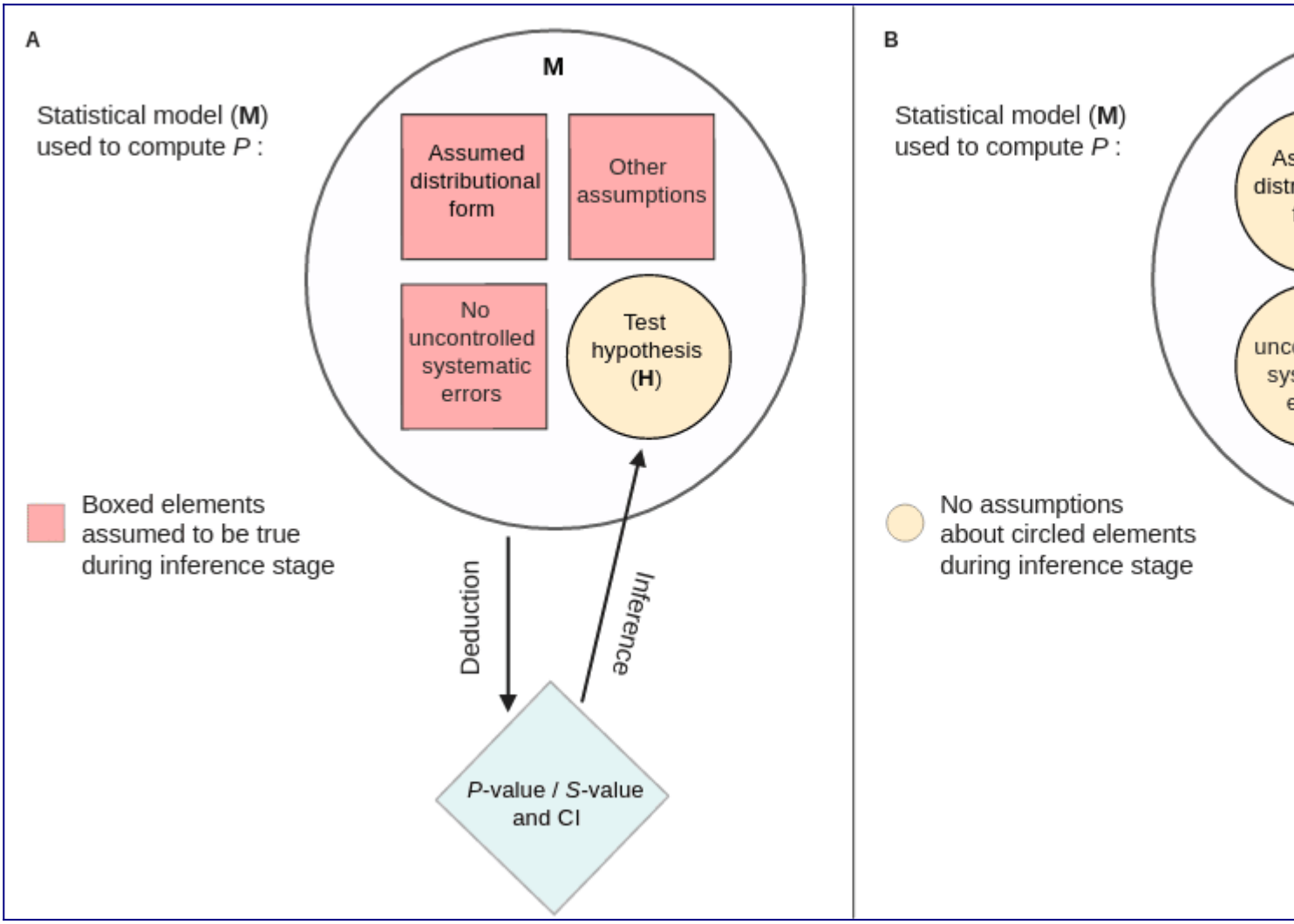
Academic conclusions differ wildly even on same data, study finds
timeshighereducation.com
New study raises questions over validity of single studies and suggests detailed tracking of researchers' decisions during analysis
Shared by Retraction Watch

Color schemes in data graphics « Statistical Modeling, Causal Inference, and Social Science
statmodeling.stat.columbia.edu
Natesh Pillai points us to this recent article, "The misuse of colour in science communication," which begins: The accurate representation of data is essential in science communication. However, colo…
Shared by Andrew Gelman et al.
S. Senn: "Automatic for the people? Not quite" (Guest post)
errorstatistics.com
Stephen Senn Head of Competence Center for Methodology and Statistics (CCMS) Luxembourg Institute of Health Twitter @stephensenn Automatic for the people? Not quite What caught my eye was the estimab…
Shared by Frank Harrell



Is stratified randomisation in trials (at least large ones) pointless?
thestatsgeek.com
I recently wrote a blog post giving some comments about the revised FDA guidance on covariate adjustment in randomised trials. One of the things I commented on was recent work by different authors on…
Shared by Stephen John Senn

P-Values Are Tough and S-Values Can Help
lesslikely.com
An extensive discussion about what P-values are, their properties, common interpretations, misinterpretations, and how a measure called an S-value may better help us interpret them.
Shared by Darren Dahly, PhD



[Extensions of 'worst-rank' score methods with outcomes truncated by death](#)
discourse.datamethods.org
I am helping with the planning of a clinical trial looking at an intervention in a particular group of newborns. Primary interest is in the effect on cognitive delay, and will be assessed using a sta…

[Language for communicating frequentist results about treatment effects](#)
discourse.datamethods.org
Editorial Notes As Sander Greenland has so well stated below, this initial draft was still too “dichotomous” in that it was written assuming we would have different language for “positive” vs. “negat…
Shared by Frank Harrell

A grim warning from Israel: Vaccination blunts, but does not defeat Delta
sciencemag.org
With early vaccination and outstanding data, country is the world's real-life COVID-19 lab
Shared by C. Michael Gibson MD

Forecast displays that emphasize uncertainty « Statistical Modeling, Causal Inference, and Social Science
statmodeling.stat.columbia.edu
This is Jessica. G. Elliot Morris points me to some recent visualizations of the Economist's forecast for the upcoming election in Germany, which emphasize a range of plausible outcomes over point es…
Shared by Andrew Gelman et al.

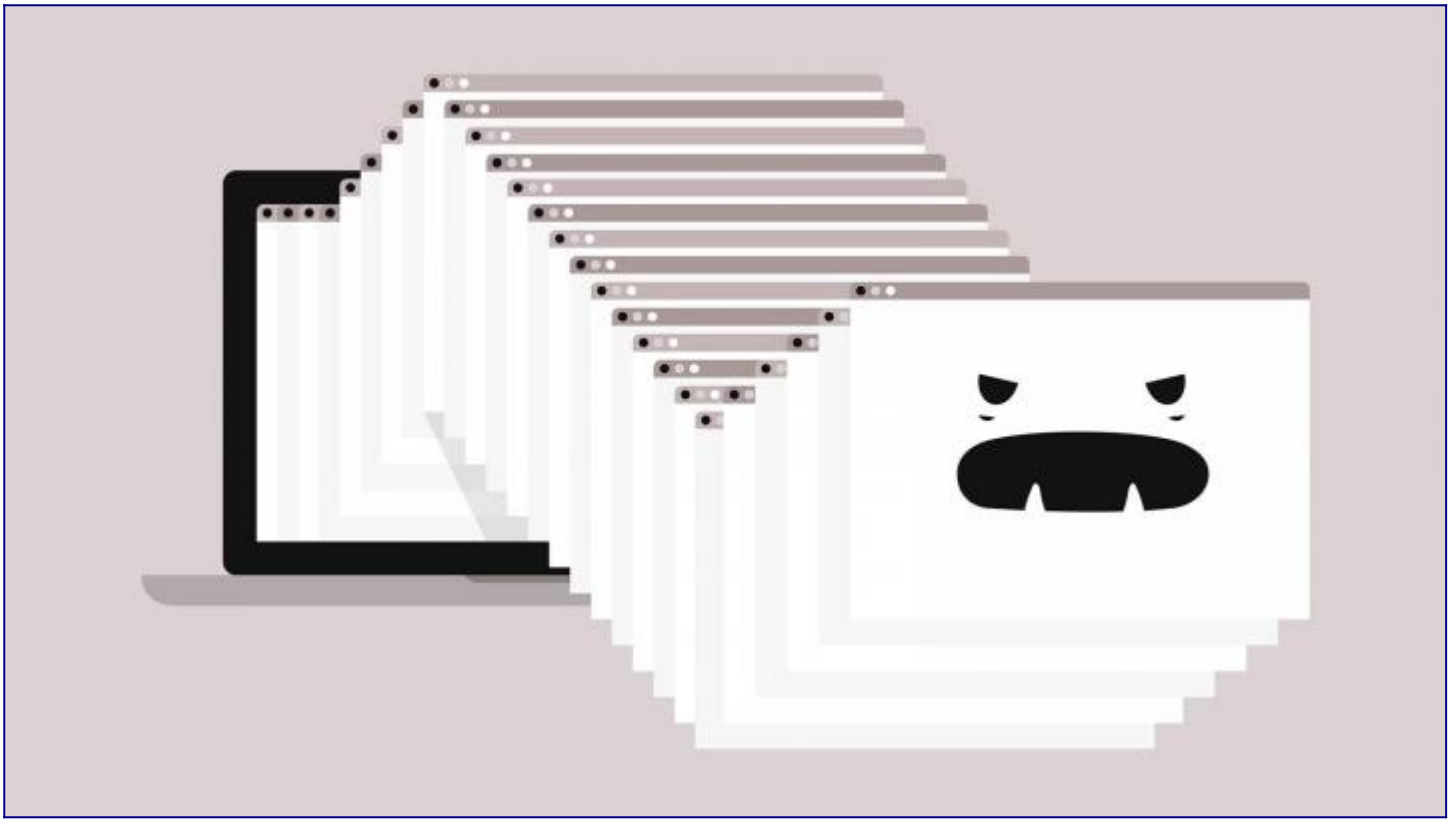
R / Medicine 2021
rviews.rstudio.com
An R community blog edited by RStudio
Shared by RStudio

The Complicated Legacy of an Anti-Tobacco Hero
undark.org
A lauded tobacco scientist's crusade against vaping has some critics — and former allies — questioning his research.
Shared by Retraction Watch

P-Hacking Tutorial
youtu.be
Garden of forking paths; analysis to a foregone conclusion; multiplicity; investigator degrees of freedom; thanks to James Smith
Shared by Frank Harrell
Telling good science from bad – a user's guide to navigating the scientific literature
wiringthebrain.com
"Did you find it convincing?" That's what one of my genetics professors used to ask us, a small group of undergraduates who blinked in respo...
Shared by Darren Dahly, PhD
Fair shares: sexual justice in patient recruitment in clinical trials
errorstatistics.com
Stephen SennConsultant StatisticianEdinburgh, Scotland It is hard to argue against the proposition that approaches to clinical research should treat not only men but also women fairly, and of course …
Shared by Stephen John Senn

[Extraordinary claims, extraordinary evidence? A discussion](#)
[link.springer.com](#)
[Roberts (2020, Learning & Behavior, 48[2], 191–192) discussed research claiming honeybees can do arithmetic. Some readers of this research might](#)
Shared by [EJ Wagenmakers](#)

[Down With the Social Media Platforms](#)
[medpagetoday.com](#)
[Simply policing misinformation doesn't address the real issue at hand](#)
Shared by [Vinay Prasad, MD MPH](#)



Clinician Conceptualization of the Benefits of Treatments for Individual Patients
jamanetwork.com
Question How do clinicians conceptualize the benefits of treatments for common diseases? Findings In this survey study of 542 clinicians, most respondents significantly overestimated the benefits of …
Shared by Edward Tufte

On zombies, struldbrugs, and other horrors of the scientific literature
pnas.org
When I signed on as Editor-in-Chief of PNAS, I had no idea that killing zombies would be part of the job. These zombies aren't the spirits of Haitian mythology doomed to roam the earth in thrall to s…
Shared by Retraction Watch

Rediscovery of calculus in 1994: what should have happened to that paper?
academia.stackexchange.com
There's a well-known paper in academic circles that features a rediscovery of the trapezoidal rule for numerical integration by a medical researcher: “A Mathematical Model for the Determination of
Shared by Boback Ziaeian

“A calculated risk”: the Salk polio vaccine field trials of 1954
ncbi.nlm.nih.gov
Competing interests: MM has presented papers at meetings sponsored in part by pharmaceutical companies. The polio vaccine field trials of 1954, sponsored by the National Foundation for Infantile Para…
Shared by Venk Murthy

Is stratified randomisation in trials (at least large ones) pointless?
thestatsgeek.com
I recently wrote a blog post giving some comments about the revised FDA guidance on covariate adjustment in randomised trials. One of the things I commented on was recent work by different authors on…
Shared by Andrew Althouse
Hundreds of AI tools have been built to catch covid. None of them helped.
technologyreview.com
The AI community rushed to create predictive covid-19 tools. None of them made a real difference, and some were potentially harmful. But it could make AI more collaborative.

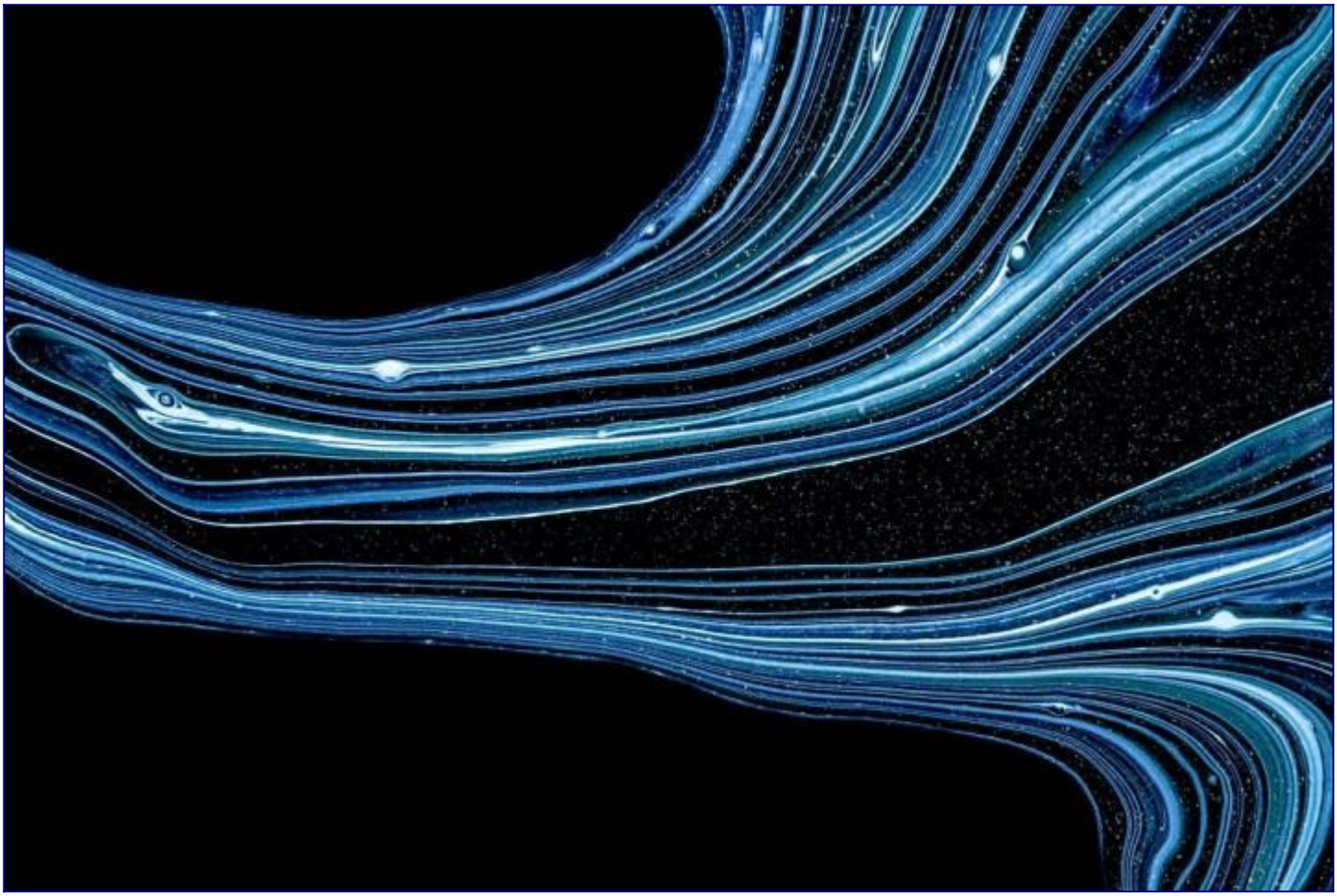
Shared by Gary Collins

Microbiologist Elisabeth Bik queried Covid research – that's when the abuse and trolling began
theguardian.com
Bik's home address was posted on Twitter and she faced a barrage of attacks after dissecting a paper endorsing hydroxychloroquine as treatment When Dr Elisabeth Bik raised serious concerns about the …

Shared by Retraction Watch

Covid-19—anatomy of a methodology crisis. It all starts with data collection
towardsdatascience.com
Abstract: Without reliable measures of prevalence (viral circulation) and severity (the probability that infections lead to severe forms such as hospitalisation), the medium-term impact of viral…

Shared by Towards Data Science

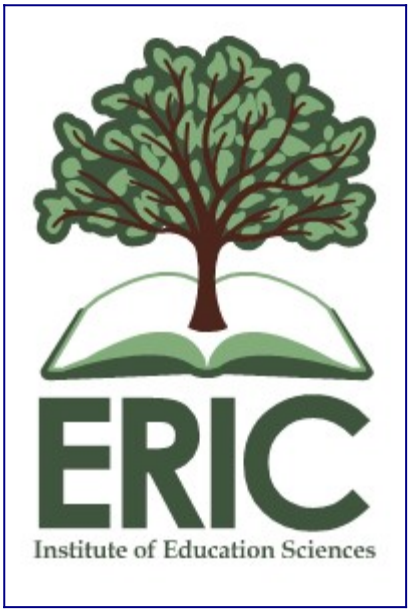


ERIC - EJ1152511 - Against Inferential Statistics: How and Why Current Statistics Teaching Gets It Wrong, Statistics Education Research Journal, 2017…

eric.ed.gov

Recent concerns about a shortage of capacity for statistical and numerical analysis skills among social science students and researchers have prompted a range of initiatives aiming to improve teachin…

Shared by Frank Harrell



Reference Collection to push back against "Common Statistical Myths"
discourse.datamethods.org
Note: This topic is a wiki, meaning that this main body of the topic can be edited by others. Use the Reply button only to post questions or comments about material contained in the body, or to sugge…
Shared by German Demidov

Why is accuracy not the best measure for assessing classification models?
stats.stackexchange.com
This is a general question that was asked indirectly multiple times in here, but it lacks a single authoritative answer. It would be great to have a detailed answer to this for the reference. Accu...
Shared by Frank Harrell

The review I wanted to write…
statsepi.substack.com
What happens when people who know better still try to publish their garbage?

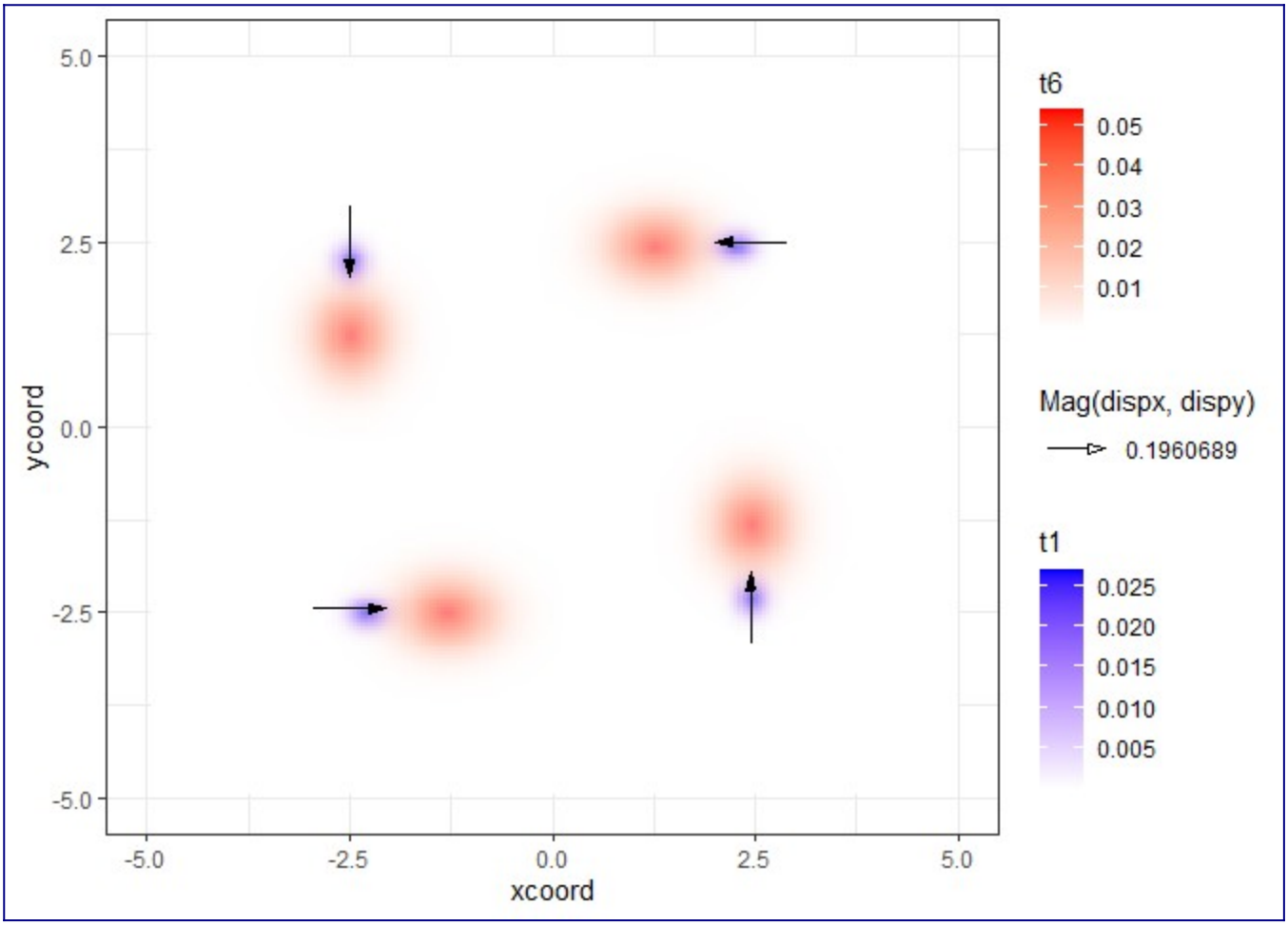
Shared by Darren Dahly, PhD

Surrogate endpoints in oncology: when are they acceptable for regulatory and clinical decisions, and are they currently overused?
pubmed.ncbi.nlm.nih.gov
The use of surrogate outcomes should be limited to situations where a surrogate has demonstrated robust ability to predict meaningful benefits, or where cases are dire, rare or with few treatment opt…
Shared by Vinay Prasad, MD MPH

Invited Commentary: Treatment Drop-in: Making the Case for Causal Prediction
academic.oup.com
Abstract. Clinical prediction models (CPMs) are often used to guide treatment initiation, with individuals at high risk offered treatment. This implicitly assum
Shared by Matthew Sperrin
What does public health really mean? Lessons from covid-19 - The BMJ
blogs.bmj.com

Throughout the covid-19 pandemic, there have been fierce debates about the scope of public health and the differential utility of recommendations versus mandates. With the rapid development of highly…
Shared by Vinay Prasad, MD MPH

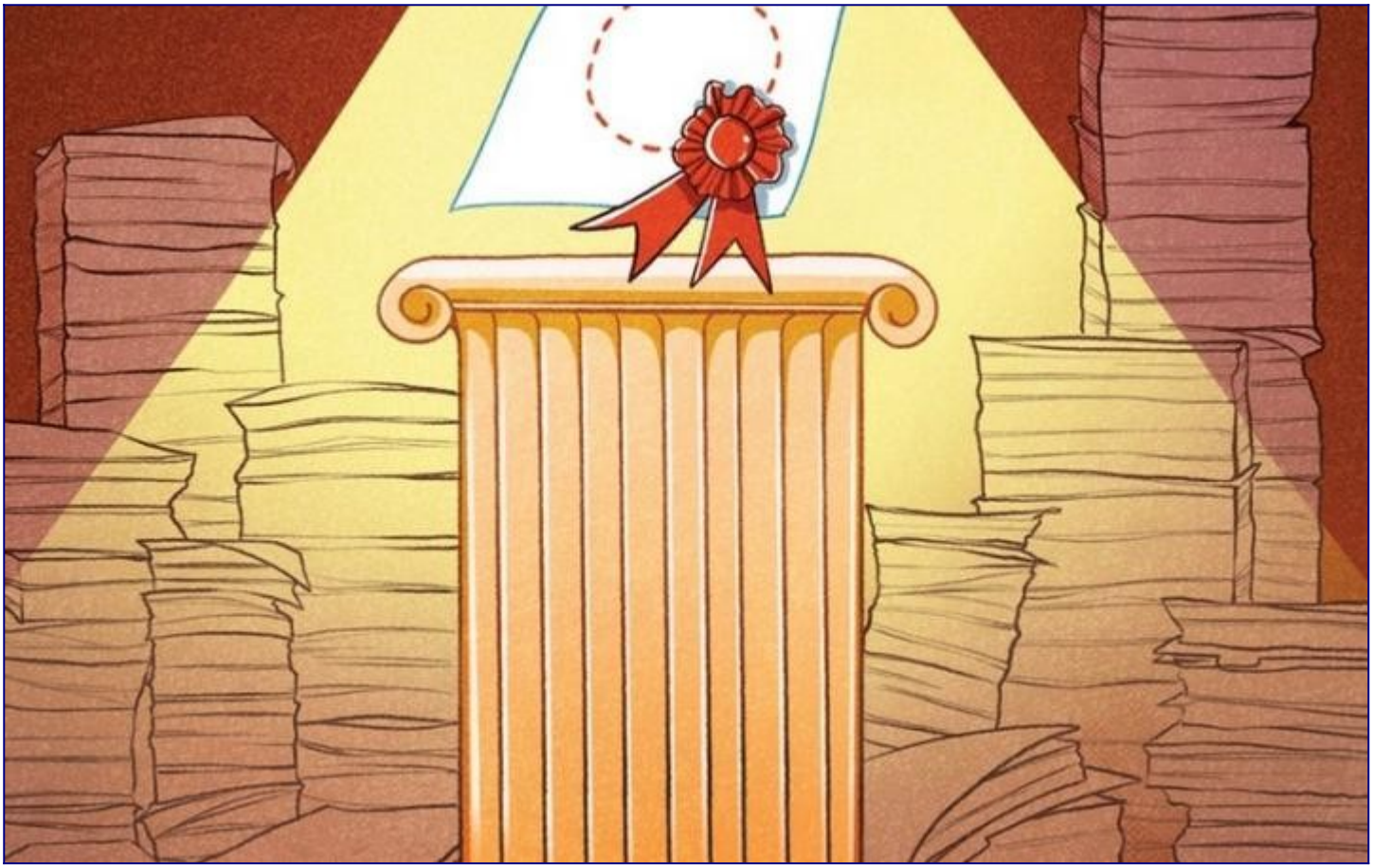
June 2021: "Top 40" New CRAN Packages
rviews.rstudio.com
An R community blog edited by RStudio
Shared by RStudio



Recent Articles in Statistics Journals
hbiostat.org
Articles largely published in the past 60d that are recommended because of their applicability to current problems. A few of the articles are critiqued.
Critiques are in the Notes section. A small minority of articles are recommended because they serve as examples of how not to solve a problem.
Shared by Frank Harrell

Top 5 Statistical Faux Pas
methodspace.com
MethodSpace is a multidimensional online network for the community of researchers, from students to professors, engaged in research methods. Sponsored by SAGE Publishing, a leading publisher of books…
Shared by Frank Harrell

Why Bad Science Is Sometimes More Appealing Than Good Science
scientificamerican.com
Scientific American is the essential guide to the most awe-inspiring advances in science and technology, explaining how they change our understanding of the world and shape our lives.
Shared by Retraction Watch



‘Tortured phrases’, lost in translation: Sleuths find even more problems at journal that just flagged 400 papers

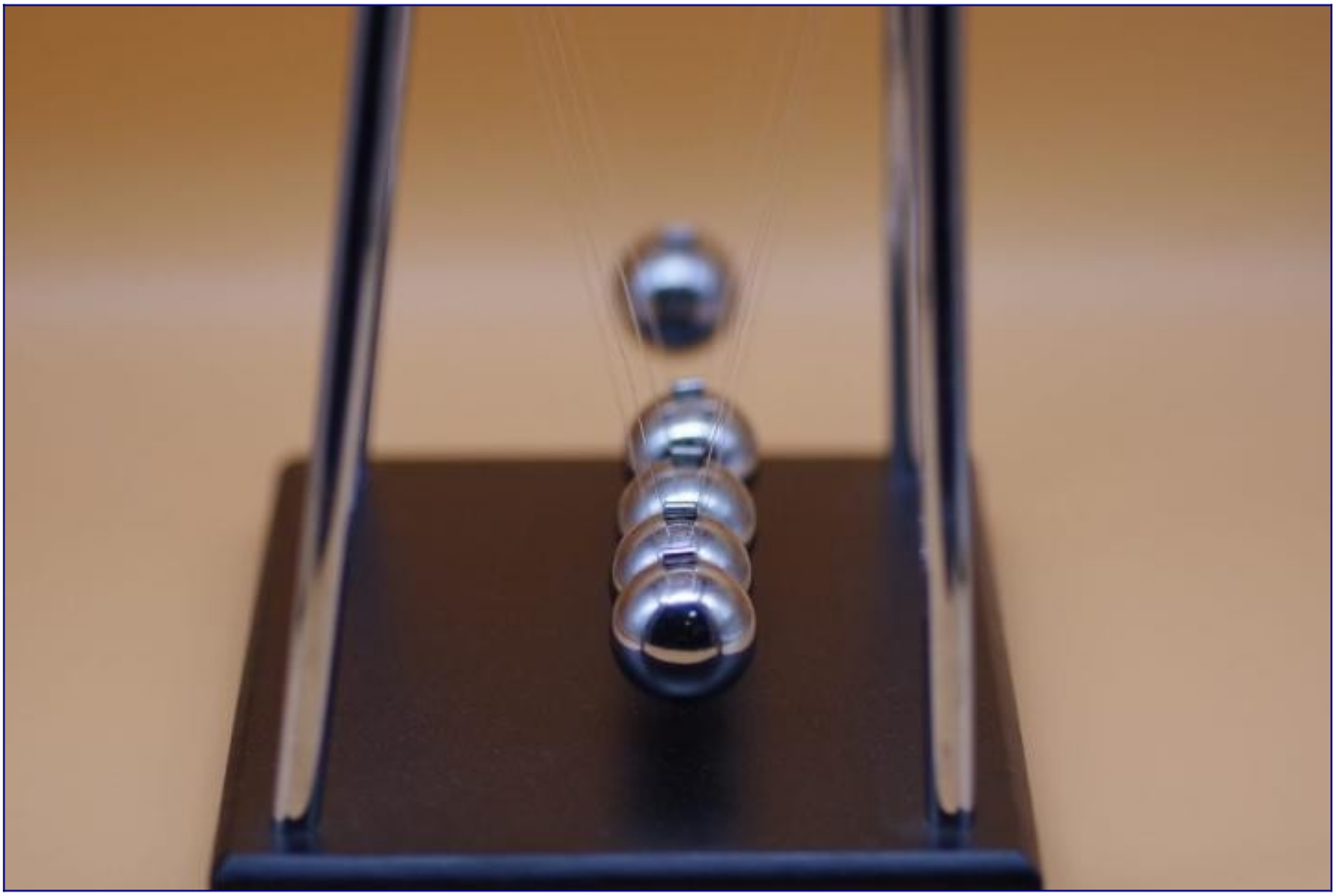
retractionwatch.com

Guillaume Cabanac What do subterranean insect provinces and motion to clamor have to do with microprocessors and microsystems? That’s an excellent question. Read on, dear reader. Earlier this month, …

Shared by Retraction Watch

[How to update a prognostic model of COVID-19 for the effect of vaccination](#)
discourse.datamethods.org
We have just put up uploaded a pre-print (https://www.medrxiv.org/content/10.1101/2021.07.19.21260759v1) that describes the derivation and validation of a prognostic model for Covid-19 severity in a …

[No More “What” Without the “Why”:](#)
medium.com
What a ride! Throughout the last months, I had the chance to enable various organizations and leaders leveraging their large databases with machine learning. I was particularly engaging with member…
Shared by Umair Akbar



[Commentary on Improving Precision and Power in Randomized Trials](#)
fharrell.com
This is a commentary on the paper by Benkeser, Díaz, Luedtke, Segal, Scharfstein, and Rosenblum

Shared by Frank Harrell

Guidelines for Covariate Adjustment in RCTs

discourse.datamethods.org

i’m susceptible to concise recommendations, thus i like this paper and tempted to share it. “Covariate adjustment for randomized controlled trials revisited” Although it is very difficult to make gen…

Shared by Frank Harrell

How bad statistical practices drive away the good ones

statsepi.substack.com

I am interested in research integrity and reproducibility. I believe that a lack of statistical expertise throughout the sciences is a substantial driver of problems in these areas (poor data practic…

Shared by Darren Dahly

Let's stop reviewing for venues that sell research - Joseph Paul Cohen PhD

josephpcohen.com

What’s the deal with how we review papers for venues (like conferences and journals) for free and then they go on to sell and restrict [...]

Shared by Retraction Watch

[Just 12 People Are Behind Most Vaccine Hoaxes On Social Media, Research Shows](#)
[npr.org](#)
[The majority of false claims about COVID-19 vaccines on social media trace back to just a handful of influential figures. So why don't the companies just shut them down?](#)
Shared by [Boback Ziaeian](#)

Violation of Proportional Odds is Not Fatal | Statistical Thinking
fharrell.com
Many researchers worry about violations of the proportional hazards assumption when comparing treatments in a randomized study. Besides the fact that this frequently makes them turn to a much worse a…
Shared by Frank Harrell
CQS Summer Institute | Center for Quantitative Sciences
vumc.org
CQS Summer Institute courses will be offered virtually via Zoom in 2021. Our CQS Summer Institute offers three great courses that will sharpen your quantitative research skills and give you methods a…
Shared by Vanderbilt Biostatistics

Full Episode List — Stats + Stories

statsandstories.net

View dozens of episodes about science, sports, politics and much more. From the podcast with the stats behind the stories and the stories behind the statistics.

Shared by Stats + Stories

Huge study supporting ivermectin as Covid treatment withdrawn over ethical concerns
theguardian.com
The preprint endorsing ivermectin as a coronavirus therapy has been widely cited, but independent researchers find glaring discrepancies in the data
Shared by Retraction Watch
How Many Numbers Exist? Infinity Proof Moves Math Closer to an Answer.
quantamagazine.org

For 50 years, mathematicians have believed that the total number of real numbers is unknowable. A new proof suggests otherwise.

Shared by Edward Tufte

Corporations, high-stakes biomedical research, and research misconduct: yes they can (and sometimes do)

academic.oup.com

Abstract. Science has long been vulnerable to research misconduct (RM). Biomedical sciences, with vast financial stakes, carry heightened temptations. However,

Shared by Retraction Watch

Tennessee's former top vaccine official: 'I am afraid for my state'

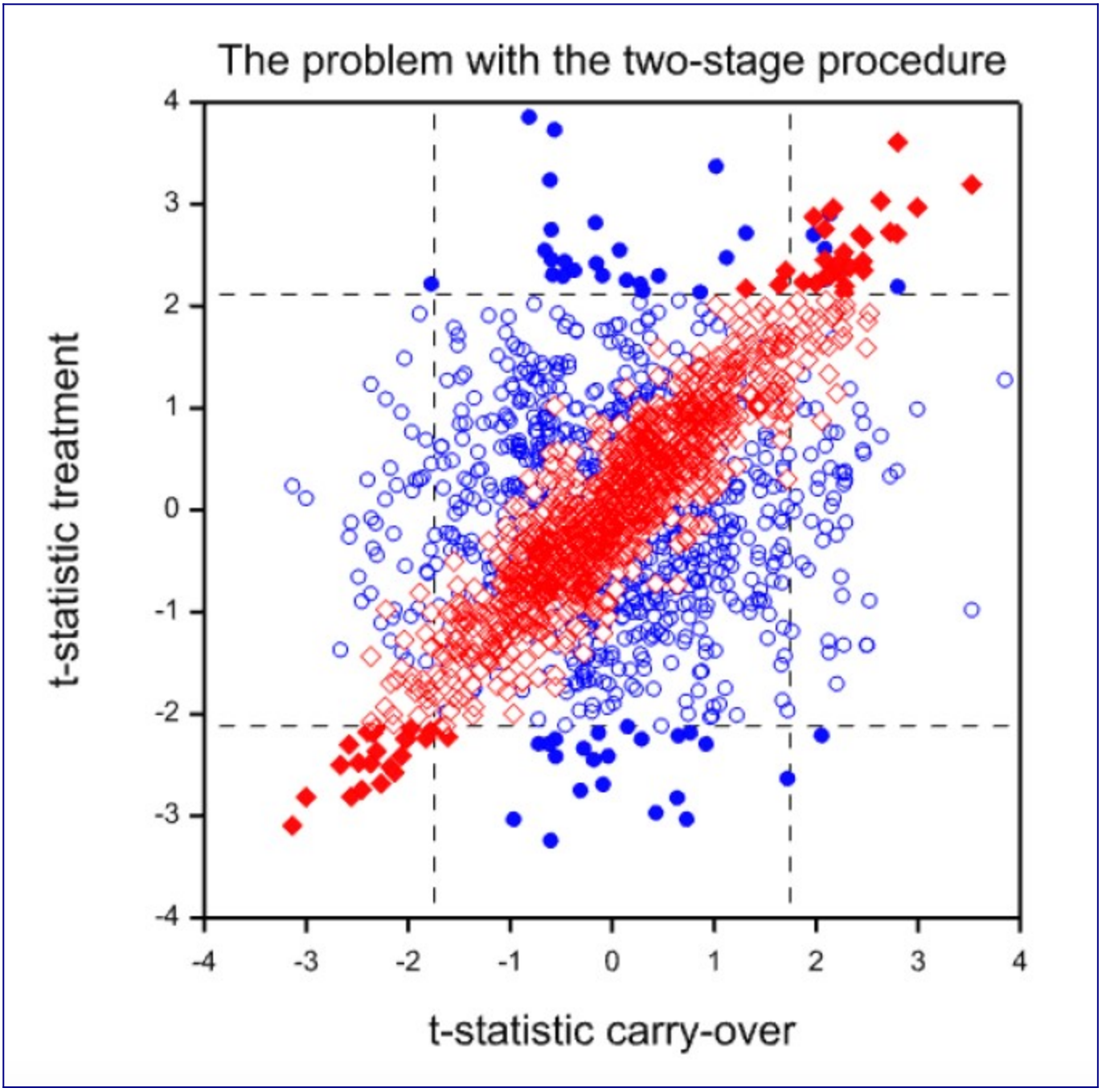
tennessean.com

Tennessee's top vaccination official, Dr. Michelle Fiscus, was fired on July 12 amid questions on youth COVID-19 vaccines. This is her explanation.

Shared by Mike Mike Mike Mike

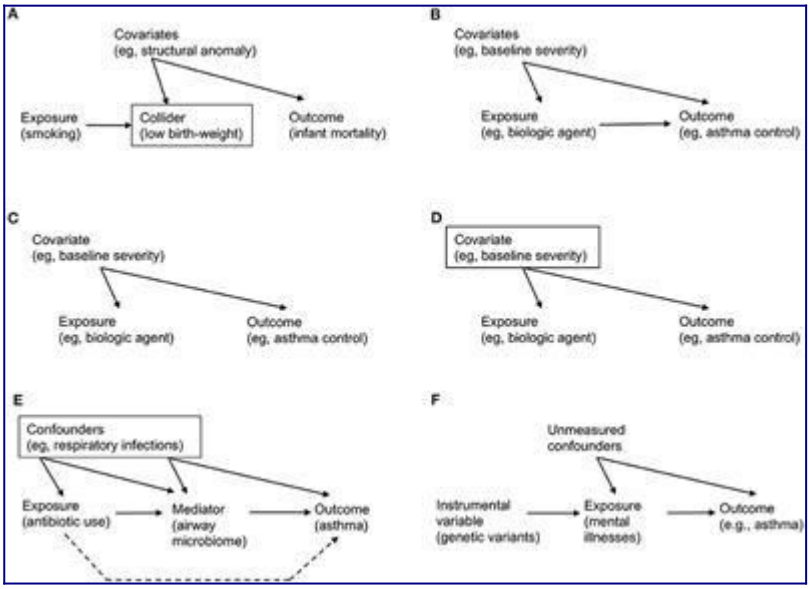
On fatally-flawed, valueless papers that journals refuse to retract « Statistical Modeling, Causal Inference, and Social Science
statmodeling.stat.columbia.edu
Commenter Carlos pointed us to this story (update here) of some scientists—Florin Moldoveanu, Richard Gill, and five others—all of whom seem to know what they're talking about and who are indignant t…
Shared by Andrew Gelman et al.
I critiqued my past papers on social media — here's what I learnt
nature.com
The systems of science must reward honesty about mistakes to speed progress.
Shared by Retraction Watch

S. Senn: Testing Times (Guest post)
errorstatistics.com
Stephen Senn Consultant Statistician Edinburgh, Scotland Testing Times Screening for attention There has been much comment on Twitter and other social media about testing for coronavirus and the rela…
Shared by Darren Dahly



She sent a letter pointing out problems with a published article, the reviewers agreed that her comments were valid, but the journal didn't publish h…
statmodeling.stat.columbia.edu
The journal in question is called The Economic Journal. To add insult to injury, the editor wrote the following when announcing they wouldn't publish the letter: My [the editor's] assessment is that …
Shared by Retraction Watch
Big Data, Data Science, and Causal Inference: A Primer for Clinicians
frontiersin.org
Clinicians handle a growing amount of clinical, biometric, and biomarker data. In this "big data" era, there is an emerging faith that the answer to all clinical and scientific questions reside in "b…

Shared by Kohei Hasegawa | 長谷川耕平



Ignoring mental health infrastructure will be a costly mistake
statnews.com
President Biden's ambitious plan makes no effort to address the pitiful state of the country's mental health infrastructure.

Shared by Allen Frances

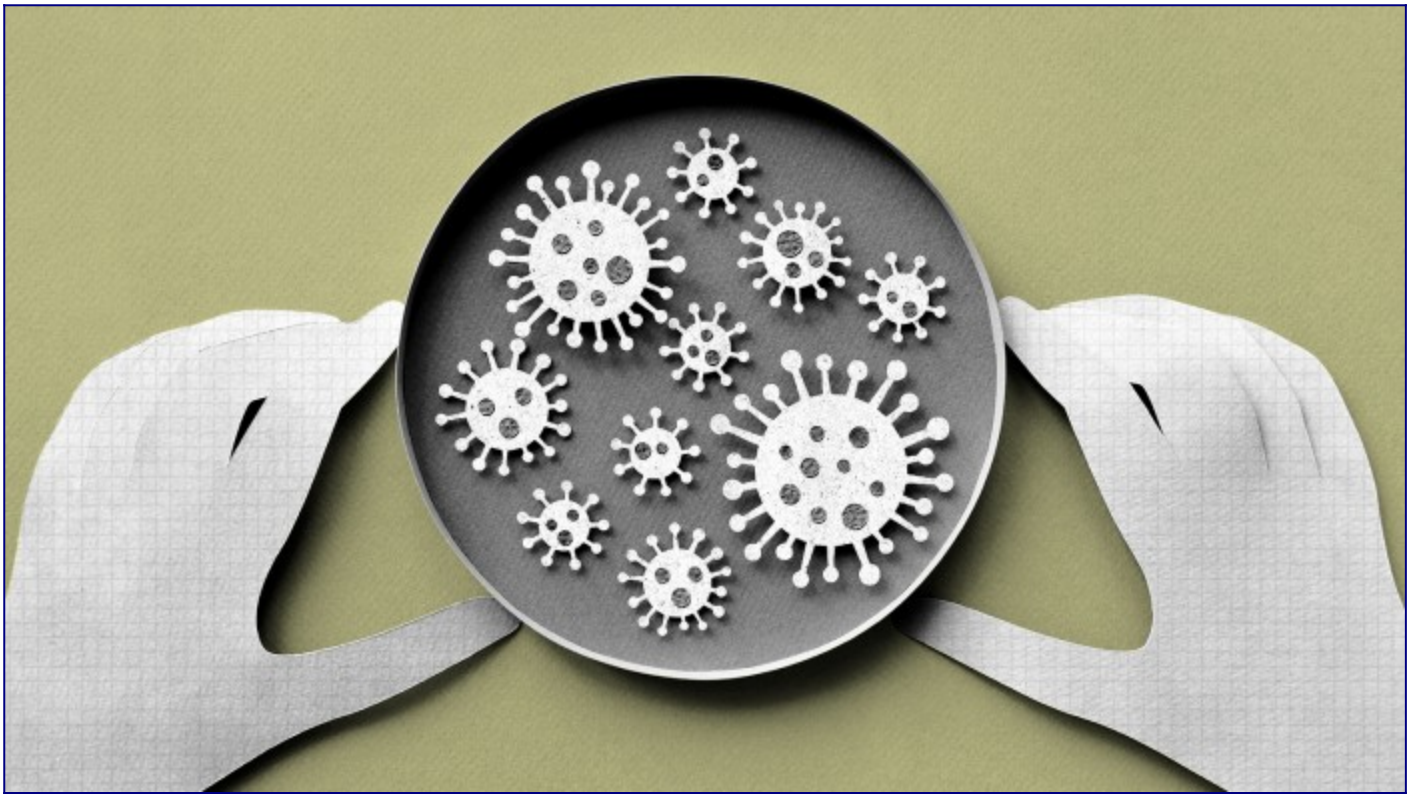
Top 10 Ideas in Statistics That Have Powered the AI Revolution « Statistical Modeling, Causal Inference, and Social Science
statmodeling.stat.columbia.edu

Aki and I put together this listsicle to accompany our recent paper on the most important statistical ideas of the top 50 years. Kim Martineau at Columbia, who suggested making this list, also had th…
Shared by Andrew Gelman et al.

Time to assume that health research is fraudulent until proven otherwise? - The BMJ
blogs.bmj.com
Health research is based on trust. Health professionals and journal editors reading the results of a clinical trial assume that the trial happened and that the results were honestly supported. [...]M…
Shared by Retraction Watch

Landmark research integrity survey finds questionable practices are surprisingly common
sciencemag.org
Largest-ever survey finds one in 12 Dutch scientists have recently committed fraud
Shared by Boback Ziaeian

How Science Moved Beyond Peer Review During The Pandemic
fivethirtyeight.com
And what scientists learned they still needed it for.
Shared by Retraction Watch

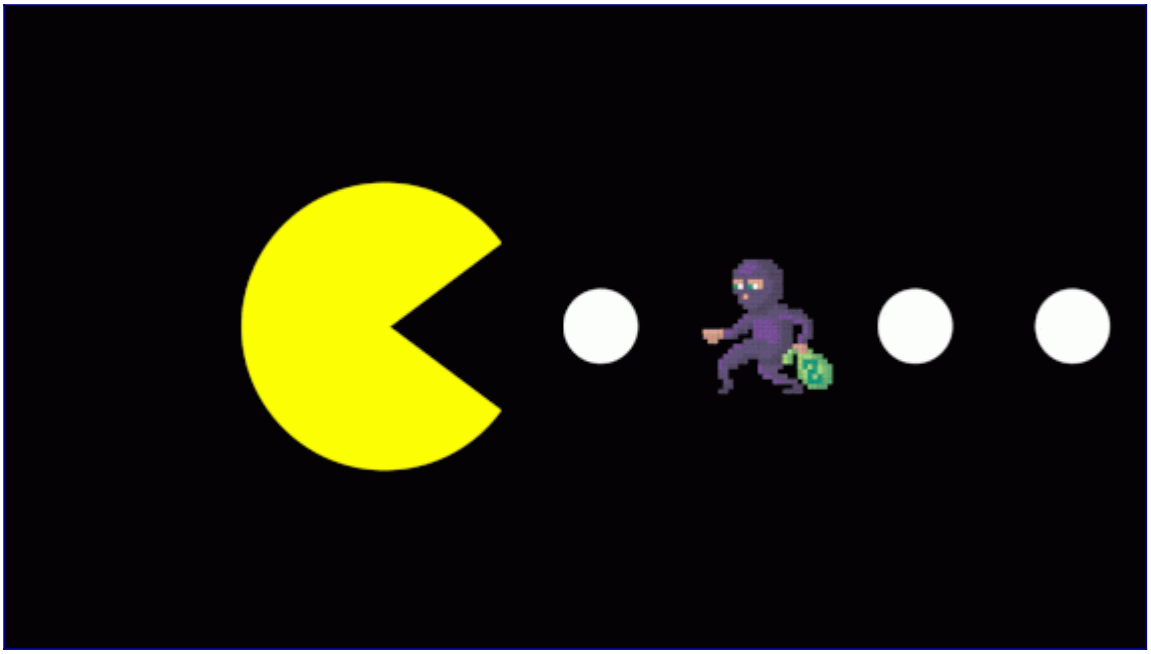


Why some biologists and ecologists think social media is a risk to humanity
vox.com
One challenge is how little we know about the dangers.
Shared by C. Michael Gibson MD

Why Are Gamers So Much Better Than Scientists at Catching Fraud?
theatlantic.com

A pair of recent cheating scandals—one in the “speedrunning” community of gamers, and one in medical research—call attention to an alarming contrast.

Shared by Retraction Watch



Observational studies and the difficult quest for causality: lessons from vaccine effectiveness and impact studies

ncbi.nlm.nih.gov

Although randomized placebo-controlled trials (RCT) are critical to establish efficacy of vaccines at the time of licensure, important remaining questions about vaccine effectiveness (VE)—used here t…

Shared by Dr Ellie Murray

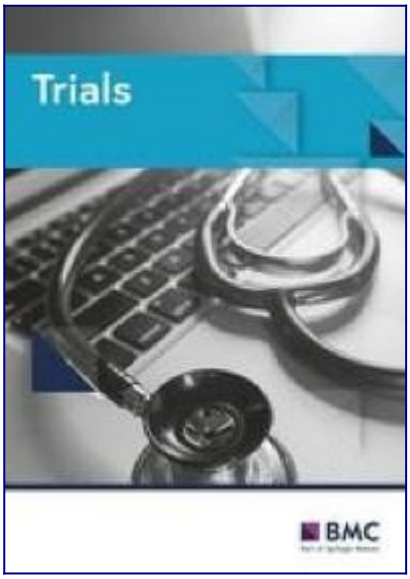
The continuing misrepresentations coming from the University of California sleep researcher and publicized by Ted and NPR « Statistical Modeling, Cau…

statmodeling.stat.columbia.edu

Just when I had put the “Matthew Walker fake news” into a comfortable place of oblivion, NPR sends me this suggested story. How disappointing that NPR’s fact check is no better than other media outle…

Shared by Andrew Gelman et al.

Comparisons against baseline within randomised groups are often used and can be highly misleading
trialsjournal.biomedcentral.com
In randomised trials, rather than comparing randomised groups directly some researchers carry out a significance test comparing a baseline with a final measurement separately in each group. We give s…
Shared by Darren Dahly



The Case for Explainable AI (XAI)
infoq.com
Black box algorithms have precipitated high-profile controversies arising from the inability to understand their inner workings. Explainable AI (XAI) seeks to provide more transparency in this regard.
Shared by Mehmet Kesim

Hidden in Plain Sight — Reconsidering the Use of Race Correction in Clinical Algorithms
nejm.org
Hidden in Plain Sight Diagnostic algorithms and practice guidelines that adjust or “correct” their outputs on the basis of a patient’s race or ethnicity guide decisions in ways that may direct more...
Shared by MayoClinicPsych

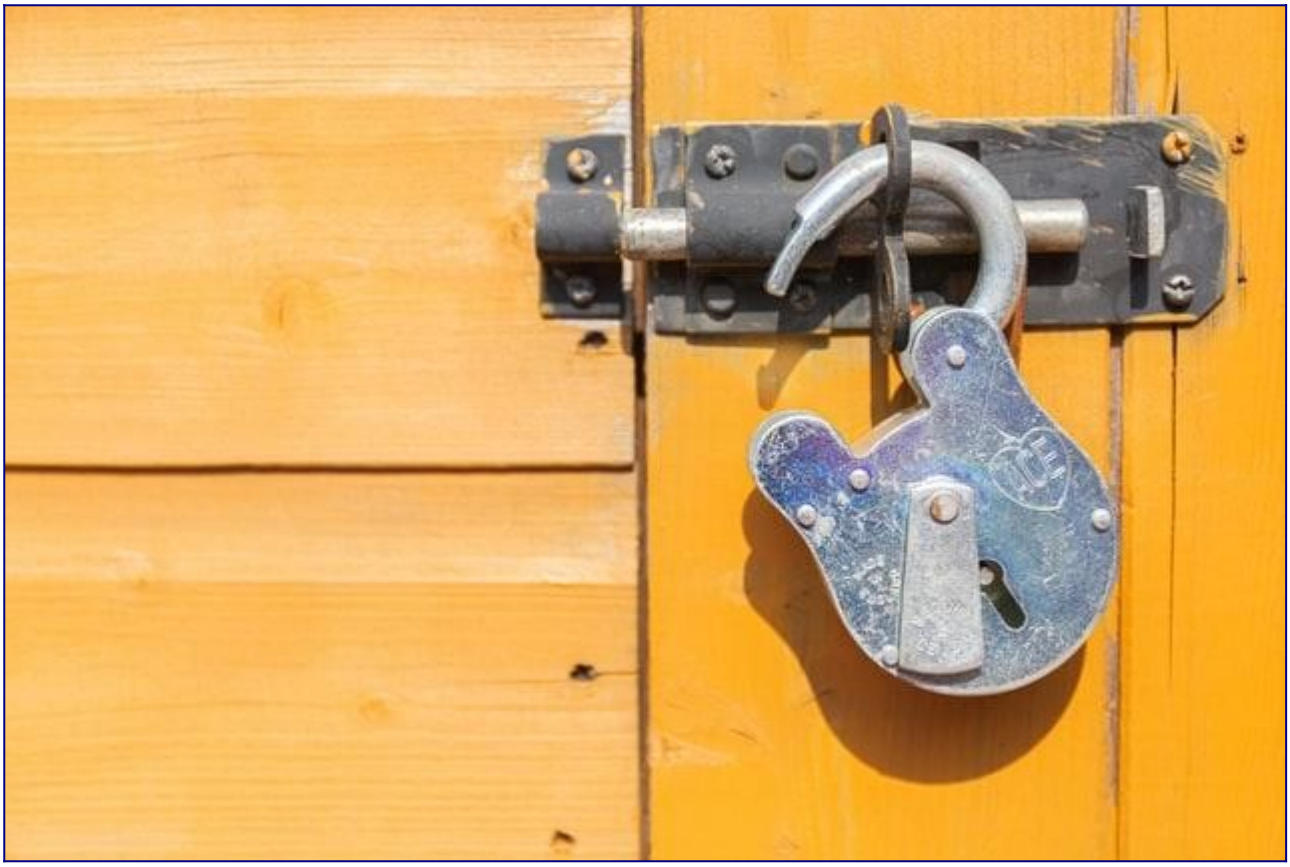
Not being able to say why you see a 2 doesn’t excuse your uninterpretable model « Statistical Modeling, Causal Inference, and Social Science
statmodeling.stat.columbia.edu
This is Jessica, but today I’m blogging about a blog post on interpretable machine learning that co-blogger Keith wrote for his day job and shared with me. I agree with multiple observations he makes…
Shared by Andrew Gelman et al.

Review papers and the creative destruction of the research literature
blogs.lse.ac.uk
Review papers play a significant role in curating the scholarly record. Drawing on a study of close to six million research articles, Peter McMahan, shows how review papers not only focus and shift…
Shared by Retraction Watch

Calculating the sample size required for developing a clinical prediction model
bmj.com
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…
Shared by TRIPODStatement

[Incorrect Covariate Adjustment May Be More Correct than Adjusted Marginal Estimates | Statistical Thinking]
fharrell.com
[This article provides a demonstration that the perceived non-robustness of nonlinear models for covariate adjustment in randomized trials may be less of an issue than the non-transportability of marg…]
Shared by Frank Harrell

[Open access at no cost? Just ditch academic journals]
researchwhisperer.org
[Abel Polese is a researcher, trainer, writer, manager and fundraiser. He is the author of “The SCOPUS Diaries and the (il)logics of Academic Survival: A Short Guide to Design Your Own Strategy and Su…]
Shared by Retraction Watch



Avoiding One-Number Summaries of Treatment Effects for RCTs with Binary Outcomes | Statistical Thinking
fharrell.com
This article presents an argument that for RCTs with a binary outcome the primary result should be a distribution and not any single number summary. The GUSTO-I study is used to exemplify risk differ…
Shared by Frank Harrell

The Pandemic Led To The Biggest Drop In U.S. Life Expectancy Since WW II, Study Finds
npr.org
Life expectancy fell by nearly 2 years between 2018 and 2020, largely due to COVID-19. Declines were most pronounced among minority groups, including Black and Hispanic people.
Shared by Boback Ziaeian

Correlation without a cause: an epidemiological odyssey
academic.oup.com
Background. In the 1980s debate intensified over whether there was a protective effect of high-density lipoprotein cholesterol (HDL-C) or an adverse effect of t
Shared by george davey smith

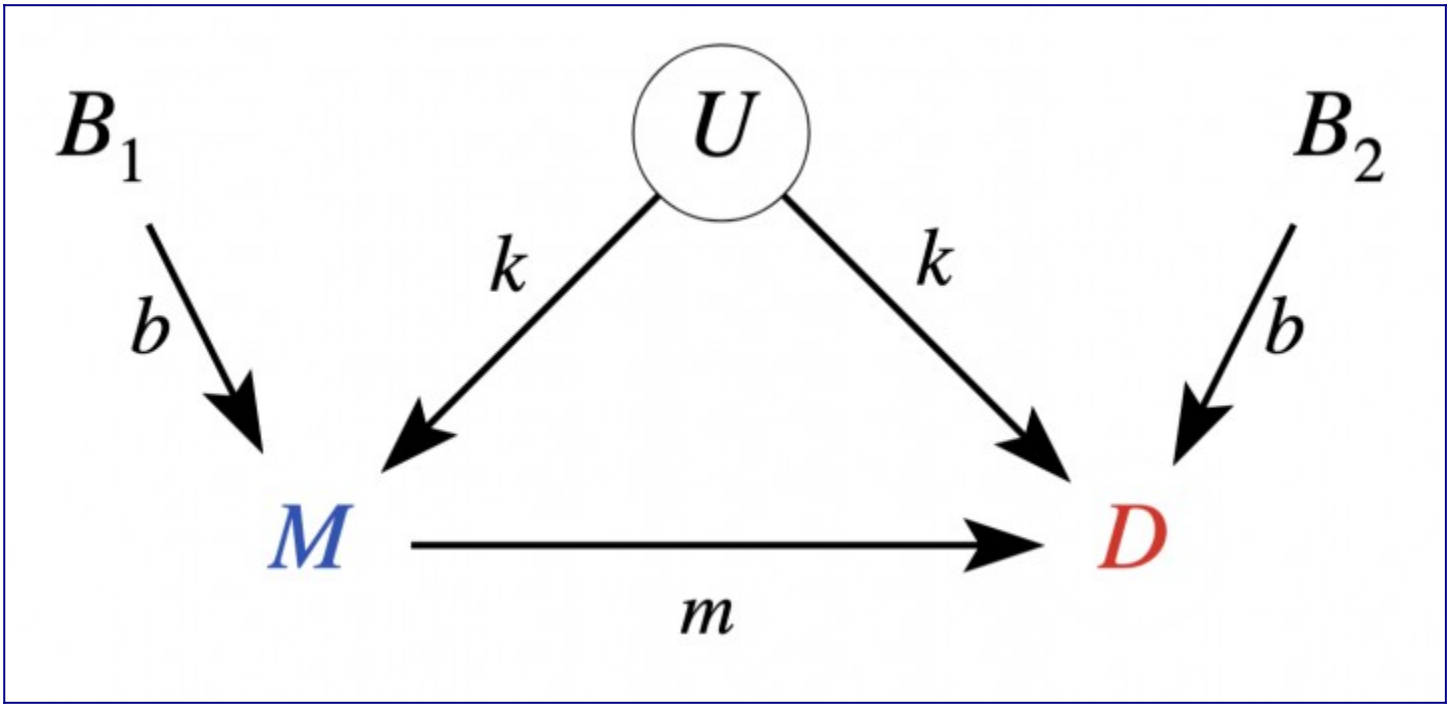
Opinion | Where Did the Coronavirus Come From? What We Already Know Is Troubling.
nytimes.com
Even if the coronavirus did not emerge from a lab, the groundwork for a potential disaster had been laid for years, and learning its lessons is essential to preventing others.
Shared by Boback Ziaeian

Guernsey McPearson's Dictionary
senns.uk
The definitions here are all original and were coined by me at the Elixir Laboratories of Panacea Pharmaceuticals. They seem to have spread rather like a virus and now often appear unattributed elsew…
Shared by shriyamite

Quality shines when scientists use publishing tactic known as registered reports, study finds
sciencemag.org
Papers accepted by journals before results are known rate higher on rigor than standard studies
Shared by Brian Nosek

How a Sharp-Eyed Scientist Became Biology's Image Detective
newyorker.com
Using just her eyes and memory, Elisabeth Bik has single-handedly identified thousands of studies containing potentially doctored scientific images.
Shared by Retraction Watch

Regression, Fire, and Dangerous Things (2/3)
elevanth.org
I'll worry about the singularity when AI isn't confused about cinnamon rolls Thinking Like a Graph
The brilliance of artificial intelligence is that it is much better than us, its creat…
Shared by Richard McElreath



External Validation of a Widely Implemented Sepsis Prediction Model in Hospitalized Patients

jamanetwork.com

This cohort study externally validates the Epic Sepsis Model in the prediction of sepsis and evaluates its potential clinical impact compared with usual care.

Shared by Cecile Janssens

Table 1. Characteristics of Patients

Characteristic	Hospitalizations, No. (%)[a] Overall (N = 38 455)	No sepsis (n = 35 903)	Sepsis (n = 2552)	*P* value[b]
Age, median (IQR), y	56 (35-69)	55 (34-68)	63 (51-73)	<.001
Sex				
Female	21 904 (57)	20 781 (58)	1123 (44)	<.001
Male	16 551 (43)	15 122 (42)	1429 (56)	
Race/ethnicity				
Black	4638 (12)	4303 (12)	335 (13)	<.001
White	30 812 (80)	28 742 (80)	2070 (81)	
Other[c]	2526 (7)	2405 (7)	121 (5)	
Unknown	479 (1)	453 (1)	26 (1)	
Comorbidities				
Congestive heart failure	7062 (18)	6189 (17)	873 (34)	<.001
Chronic obstructive pulmonary disease	7513 (20)	6763 (19)	750 (29)	<.001
Chronic kidney disease	7313 (19)	6332 (18)	981 (38)	<.001
Depression	8140 (21)	7478 (21)	662 (26)	<.001
Type 1 and 2 diabetes	8267 (22)	7325 (20)	942 (37)	<.001
Hypertension	18 066 (47)	16 369 (46)	1697 (67)	<.001
Obesity	8422 (22)	7799 (22)	623 (24)	.002
Liver disease	3213 (8)	2611 (7)	602 (24)	<.001
Solid tumor, nonmetastatic	5973 (16)	5473 (15)	500 (20)	<.001
Metastatic cancer	3255 (9)	2950 (8)	305 (12)	<.001

Abbreviation: IQR, interquartile range.

[a] For some patients, there were repeated hospitalizations.

[b] Statistical tests performed: Wilcoxon rank sum test and χ^2 test of independence.

[c] Includes American Indian or Alaska Native, Asian, and Native Hawaiian and other Pacific Islander.

Summer Conferences!

rviews.rstudio.com

An R community blog edited by RStudio

Shared by RStudio

[Pragmatism should Not be a Substitute for Statistical Literacy, a Commentary on Albers, Kiers, and Van Ravenzwaaij (2018)]
online.ucpress.edu
[Based on the observation that frequentist confidence intervals and Bayesian credible intervals sometimes happen to have the same numerical boundaries (under very specific conditions), Albers et al. (…]
Shared by Frank Harrell

[My Journey From Frequentist to Bayesian Statistics | Statistical Thinking]
fharrell.com
[The difference between Bayesian and frequentist inference in a nutshell: With Bayes you start with a prior distribution for θ and given your data make an inference about the θ-driven process generati…]
Shared by Frank Harrell

[On the Convergence of Epidemiology, Biostatistics, and Data Science]
hdsr.mitpress.mit.edu
[Epidemiology, biostatistics, and data science are broad disciplines that incorporate a variety of substantive areas. Common among them is a focus on quantitative approaches for solving intricate prob…]
Shared by Jason H. Moore, PhD

Wow, just wow. If you think Psychological Science was bad in the 2010-2015 era, you can't imagine how bad it was back in 1999 « Statistical Modeling,…
statmodeling.stat.columbia.edu
Shane Frederick points us to this article from 1999, "Stereotype susceptibility: Identity salience and shifts in quantitative performance," about which he writes: This is one of the worst papers ever…
Shared by Andrew Gelman et al.

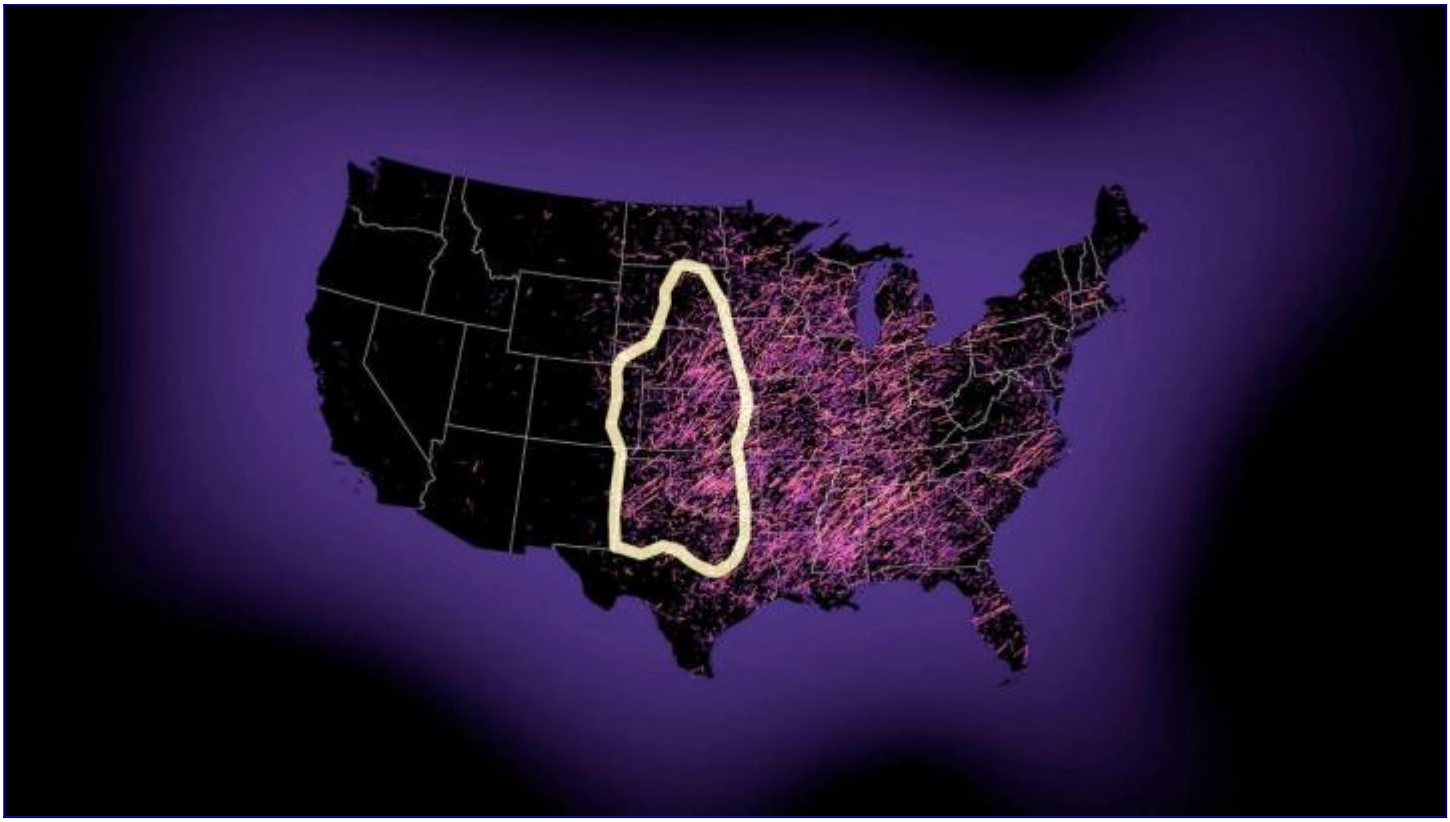
The FDA Is Failing the American People
medpagetoday.com
The agency's recent decision to approve aducanumab offers more evidence of a broken system
Shared by Vinay Prasad MD MPH

Evidence-based medicine eats itself in real time « Statistical Modeling, Causal Inference, and Social Science
statmodeling.stat.columbia.edu
This has just appeared in BMJ Evidence Based Medicine. It addresses the controversial question of whether lowering LDL using statins leads to reduced mortality and CVD rates. The researchers pull tog…
Shared by Andrew Gelman et al.

Guttman points out another problem with null hypothesis significance testing: It falls apart when considering replications. « Statistical Modeling, C…
statmodeling.stat.columbia.edu
Re-reading a classic from Louis Guttman, What is not what in statistics, I saw his “Problem 2” with new eyes given the modern replication debate: Both estimation and the testing of hypotheses have us…
Shared by Andrew Gelman et al.

Special vs. Markov models for analyzing within-patient variability
discourse.datamethods.org
This excellent paper by German et al. presents a computationally and statistically efficient way to jointly analyze the mean and the variance as a function of covariates. Their method is named WiSER …

'Tornado Alley' is expanding: Southern states see more twisters now than ever before
usatoday.com
We talked with scientists and examined years of data to explain why millions across the South are at risk
Shared by Mike Mike Mike Mike



Unadjusted Odds Ratios are Conditional | Statistical Thinking
fharrell.com
This article discusses issues with unadjusted effect ratios such as odds ratios and hazard ratios, showing a simple example of non-generalizability of unadjusted odds ratios.
Shared by Frank Harrell
NIH should boost rigor of animal studies with stronger statistics, pilot studies, experts say
sciencemag.org
Working group recommends a range of steps to bolster reproducibility
Shared by Boback Ziaeian
Log Log Regression
darrendahly.github.io
Principal Statistician | Senior Lecturer

Shared by Darren Dahly

The Brier score does not evaluate the clinical utility of diagnostic tests or prediction models
diagnprognres.biomedcentral.com
A variety of statistics have been proposed as tools to help investigators assess the value of diagnostic tests or prediction models. The Brier score has been recommended on the grounds that it is a p…

Shared by Andrew Vickers

Regression, Fire, and Dangerous Things (1/3)
elevanth.org
Okay so correlation does not imply causation. What now? Three Disappointing Things It isn't my job to disappoint people, but I'm good at it. Other researchers are out there writing book…

Shared by Maarten van Smeden

Concerns about Marseille’s IHUMI/AMU papers – Part 1
scienceintegritydigest.com
Last March, I shared my concerns about a paper from the IHU-Méditerranée Infection (IHUMI)/ Aix-Marseille University (AMU) claiming that Hydroxychloroquine in combination with Azithromycin could redu…
Shared by Retraction Watch

Figure 1 and Figure 2 look remarkably similar. Based on the legends, Figure 2 maybe should have shown something else.

Could the authors please check?

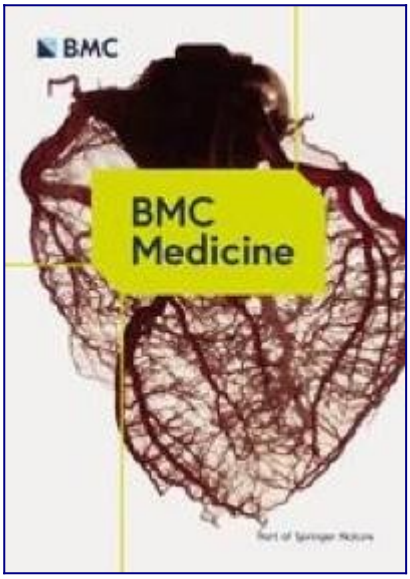
They might also want to check the spelling of the affiliated city, "*Marrseille*".

Raoult et al., Am J Clin Pathol (1994)
DOI: 10.1093/ajcp/101.3.318

FIG. 1. Immunofluorescent staining of *C burnetii* in the aortic valve of a patient with chronic Q fever endocarditis using Cox1D8 monoclonal antibody (magnification × 400).

FIG. 2. Immunofluorescent detection of *C burnetii* in shell vial using Cox1D8 monoclonal antibody (6 days incubation, magnification × 1000).

Recommended Stat Articles
hbiostat.org
Recommended articles in stat journals, primarily published in the past 60d
Shared by Frank Harrell
Meet a sleuth whose work has led to the identification of hundreds of fraudulent papers
retractionwatch.com
John Loadsman Last month, Retraction Watch reported on the case of Hironobo Ueshima, an anesthesiology researcher found guilty of misconduct in more than 140 papers. A journal editor, John Loadsman, …
Shared by Retraction Watch
Calibration: the Achilles heel of predictive analytics
bmcmedicine.biomedcentral.com
The assessment of calibration performance of risk prediction models based on regression or more flexible machine learning algorithms receives little attention. Herein, we argue that this needs to cha…
Shared by Maarten van Smeden



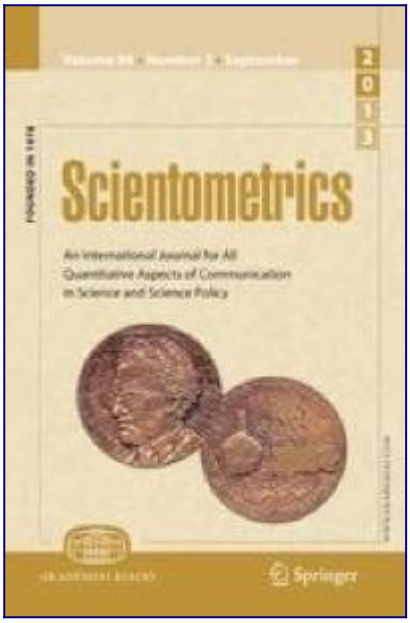
[Clinical Trial Outcomes Interrupted by Other Outcomes]
discourse.datamethods.org
[This is a discussion of how best to analyze and to present clinical outcomes when an outcome of major interest can be interrupted by another outcome making the first outcome unobservable. The most co…]
Shared by Frank Harrell

[Reporting on scientific failures, holding the science community accountable]
journalistsresource.org
[Scholars offer tips to help news outlets report accurately on scientists' work and hold the scientific community accountable for its failures.]
Shared by Retraction Watch

[The Absurdity of Peer Review]
elemental.medium.com
[I was reading my umpteenth news story about Covid-19 science, a story about the latest research into how to make indoor spaces safe from infection, about whether cleaning surfaces or changing the air…]
Shared by [Retraction Watch]

[Upcoming Webinar - Incorporating R into your Clinical Legacy Workflows]
blog.rstudio.com
When it comes to adopting open source tools into clinical processes, there are many questions that have to be addressed, such as how do we migrate legacy workflows over to new programming languages? …
Shared by RStudio

[Research ethics: a profile of retractions from world class universities]
link.springer.com
This study aims to profile the scientific retractions published in journals indexed in the Web of Science database from 2010 to 2019, from researchers at t
Shared by Retraction Watch



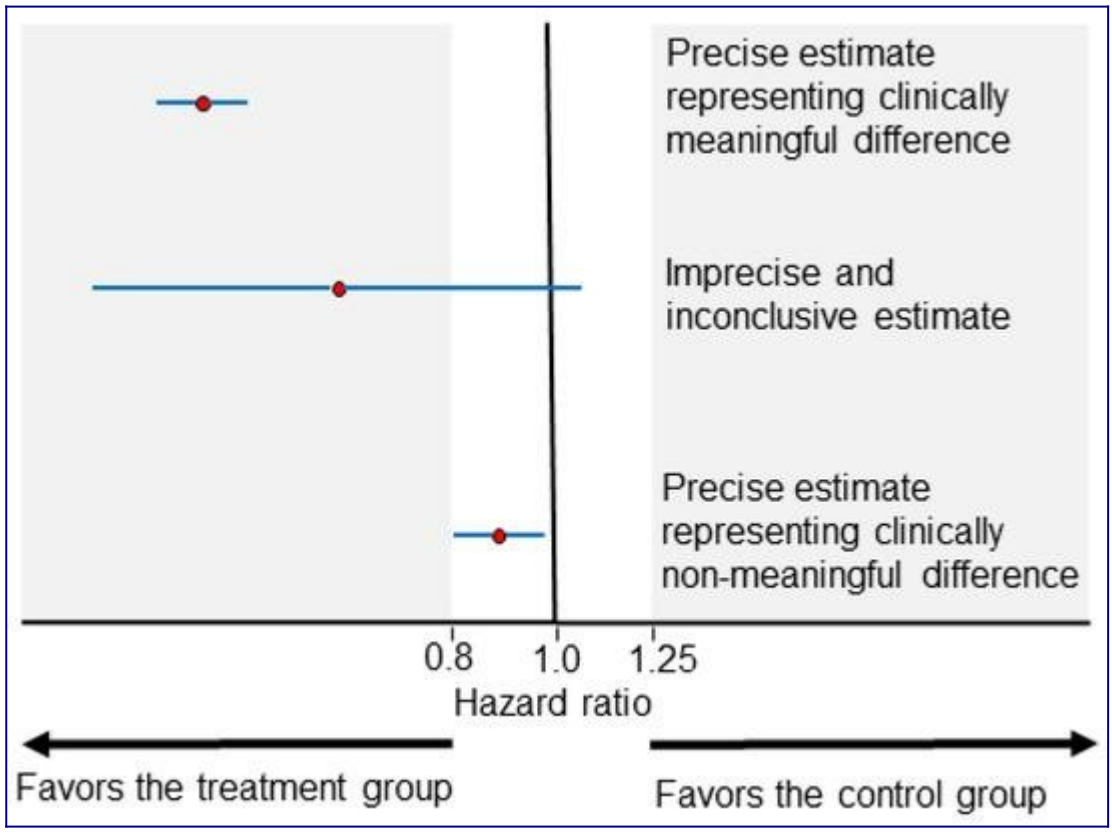
Reference Collection to push back against "Common Statistical Myths"
discourse.datamethods.org
Note: This topic is a wiki, meaning that this main body of the topic can be edited by others. Use the Reply button only to post questions or comments about material contained in the body, or to sugge…
Shared by Michael Harhay

Should one derive risk difference from the odds ratio?
discourse.datamethods.org
This paper (Doi et al. 2020) came out last year heavily advocating for the odds ratio (OR). https://www.jclinepi.com/article/S0895-4356(20)31171-9/fulltext I am aware of OR's advantages over risk rat…

Shared by Frank Harrell

Making Patient-Specific Treatment Decisions Using Prognostic Variables and Utilities of Clinical Outcomes
mdpi.com
We argue that well-informed patient-specific decision-making may be carried out as three consecutive tasks: (1) estimating key parameters of a statistical model, (2) using prognostic information to c…

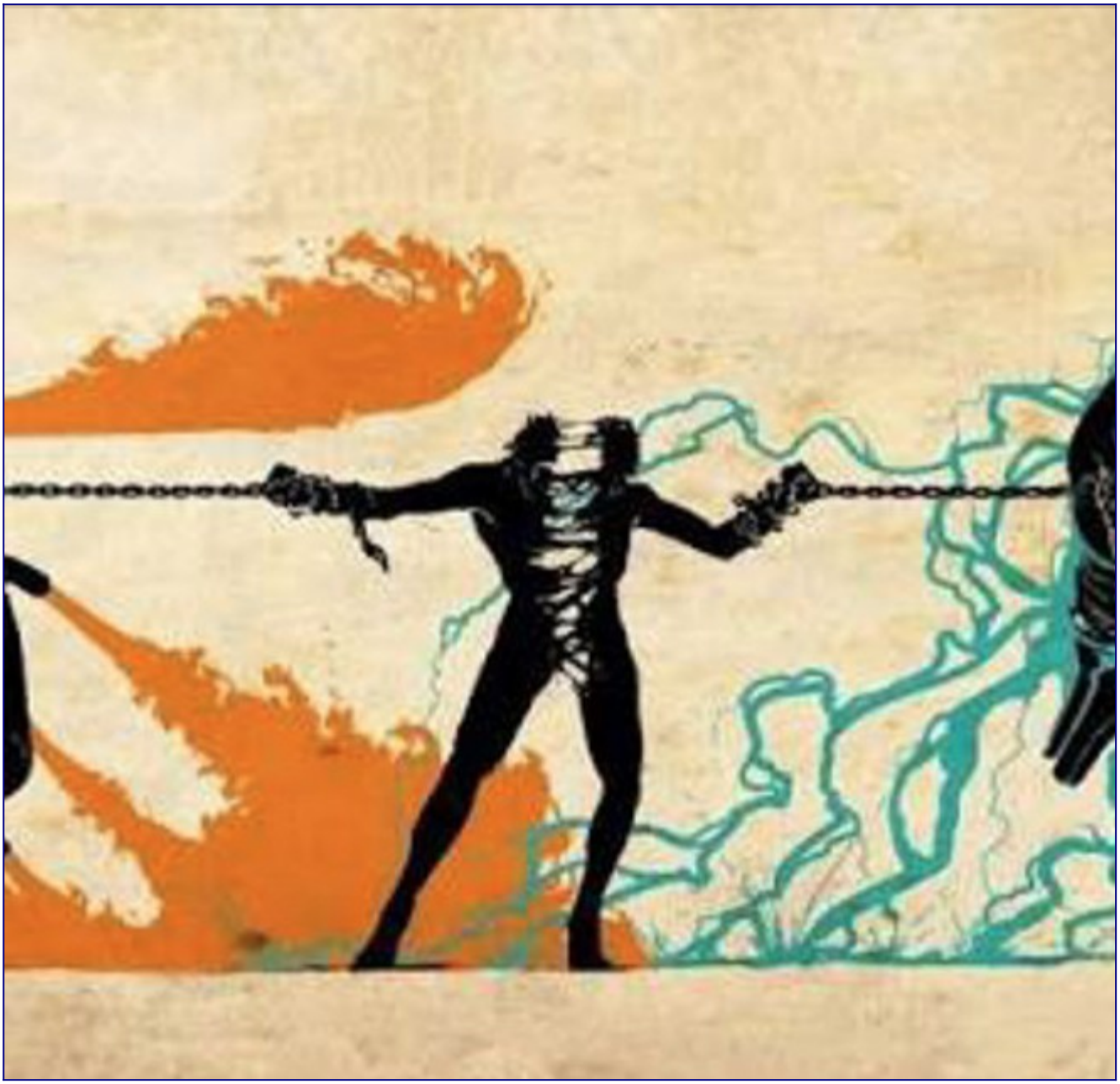
Shared by Darren Dahly



Modern modelling techniques are data hungry: a simulation study for predicting dichotomous endpoints
bmcmedresmethodol.biomedcentral.com
Modern modelling techniques may potentially provide more accurate predictions of binary outcomes than classical techniques. We aimed to study the predictive performance of different modelling techniq…

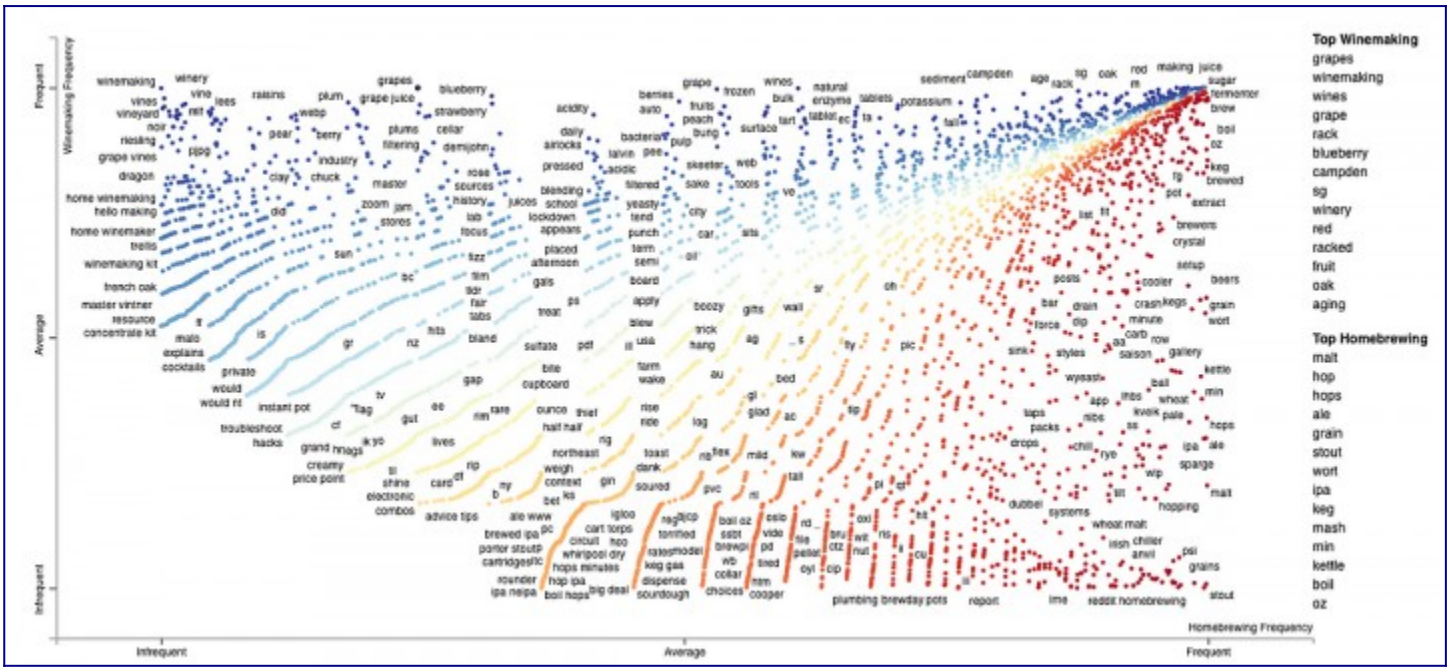
Shared by Ewout Steyerberg

Bayesian philosophers vs Bayesian statisticians: Remarks on Jon Williamson
errorstatistics.com
While I would agree that there are differences between Bayesian statisticians and Bayesian philosophers, those differences don’t line up with the ones drawn by Jon Williamson in his presentation to o…
Shared by Stephen John Senn



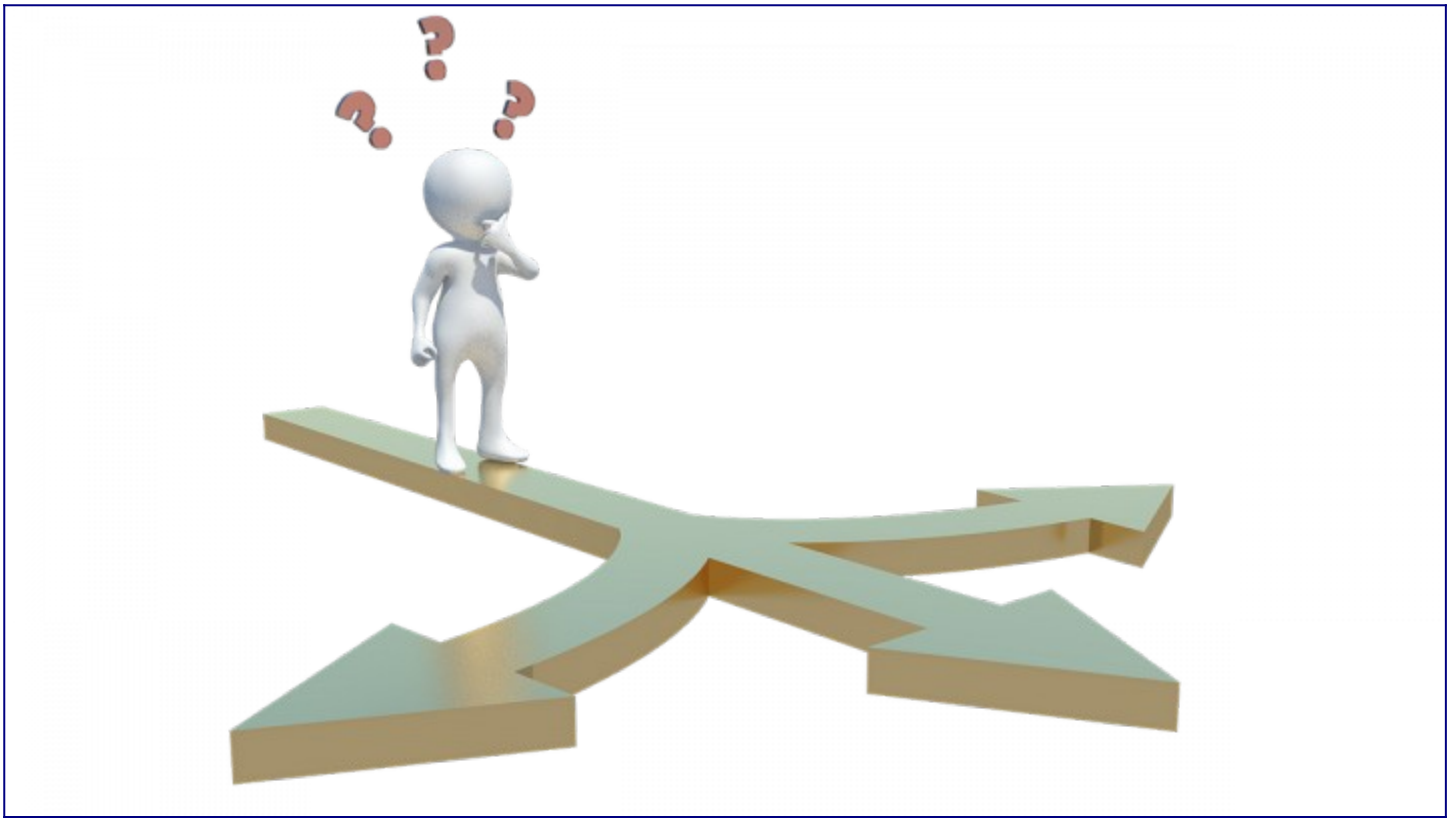
Is Medicine Mesmerized by Machine Learning? | Statistical Thinking
fharrell.com
Deep learning and other forms of machine learning are getting a lot of press in medicine. The reality doesn't match the hype, and interpretable statistical models still have a lot to offer.
Shared by Gary Lyman
Smell you later
crystalprisonzone.blogspot.com
A few weeks ago, I started a new job as a data scientist at the logistics division of a pharmaceutical company. This was primarily motivated...
Editor's note: This excellent article address the replication crisis in science and how academics are often motivated to do bad research, then contrasts this with an industrial setting.
Shared by Darren Dahly

[Interpreting Scattertext: A seductive tool for plotting text](#)
towardsdatascience.com
Scattertext Visualizations for Natural Language Processing. Jason Kessler, created a Spacy tool for plot, analyzing and interpreting text data with two classes.
Shared by Edward Tufte



[European regulators urged to crack down on missing clinical trial results](#)
bmj.com
Transparency campaigners have urged national medicines regulators across Europe to take action to reduce the high number of unreported clinical trials. An open letter1 to heads of medicines agencies …
Shared by Retraction Watch

[Interpretable or Accurate? Why Not Both?]
towardsdatascience.com
As summed up by Miller, interpretability refers to the degree to which a human can understand the cause of a decision. A common notion in the machine learning community is that a trade-off exists…
Editor's note: the proposed method is a statistical model, not strictly machine learning, and is related to multivariate adaptive regression splines (MARS).
Shared by Towards Data Science



Be Careful When Interpreting Predictive Models in Search of Causal Insights
towardsdatascience.com
A careful exploration of the pitfalls of trying to extract causal insights from modern predictive machine learning models.
Shared by ktaku

Regression discontinuity designs in healthcare research
bmj.com
Clinical decisions are often driven by decision rules premised around specific thresholds. Specific laboratory measurements, dates, or policy eligibility criteria create cut-offs at which people beco…

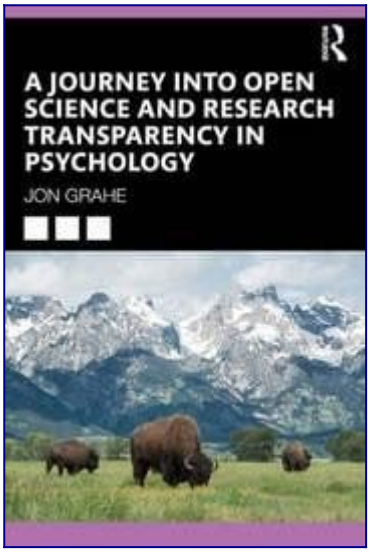
Shared by Emilie Courtin

Recommended Statistics Articles
hbiostat.org
Articles published in stat journals, mostly from the past 60d
Shared by Frank Harrell

World expert in scientific misconduct faces legal action for challenging integrity of hydroxychloroquine study
theguardian.com
Australian and international scientists publish open letter defending Dr Elisabeth Bik and calling for science whistleblowers to be protected
Shared by Retraction Watch

Does the log rank test assume proportional hazards?
thestatsgeek.com
A student asked me recently whether the log rank test for time to event data assumes that the hazard ratio between the two groups is constant over time, as is assumed in Cox's famous proportional haz…
Editors's note: Yes the logrank test assumes PH. To say that neither logrank nor Cox assume PH under the null is not that helpful, as the null is one point out of an infinity of treatment effects.
Shared by Andrew Althouse

A Journey into Open Science and Research Transparency in Psychology
routledge.com
A Journey into Open Science and Research Transparency in Psychology introduces the open science movement from psychology through a narrative that integrates son
Shared by EJ Wagenmakers

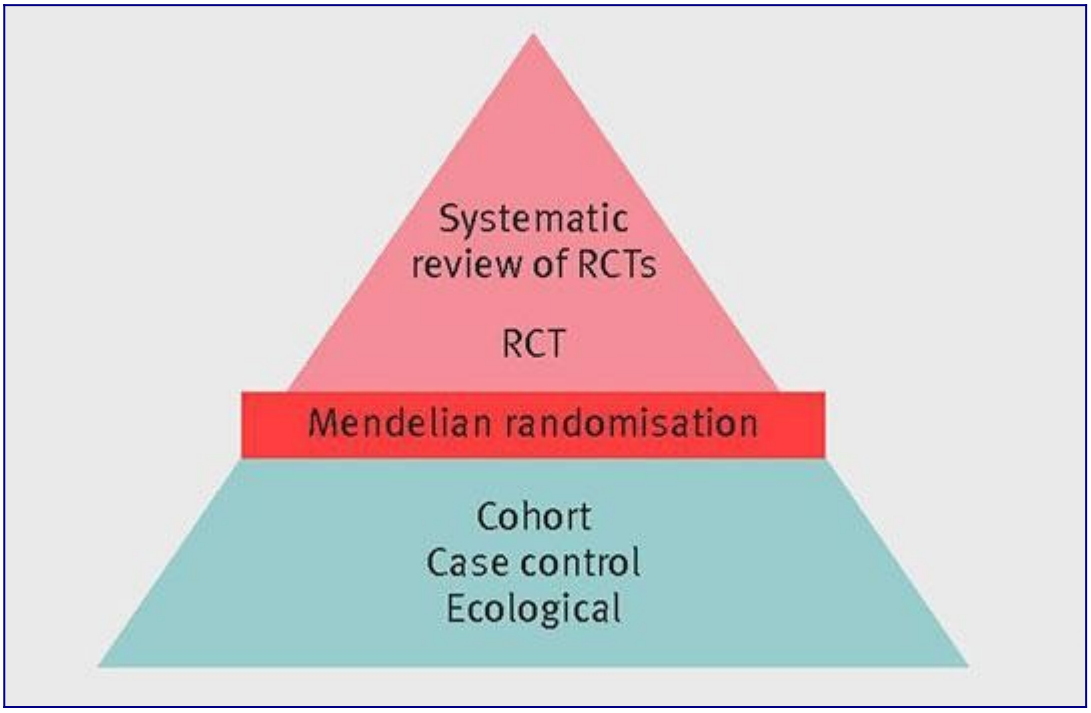


Absence of evidence is not evidence of absence
pubmed.ncbi.nlm.nih.gov
Absence of evidence is not evidence of absence
Editor's note: This is a classic must-read article for anyone interested in evidence from data
Shared by Darren Dahly

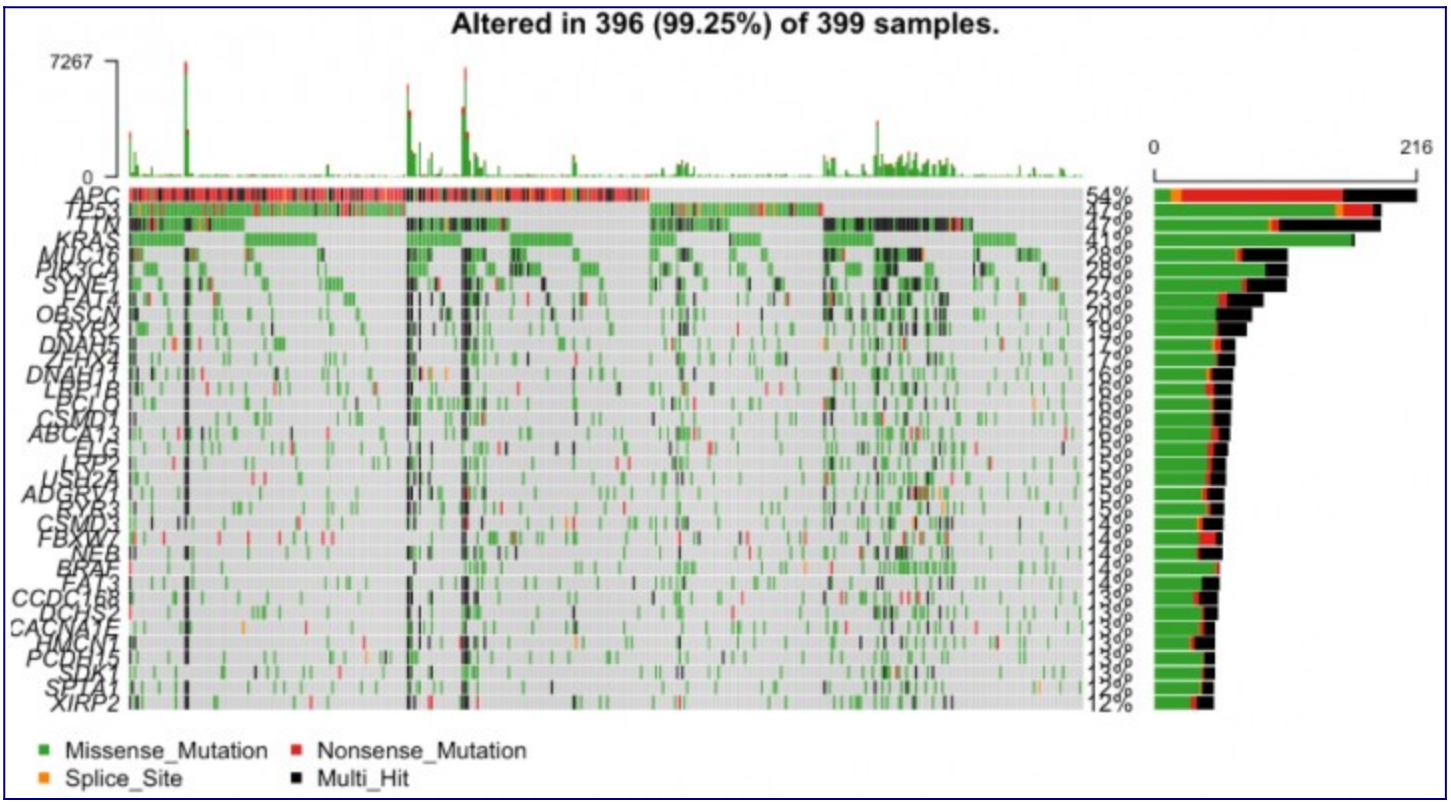
2021 Commencement Address to the Yale Data Science and Statistics Department
harlan-krumholz.medium.com
You have chosen well by selecting Data Science and Statistics as a major. We are in just the beginning of an age of data science, and this age will bring on major changes in almost every facet of…
Shared by Harlan Krumholz
Guidelines for Covariate Adjustment in RCTs
discourse.datamethods.org

i’m susceptible to concise recommendations, thus i like this paper and tempted to share it. “Covariate adjustment for randomized controlled trials revisited” Although it is very difficult to make gen…
Shared by Frank Harrell

Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians
bmj.com
Mendelian randomisation uses genetic variation as a natural experiment to investigate the causal relations between potentially modifiable risk factors and health outcomes in observational data. As wi…
Shared by Ash Paul



April 2021: "Top 40" New CRAN Packages
rviews.rstudio.com
An R community blog edited by RStudio
Shared by RStudio



[Categorizing Continuous Variables](discourse.datamethods.org)

discourse.datamethods.org

[Problems Caused by Categorizing Continuous Variables References | Key reference | Key Reference for Response Variables | More References | Myths About Risk Thresholds | Dichotomania Optimum decisions…]

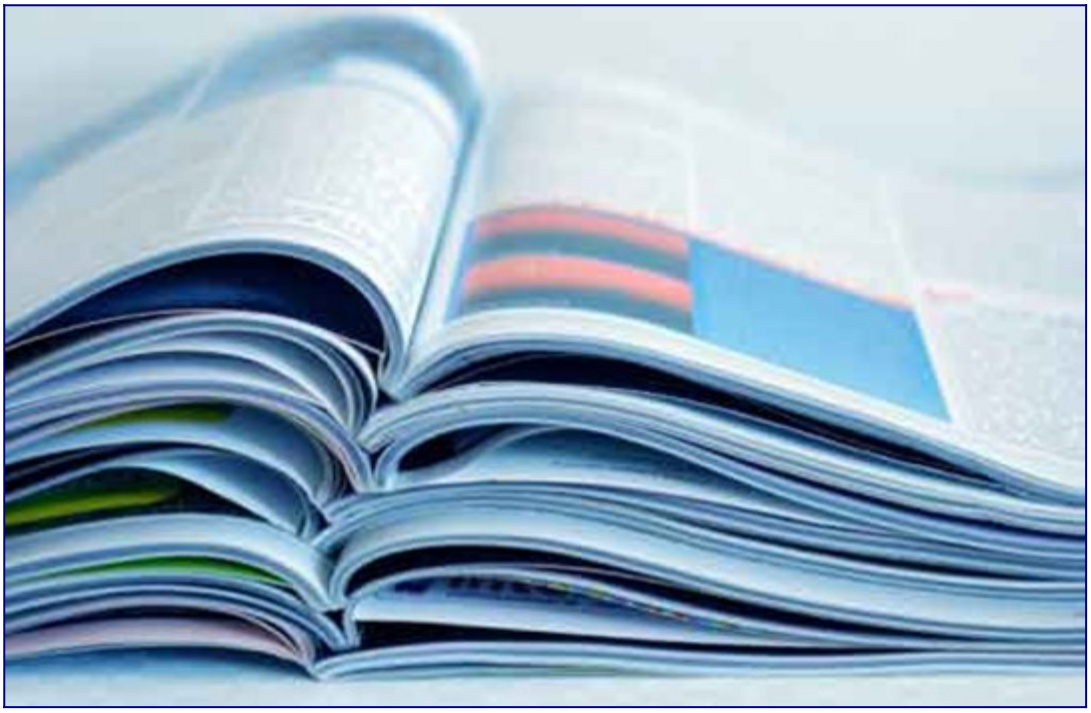
Shared by [Richard D Riley]

[Alex Nowbar's journal reviews—21 May 2021 - The BMJ](blogs.bmj.com)

blogs.bmj.com

[Alex Nowbar reviews the latest research from the top medical journals Ovarian cancer screening is out The UK Collaborative Trial of Ovarian Cancer Screening is a randomised trial comparing annual [.…]

Shared by [Alex Nowbar]

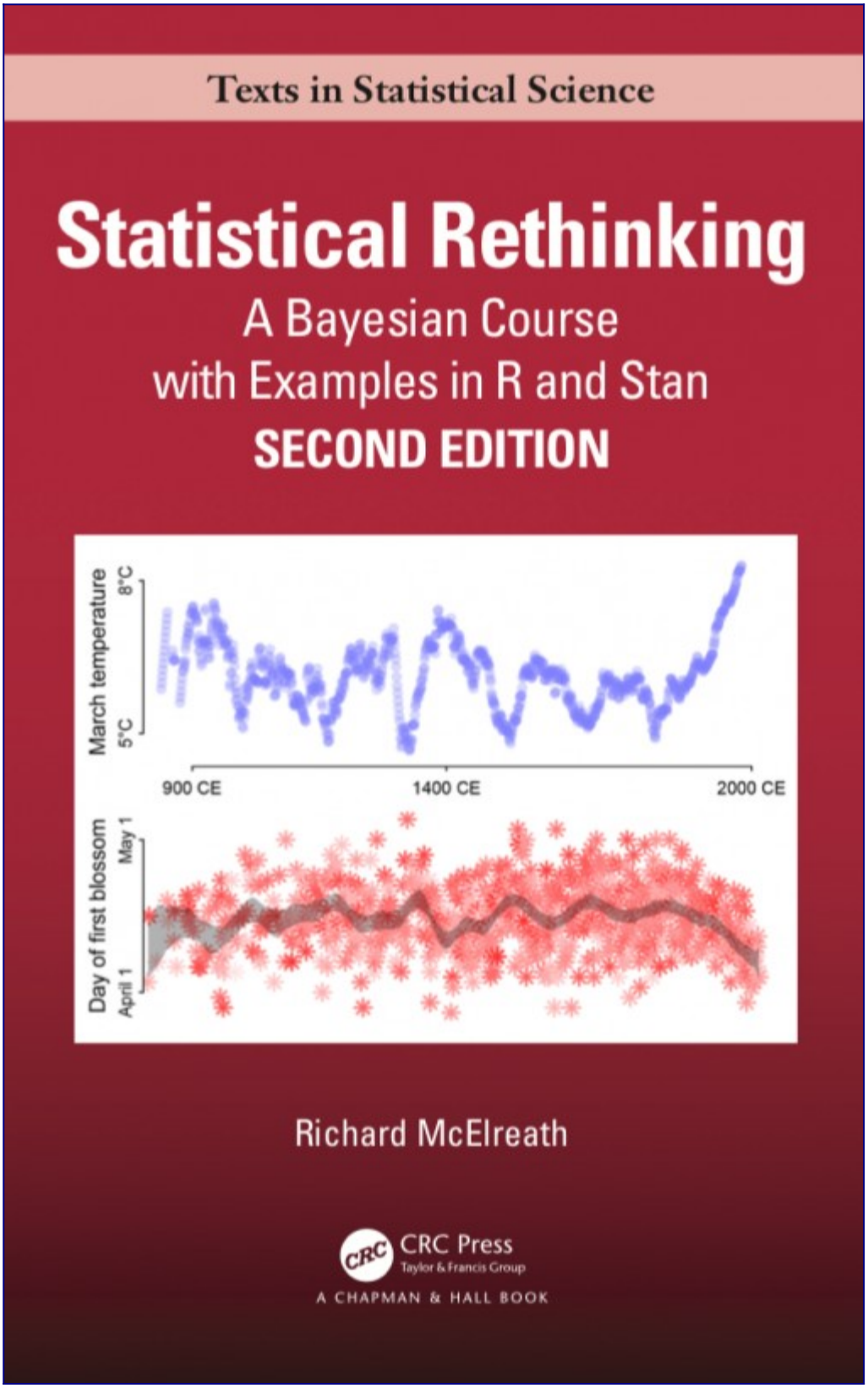


Review: Statistical Rethinking, by Richard McElreath

crosstab.io

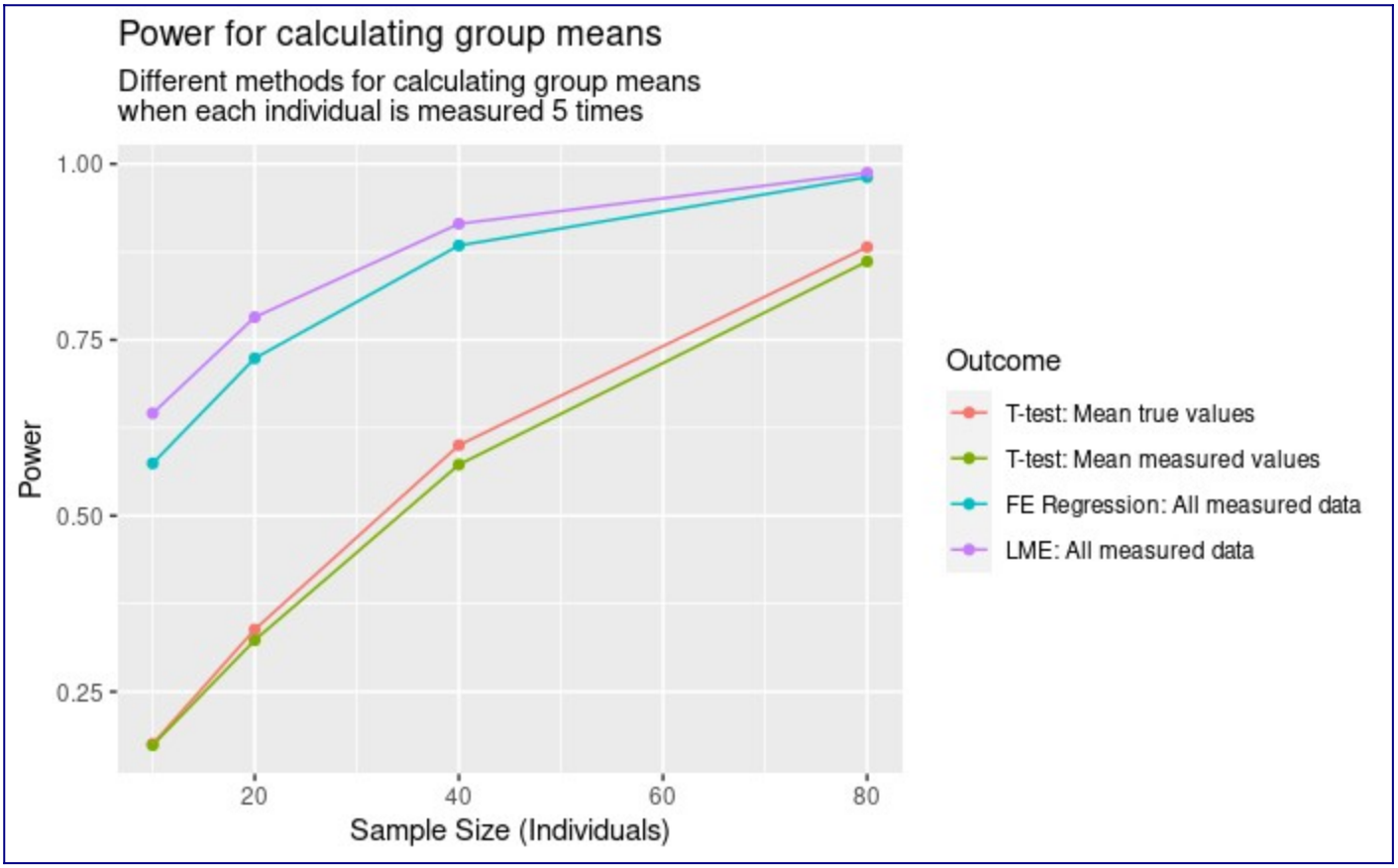
This is an absolute gem of a book. McElreath has found an elusive combination: Statistical Rethinking is not only one of the best intro textbooks for both causal and Bayesian modeling, it's also high…

Shared by The Crosstab Kite



[Statistical conundrum: multiple noisy measurements vs an oracle](discourse.datamethods.org)

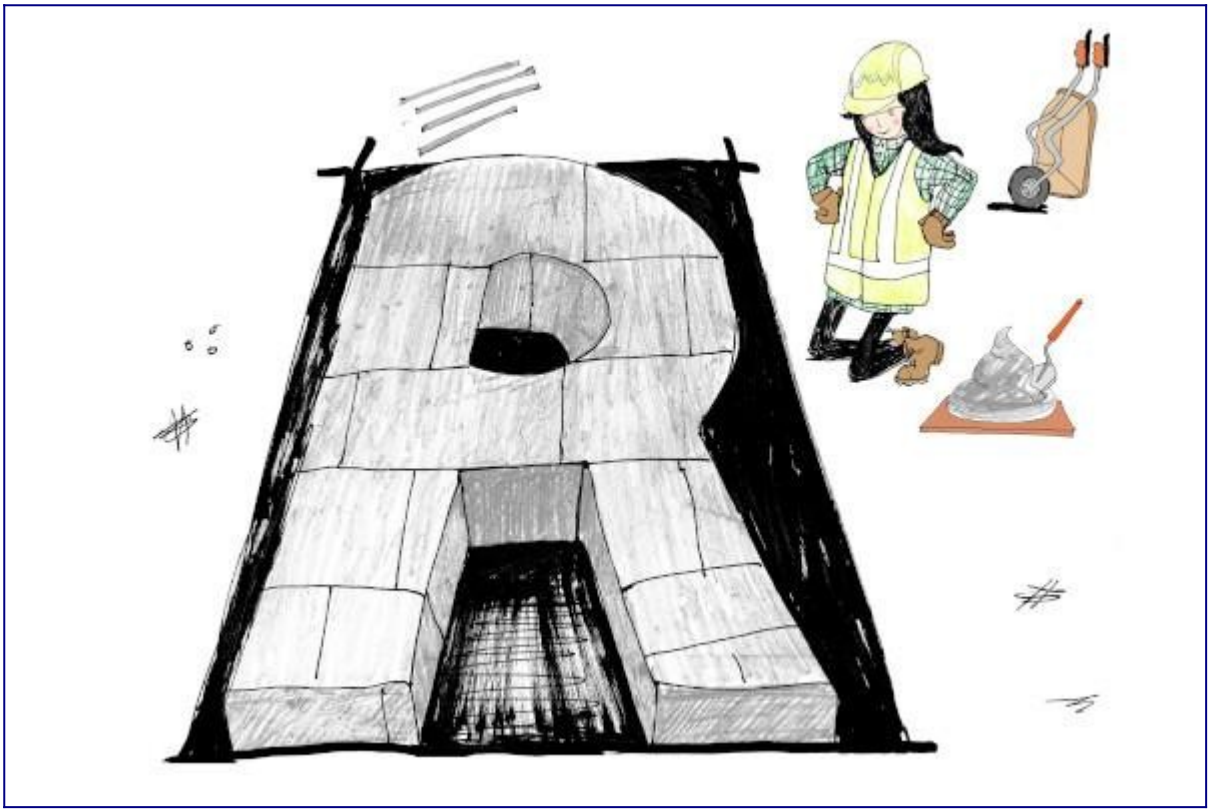
[I have a bit of a statistical conundrum that I can't quite wrap my head around. It appears very simple, but the results are very counterintuitive, and I just can't find an intuitive explanation. This…]



Consistency of causal claims in observational studies: a review of papers published in a general medical journal
bmjopen.bmj.com
Objective To evaluate the consistency of causal statements in observational studies published in The BMJ . Design Review of observational studies published in a general medical journal. Data source C…
Shared by Gary Collins

Modern theories of human evolution foreshadowed by Darwin's Descent of Man
science.sciencemag.org
Charles Darwin's The Descent of Man was published in 1871. Ever since, it has been the foundation stone of human evolutionary studies. Richerson et al. reviewed how modern studies of human biological…
Shared by Allen Frances

New features in R 4.1.0
jumpingrivers.com
R-4.1.0 is released! Rejoice! A new R release (v 4.1.0) is due on 18th May 2021. Typically most major R releases don't contain that many new features, but this release does contain some interesting a…
Shared by Venk Murthy



[Op-ed: The important role data scientists have in defeating COVID vaccine skepticism](#)
indystar.com
[A data-science divide in our nation makes it hard to accept data-driven messages. Here's how scientists can bridge it, one expert writes.](#)
Shared by [ASA Science Policy](#)

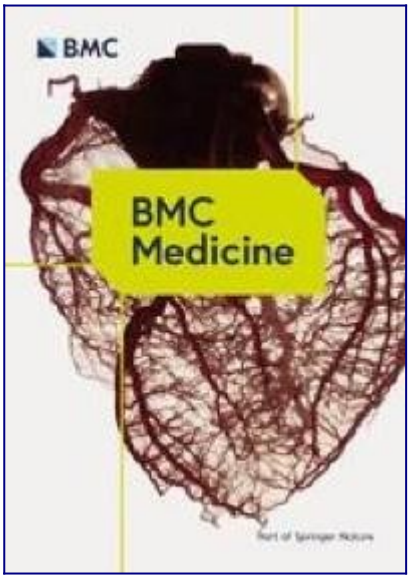
The Teeny, Tiny Scientific Screwup That Helped Covid Kill

wired.com

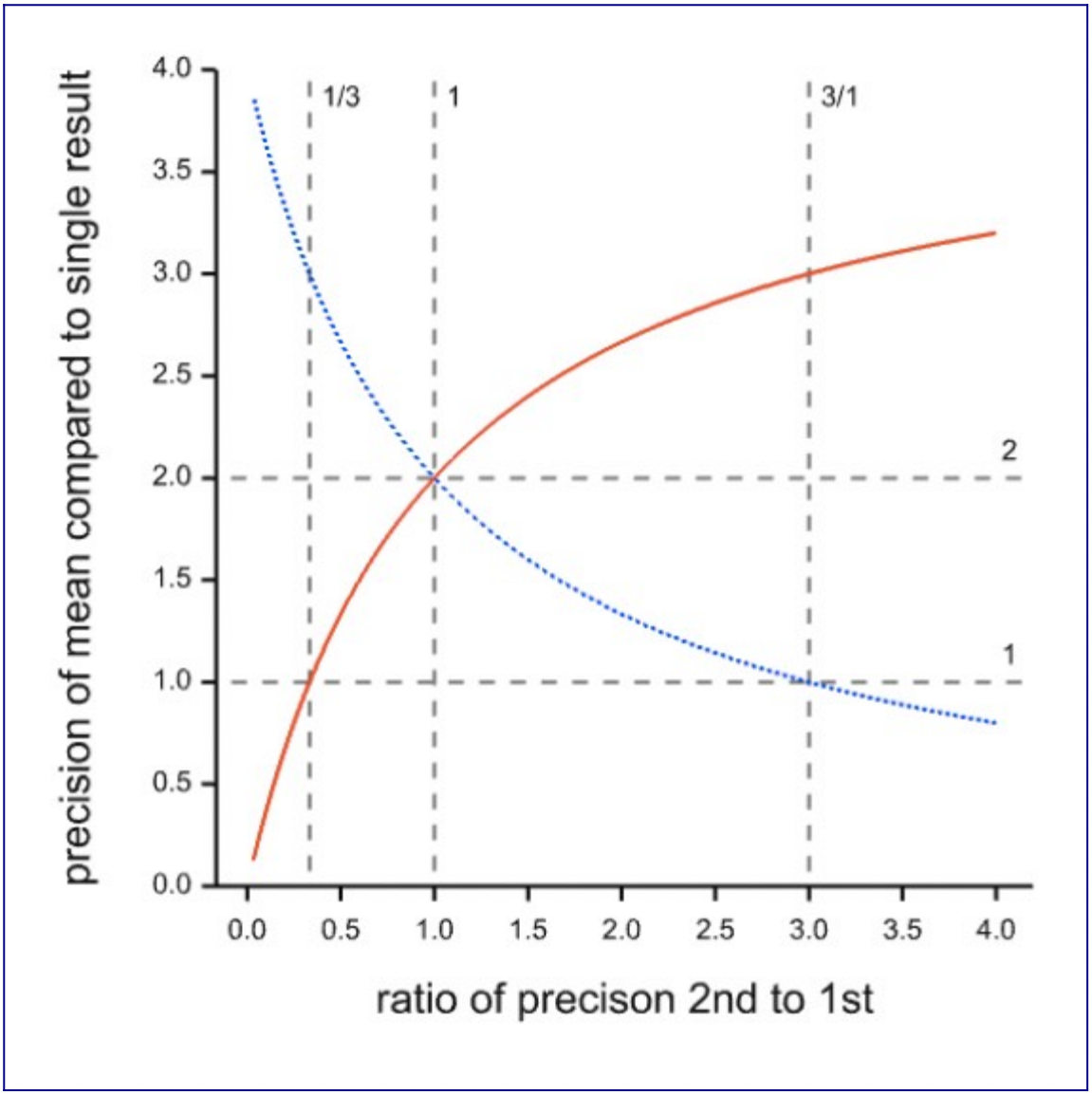
All pandemic long, scientists brawled over how the virus spreads. Droplets! No, aerosols! At the heart of the fight was a mysterious error in decades-old research.

Shared by Retraction Watch

Legal Theory Bookworm: "Noise" by Kahneman, Sibony, & Sunstein
lsolum.typepad.com
The Legal Theory Bookworm recommends Noise: A Flaw in Human Judgment by Daniel Kahneman, Olivier Sibony, & Cass R. Sunstein. Here is a description: Imagine that two doctors in the same city give diff…
Shared by Cass Sunstein
Up for Debate
cancertodaymag.org
Hematologist-oncologist Vinayak K. Prasad describes his positions on cancer research and policy in a recently published book.
Shared by Vinay Prasad MD MPH
Why we need a small data paradigm
bmcmedicine.biomedcentral.com
There is great interest in and excitement about the concept of personalized or precision medicine and, in particular, advancing this vision via various 'big data' efforts. While these methods are nec…
Shared by Tom Freeman



How good are we at predicting the pandemic?
theguardian.com
Models have been useful, especially as humans are far too optimistic and confident
Shared by David Spiegelhalter
S. Senn: "Error point: The importance of knowing how much you don't know" (guest post)
errorstatistics.com
Stephen Senn Consultant Statistician Edinburgh 'The term "point estimation" made Fisher nervous, because he associated it with estimation without regard to accuracy, which he regarded as ridiculous.'…
Shared by Stephen John Senn



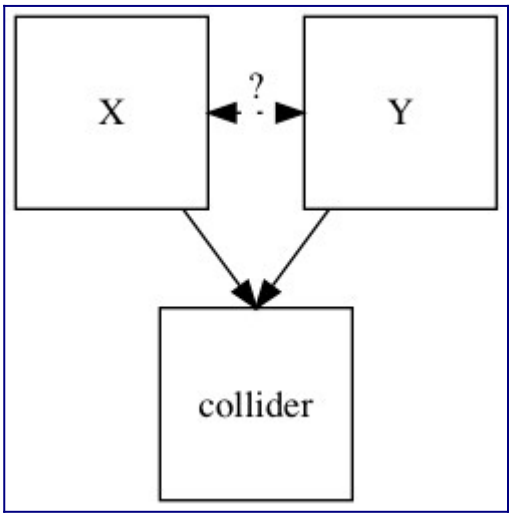
Casual Inference

podcasts.google.com

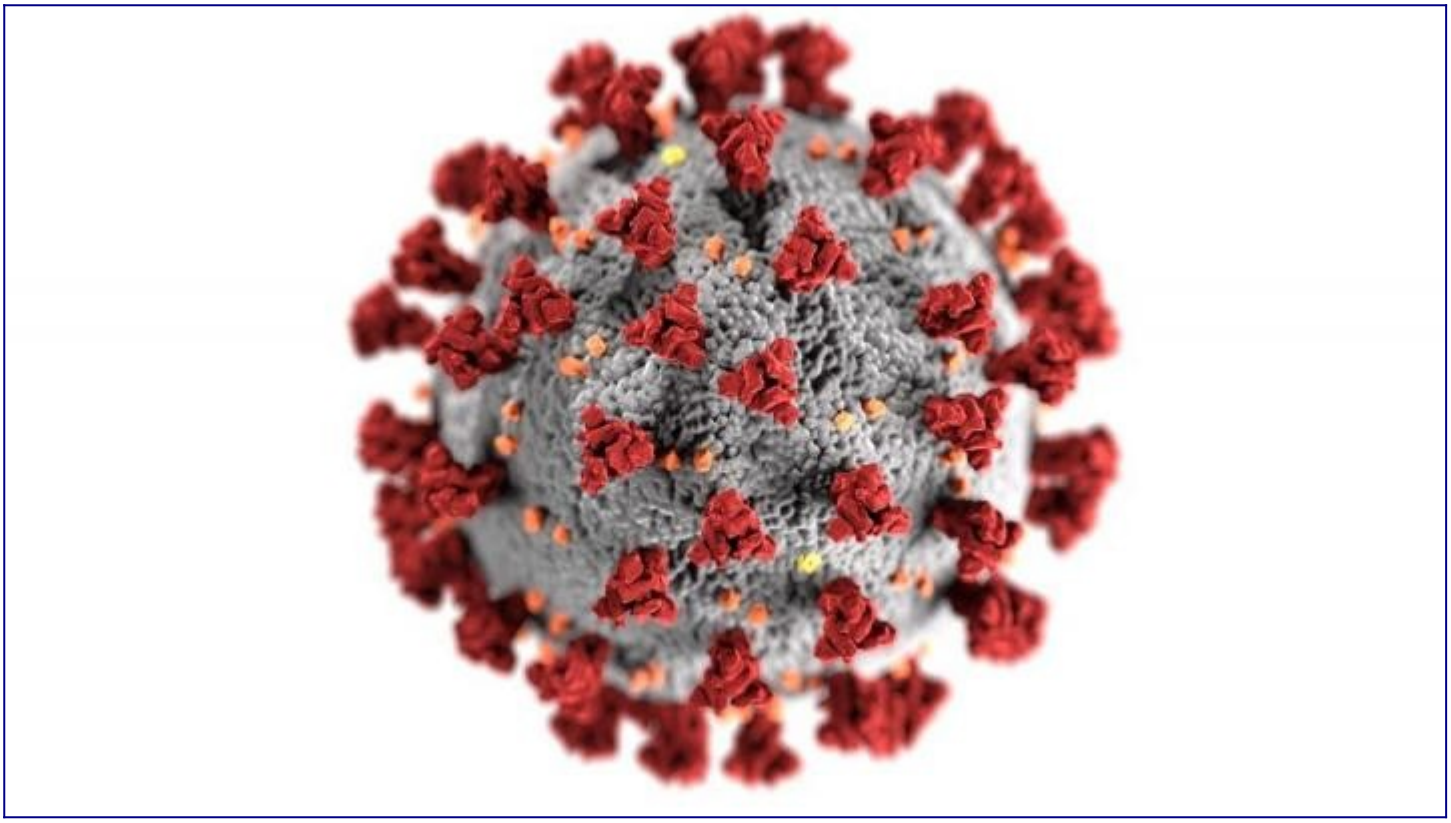
Keep it casual with the Casual Inference podcast. Your hosts Lucy D'Agostino McGowan and Ellie Murray talk all things epidemiology, statistics, data science, causal inference, and public health. Spon…

Shared by Leyla Khazaei

Collider (statistics)
en.wikipedia.org
In statistics and causal graphs, a variable is a collider when it is causally influenced by two or more variables. The name "collider" reflects the fact that in graphical models, the arrow heads from…
Shared by Richard McElreath

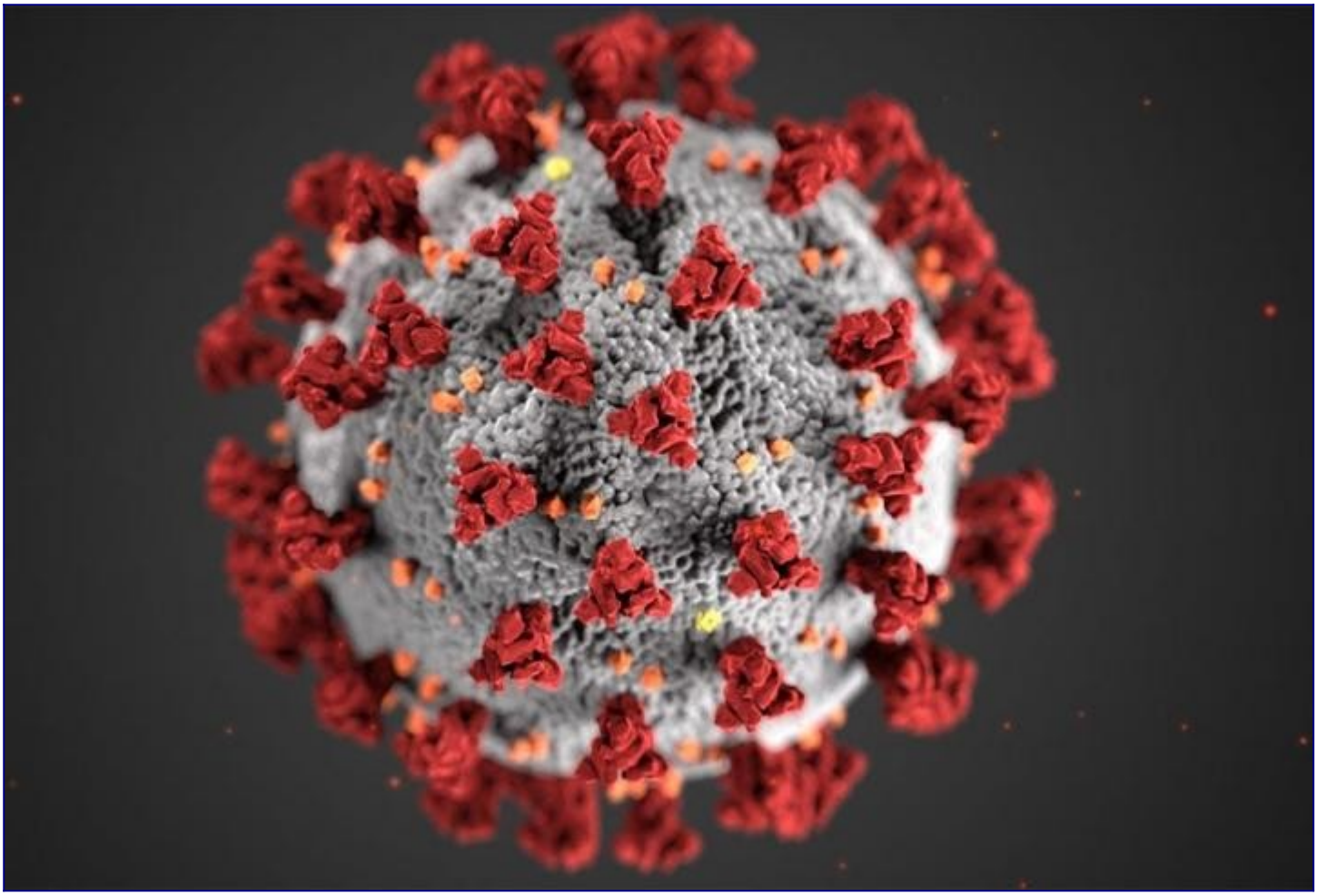


Open Source COVID-19 ML Datasets, Models, and Tools for Data Scientists
hackernoon.com
A look at machine learning-based approaches to global management of COVID-19—including open source project initiatives supported by Google and more.
Shared by Fawad Khan



Recommended Statistics Articles
hbiostat.org
These articles, primarily published online in the past 60d, are recommended because of their potential for use.
Shared by Frank Harrell
Andrew Ng Criticizes The Culture Of Overfitting In Machine Learning
unite.ai
Andrew Ng, one of the most influential voices in machine learning over the last decade, is currently voicing concerns about the extent to which the sector emphasizes innovations in model architecture…
Shared by Richard D Riley
Platform trials: full steam ahead?
rdsblog.org.uk
James Wason, Professor of Biostatistics, RDS North East and North Cumbria Clinical trials provide the highest quality evidence on the benefits and harms of new healthcare interventions. Nevertheless,…
Shared by Victoria Cornelius

Methodological issues plague studies of early autism interventions | Spectrum | Autism Research News
spectrumnews.org
Multiple types of bias and an overreliance on caregiver reports have clouded research on the effectiveness of early interventions for autism for nearly three decades.
Shared by Retraction Watch
Origin of Covid—Following the Clues
nicholaswade.medium.com
The Covid-19 pandemic has disrupted lives the world over for more than a year. Its death toll will soon reach three million people. Yet the origin of pandemic remains uncertain: the political agendas…
Shared by C. Michael Gibson MD



Statistical Misinterpretations and Clinical Practice Guidelines: any examples?
discourse.datamethods.org
In this important paper on the interpretation of statistics in medicine, the authors write: Blockquote
Misinterpretation and abuse of statistical tests, confidence intervals, and statistical power ha…

Indeed, the standard way that statistical hypothesis testing is taught is a 2-way binary grid. Both these dichotomies are inappropriate. « Statistica…
statmodeling.stat.columbia.edu
I originally gave this post the title, “New England Journal of Medicine makes the classic error of labeling a non-significant difference as zero,” but was I was writing it I thought of a more general…
Shared by Andrew Gelman

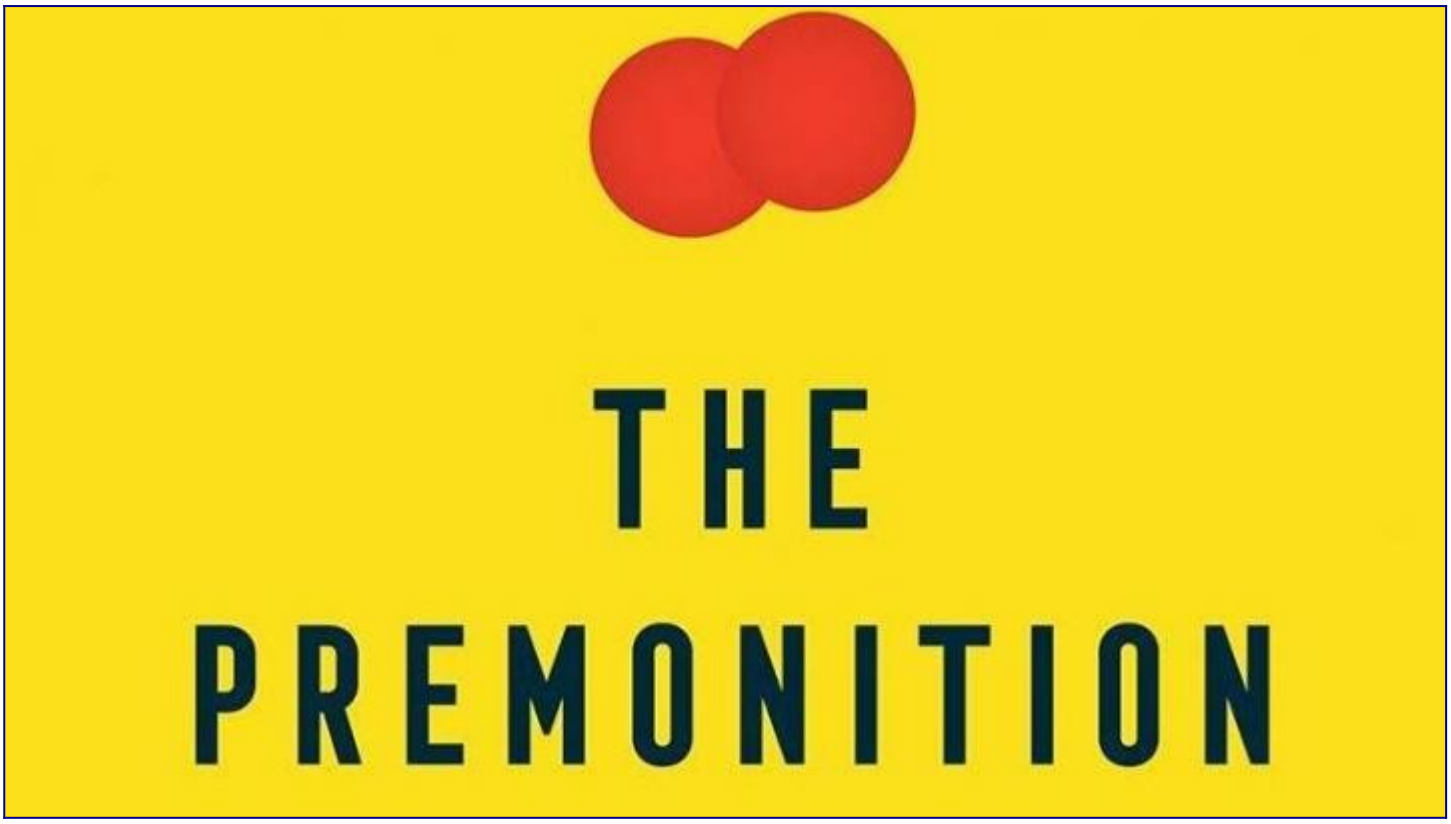
RMS
hbiostat.org
With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis Full Course | Short Course | R rms Package | Course Notes | Study Questions | Why RMS | Discussion Board Boo…
Shared by Frank Harrell

A scoping review of the literature featuring research ethics and research integrity cases
bmcmedethics.biomedcentral.com
The areas of Research Ethics (RE) and Research Integrity (RI) are rapidly evolving. Cases of research misconduct, other transgressions related to RE and RI, and forms of ethically questionable behavi…
Shared by Retraction Watch

CRFCSDAU/EH6126_data_analysis_tutorials
github.com
Contribute to CRFCSDAU/EH6126_data_analysis_tutorials development by creating an account on GitHub.
Shared by Darren Dahly

GlaxoSmithKline (GSK) Joins R Consortium
r-consortium.org
GSK providing COVID era leadership, helping adoption of R Language as standard tool for pharmaceutical industry SAN FRANCISCO, May 4, 2021 – The R Consortium, a Linux Foundation project supporting...
Shared by David Smith

Michael Lewis' 'The Premonition' Is A Sweeping Indictment Of The CDC
npr.org
In a new book, author Michael Lewis writes about public health officers who tried to get others to look at the data on Covid-19 and act to make sure the virus didn't spread.
Shared by Boback Ziaeian



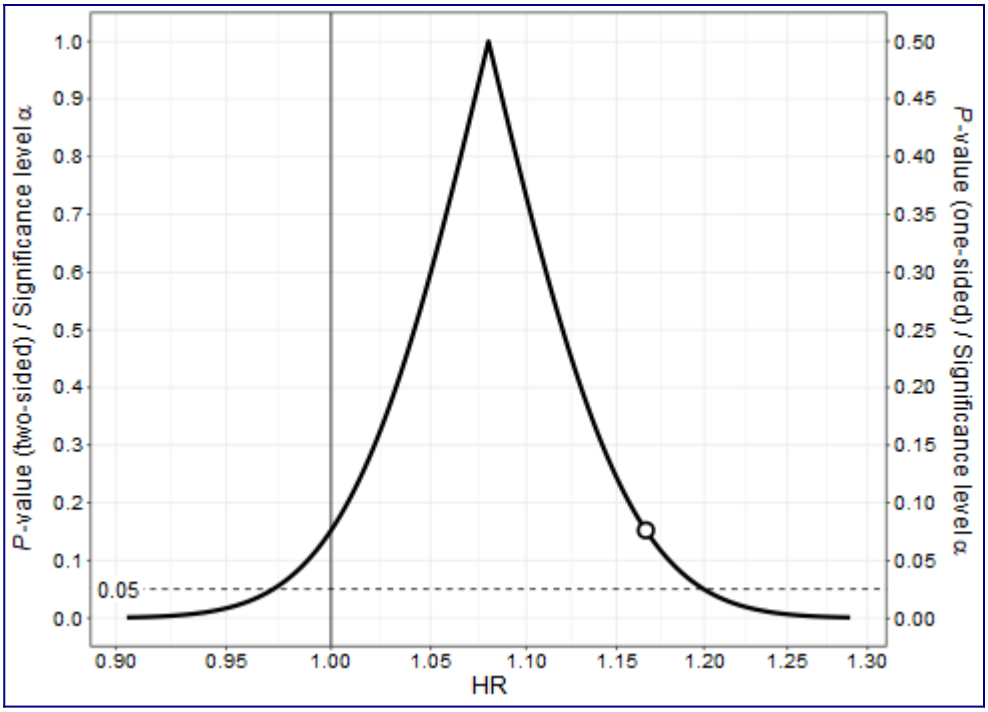
Threshold Science, OSA, and the Statisticians Role
discourse.datamethods.org

I have discussed the need for statisticians to assure that the measurements to which they apply statistics are reproducible and based on evidence rather than opinion. This is what happens when they d…

Another null misinterpretation in a study about the association of epidural analgesia and offspring risk of autism

discourse.datamethods.org

A recent cohort study published in Jama Pediatrics investigated the possible association between epidural labor analgesia (ELA) and offspring risk of autism spectrum disorder (ASD). They included 123…



Time-dependent heterogeneity leads to transient suppression of the COVID-19 epidemic, not herd immunity

pnas.org

Epidemics generally spread through a succession of waves that reflect factors on multiple timescales. Here, we develop a general approach bridging across these timescales and demonstrate how to incor…

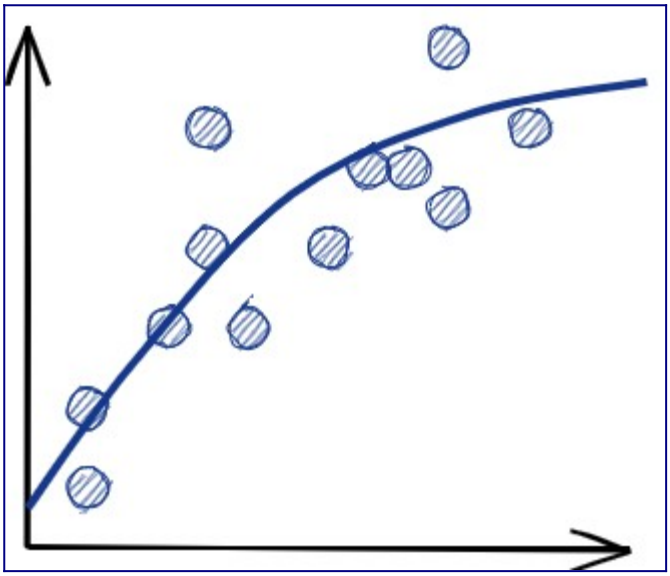
Shared by Arthur Charpentier

In defense of statistical modeling

crosstab.io

Data science remains hot but there is a persistent stream of articles that says the field is overhyped and that hiring managers and aspiring data scientists should focus more on engineering. Let's re…

Shared by The Crosstab Kite



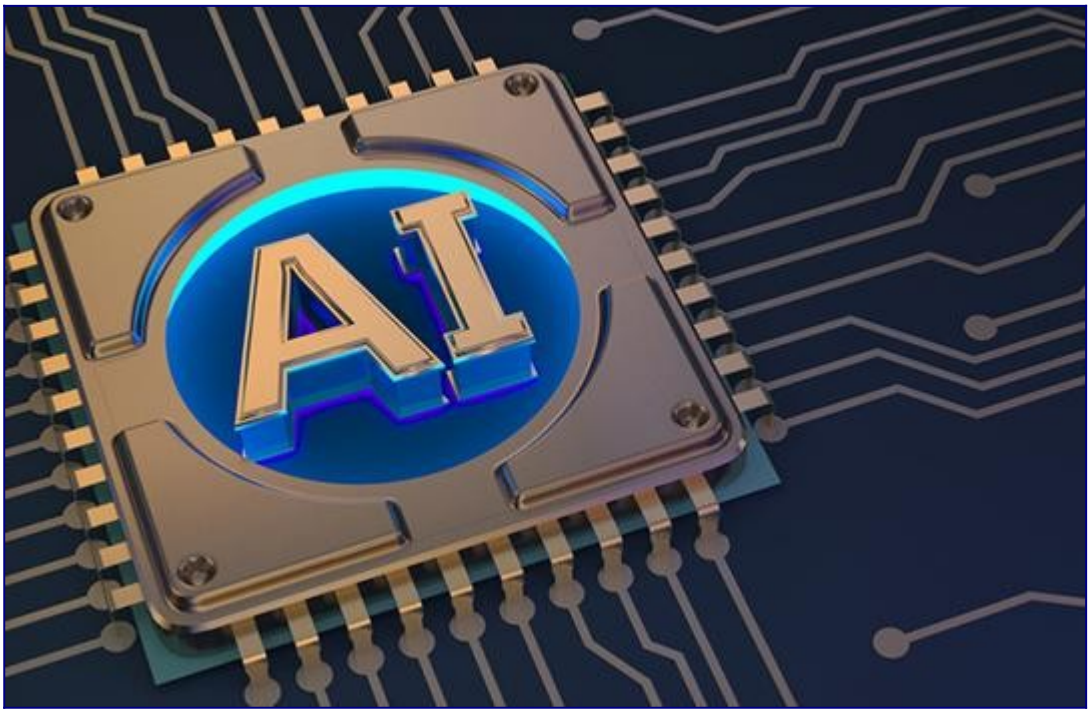
[A Review of Bayesian Perspectives on Sample Size Derivation for Confirmatory Trials](#)
tandfonline.com
[Sample size derivation is a crucial element of planning any confirmatory trial. The required sample size is typically derived based on constraints on the maximal acceptable Type I error rate and mi...](#)
Shared by [Andrew Althouse](#)

THE AMERICAN STATISTICIAN

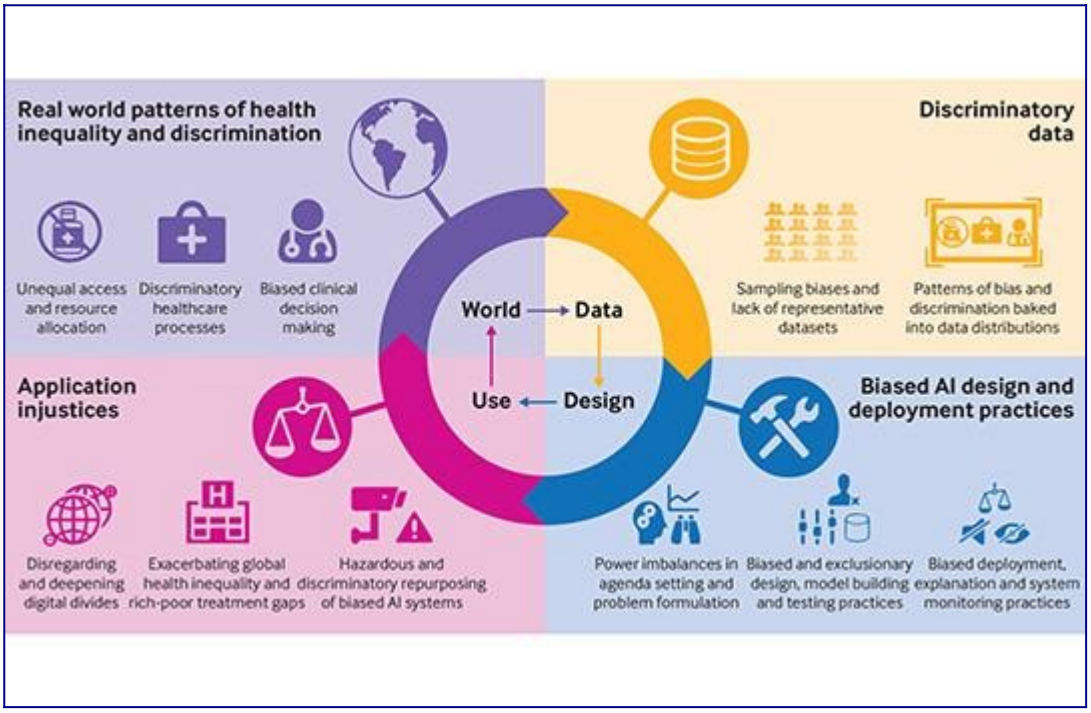
A PUBLICATION OF THE AMERICAN STATISTICAL ASSOCIATION

VOLUME 73 • NUMBER 2 MAY 2019

Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiv…
bmj.com
Machine learning, artificial intelligence, and other modern statistical methods are providing new opportunities to operationalise previously untapped and rapidly growing sources of data for patient b…
Shared by Daniel Gould



My Journey From Frequentist to Bayesian Statistics | Statistical Thinking
fharrell.com
The difference between Bayesian and frequentist inference in a nutshell: With Bayes you start with a prior distribution for θ and given your data make an inference about the θ-driven process generati…
Shared by Frank Harrell
Does “AI” stand for augmenting inequality in the era of covid-19 healthcare?
bmj.com
Artificial intelligence can help tackle the covid-19 pandemic, but bias and discrimination in its design and deployment risk exacerbating existing health inequity argue David Leslie and colleagues Am…
Shared by Darren Dahly



The obscure maths theorem that governs the reliability of Covid testing
theguardian.com
There's been much debate about lateral flow tests – their accuracy depends on context and the theories of a 18th-century cleric
Shared by Kareem Carr

Perspective | I'm an epidemiologist. Here's what I got wrong about covid.
washingtonpost.com
Many attacks on the field have been wildly unfair. But we experts made some missteps.
Shared by Dr Ellie Murray

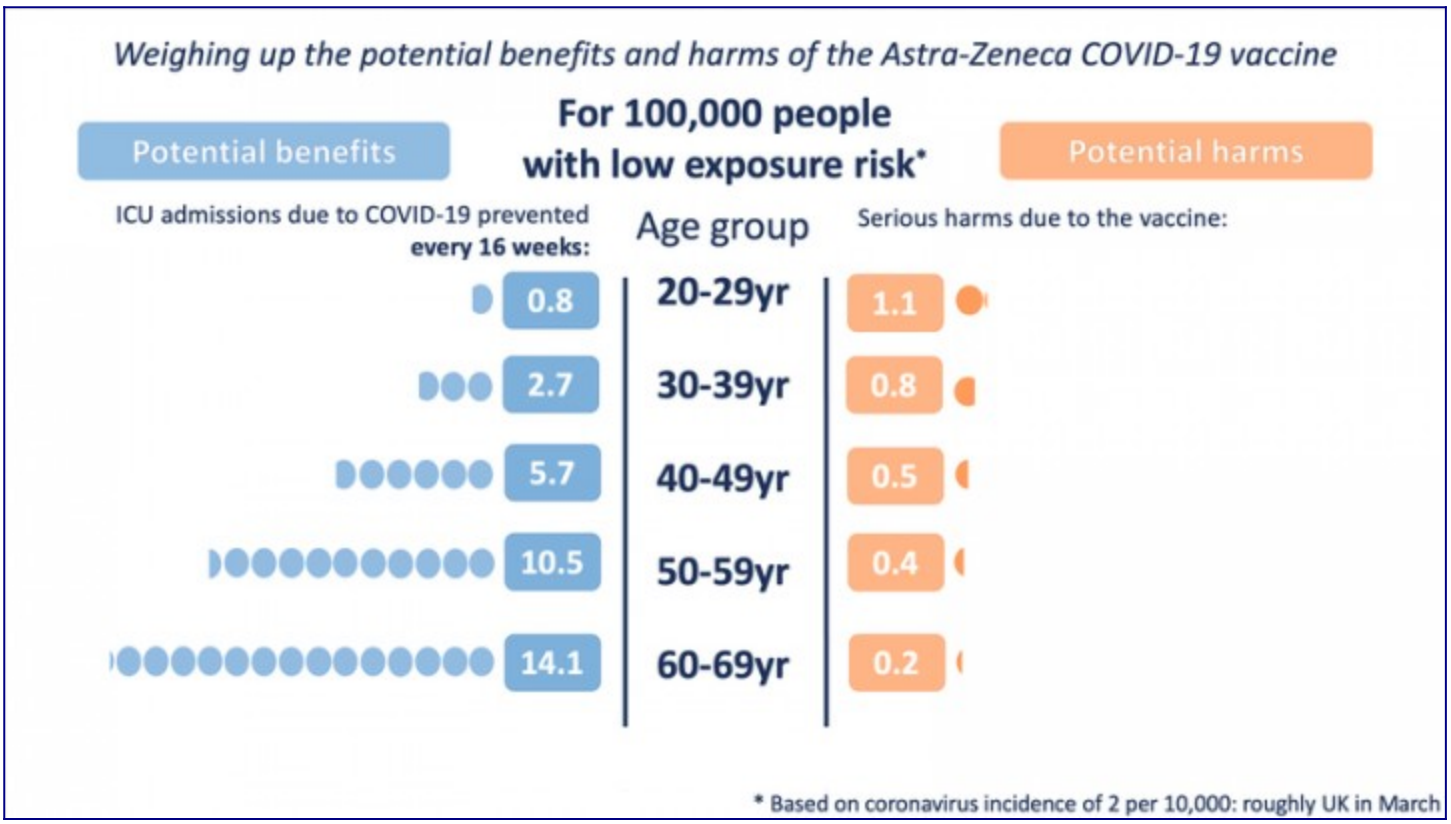
Latest news from the R Markdown family
blog.rstudio.com
Happy spring! In this post, we'll share the latest news from the R Markdown family of packages so that you know all you need to know to take advantage of the newest features. New releases include the…
Shared by RStudio

The most compatible value with the data within a confidence interval
discourse.datamethods.org
I am reading the fantastic book "Statistical Rethinking" by Richard McElreath, in which he says, "The posterior mean line is just the posterior mean, the most plausible line in the infinite universe …
Shared by Sander Greenland

Heterogeneity in RCTs with continuous outcome
discourse.datamethods.org
Hello all, I’m writing here a plea for literature recommendations or theoretical guidance. My research group work in RCTs with continuous outcomes (e.g. symptoms of depression) where the typical pres…

Rob Tibshirani, Yuling Yao, and Aki Vehtari on cross validation « Statistical Modeling, Causal Inference, and Social Science
statmodeling.stat.columbia.edu
About 9 years ago I emailed you about our new significance result for the lasso. You wrote about in your blog. For some reason I never saw that full blog until now. I do remember the Stanford-Berkele…
Shared by Andrew Gelman
Why I Won't Review or Write for Elsevier and Other Commercial Scientific Journals - The Wire Science
science.thewire.in
Can scientists who are so meticulous in preparing their papers and so generous with their time in reviewing them for free not find better ways to advance science than relying on profiteering journals?
Shared by Retraction Watch

Winton Centre Cambridge
wintoncentre.maths.cam.ac.uk
Winton Centre for Risk and Evidence Communication
Shared by David Spiegelhalter



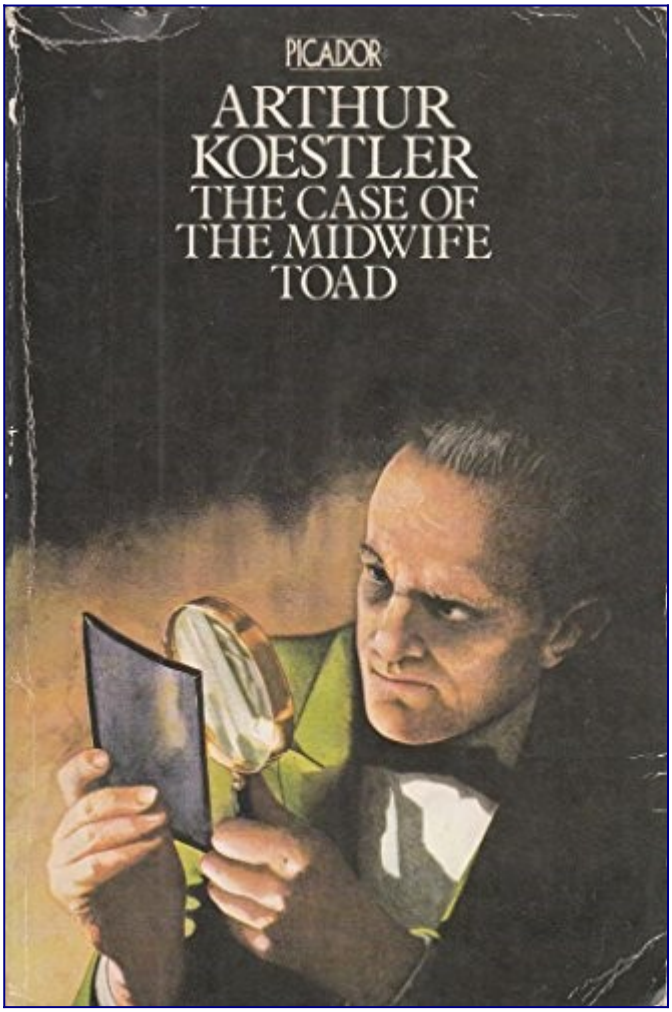
The Art and Science of Biostatistics
now.tufts.edu
Biostatisticians are behind countless medical and public health breakthroughs
Shared by Saqib Ali

Air pollution and COVID-19 mortality in the United States: Strengths and limitations of an ecological regression analysis
advances.sciencemag.org
Assessing whether long-term exposure to air pollution increases the severity of COVID-19 health outcomes, including death, is an important public health objective. Limitations in COVID-19 data availa…
Shared by MrKW

Characterizing physician EHR use with vendor derived data: a feasibility study and cross-sectional analysis
academic.oup.com
AbstractObjective. To derive 7 proposed core electronic health record (EHR) use metrics across 2 healthcare systems with different EHR vendor product installati
Shared by Harlan Krumholz

Current statistics journal articles
hbiostat.org
Recommended articles from stat journals primarily from the past 60 days
Shared by Frank Harrell

George Spencer Brown and Fisher’s Experimental Paradox.
bayesicsolutions.wordpress.com
George Spencer Brown, who died in 2016, was, to say the least, an eccentric. But also a polymath in the grand, 19th century British tradition. George Spencer Brown (Copyright Markus Heidingsfelder) B…
Shared by Andy Grieve



Data Cleaning IS Analysis, Not Grunt Work

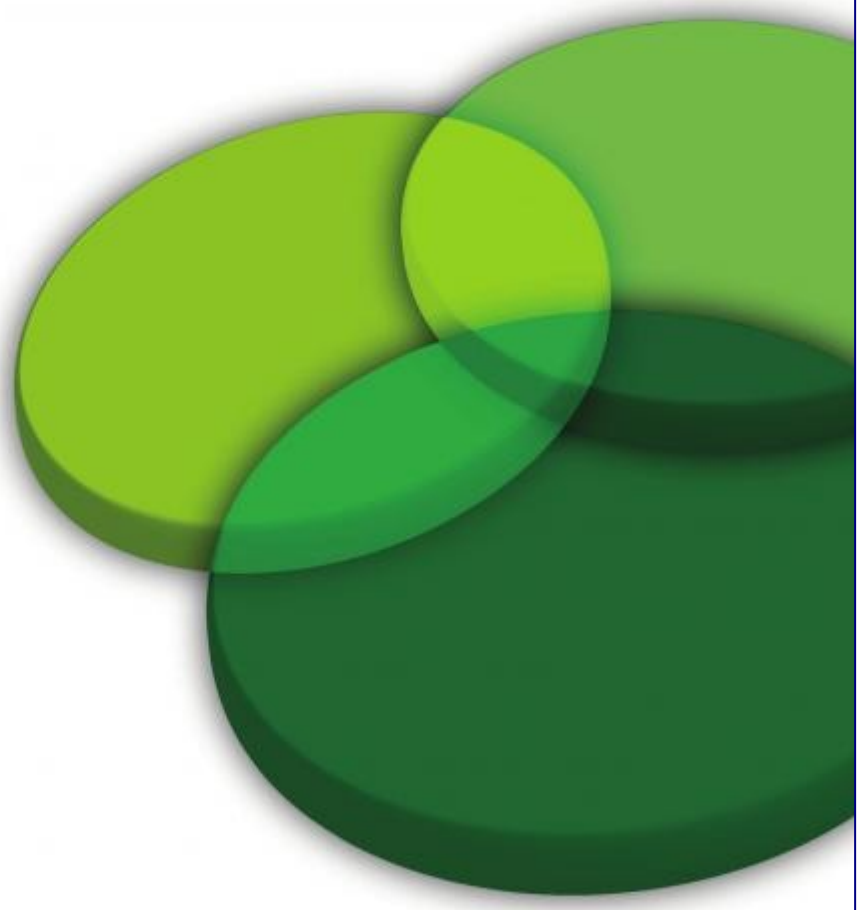
counting.substack.com
Also, most data cleaning articles suck
Shared by Asmae Toumi

Split-Sample Model Validation | Statistical Thinking
fharrell.com
Methods used to obtain unbiased estimates of future performance of statistical prediction models and classifiers include data splitting and resampling. The two most commonly used resampling methods a…
Shared by Pedro Lopez-Ayala
Teaching clinicians shared decision making and risk communication online: an evaluation study
ebm.bmj.com
Objectives To describe the development and initial evaluation of a brief e-learning course as a means of teaching shared decision making and risk communication skills to clinicians of all specialties…
Shared by David Spiegelhalter

April 2021 Volume 26 Issue 2

BMJ EVIDENCE-BASED MEDICINE

EXPERTISE VALUES EVIDENCE



Coroners' concerns to prevent harms: a series of coroners' case reports to serve patient safety and educate the public, clinicians and policy-makers

Efficacy of reminders for increasing volunteer engagement in translating Cochrane plain language summaries: a pilot randomised controlled trial

Does ursodeoxycholic acid have a role in the treatment of women with intrahepatic (obstetric) cholestasis of pregnancy anymore?

BMJ

Teaching statistics online (for the first time): Lessons learned
russpoldrack.org
We just completed the 2021 Winter Quarter at Stanford, and it was my first time to teach my introductory statistics course fully online. Th...
Shared by Darren Dahly

Methods for Evaluation of medical prediction Models, Tests And Biomarkers (MEMTAB) 2020 Symposium
diagnprognres.biomedcentral.com
Background: Guidelines recommend identifying in early pregnancy women at elevated risk of pre-eclampsia. Existing prediction tools perform poorly among nulliparous women. We aimed to 1) develop and v…
Shared by Frank Harrell

Web-Based Lectures
amstat.org
Web-based lectures.
Shared by ASA

economics of ‘more research is needed’
academic.oup.com
Abstract. Background Results from epidemiology and other health research affect millions of life-years and billions of dollars, and the research directly consum
Shared by Daniël Lakens

Casual Inference - Community Engagement, Health Disparities, and Measure Development with Melody Goodman | Episode 13
podcasts.google.com
Ellie Murray and Lucy D'Agostino McGowan discuss community engagement, health disparities, and measure development with Dr. Melody Goodman from New York University Global School of Public Health. …
Shared by Shaun McGirr

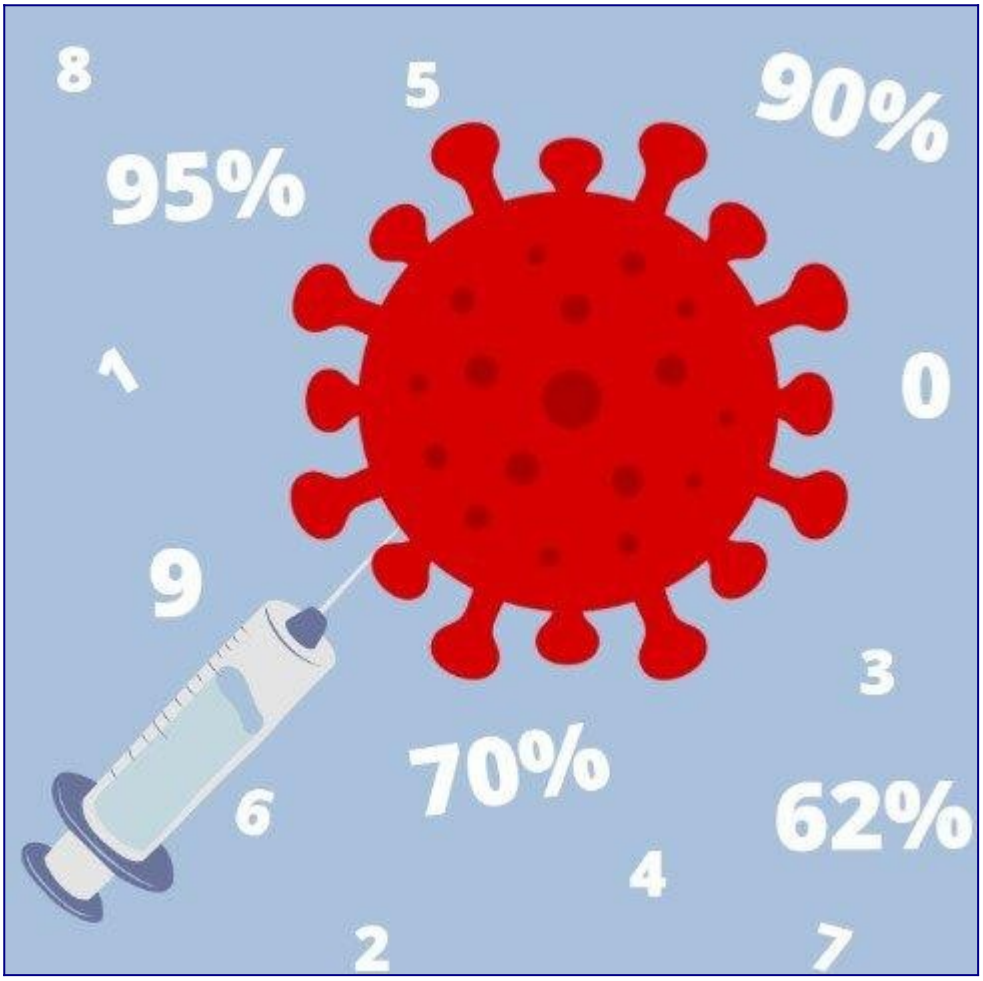
What the heck happened to John Ioannidis?
sciencebasedmedicine.org
John Ioannidis is a highly influential scientist whom we at SBM have lauded. Since the pandemic, though, he's downplayed the dangers of COVID-19. What happened?
Shared by Boback Ziaeian

The saddest equation in data science
towardsdatascience.com
Prepare a box of tissues! I’m about to drop a truth bomb about statistics and data science that’ll bring tears to your eyes. There is no magic in the world that lets you make something out of…
Shared by Kareem Carr

Research integrity in the COVID-19 era: insights from Retraction Watch co-founder Ivan Oransky
thepublicationplan.com
Read Ivan Oransky’s take on the retraction landscape and his perspectives on how to tackle research integrity issues in the future.
Shared by Retraction Watch

Reference Collection to push back against "Common Statistical Myths"
discourse.datamethods.org
Note: This topic is a wiki, meaning that this main body of the topic can be edited by others. Use the Reply button only to post questions or comments about material contained in the body, or to sugge…
Shared by Tired Old Health Stats Dude

Are the statistics from COVID-19 vaccine trials understood? — Data Saves Lives
datasaveslives.eu
Results from COVID-19 vaccine trials have been widely reported in the media. As a clinical trial statistician, I was interested to find out what members of the public thought about the results th…
Shared by Victoria Cornelius

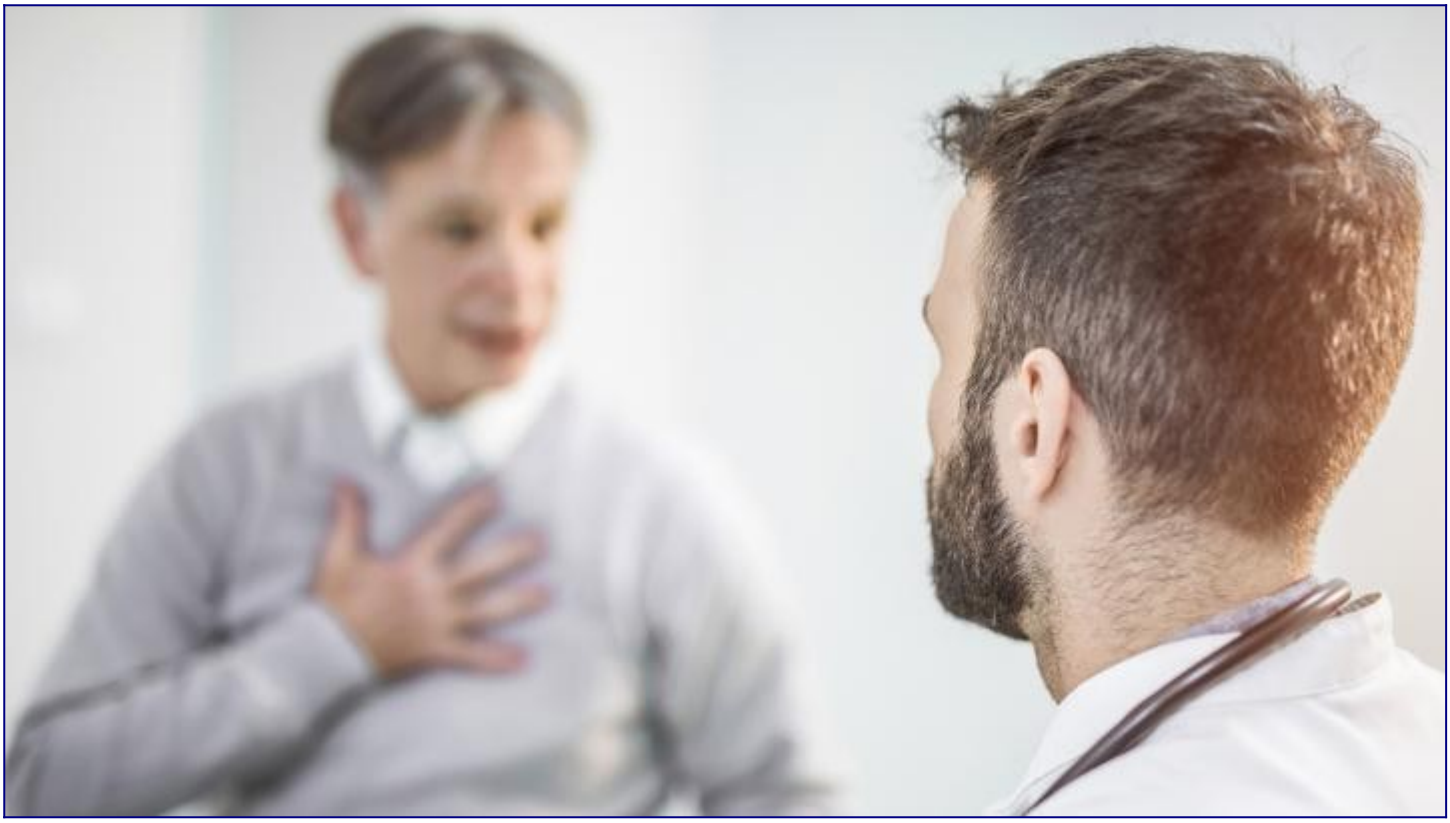


[The Folk Theorem, revisited « Statistical Modeling, Causal Inference, and Social Science](#)
statmodeling.stat.columbia.edu
[It's time to review the folk theorem, an old saw on this blog, on the Stan forums, and in all of Andrew's and my applied modeling. Andrew uses "folk" in the sense of being folksy as opposed to rigoro…](#)
Shared by [Andrew Gelman](#)

[General Multi-parameter Transformations in rms](#)
hbiostat.org
[As of version 6.2-0 of the rms package appearing in CRAN on 2021-03-18, there is a new transformation function gTrans, joining lsp, rcs, pol, etc. gTrans takes a numeric, character, or factor vector …](#)
Shared by [Frank Harrell](#)

[Data Corruption: DOJ Targets Fraud In Medical Research Trial In The Era Of COVID-19 | JD Supra](#)
jdsupra.com
[In late 2019, Gustav Eyler, the Director of the U.S. Department of Justice's ("DOJ") Consumer Protection Branch ("CPB"), cited fraudulent data in clinical research trials for new drugs...](#)
Shared by [Retraction Watch](#)

Models v. Evidence
bostonreview.net
Models v. Evidence from Boston Review. COVID-19 has revealed a contest between two competing philosophies of scientific knowledge. To manage the crisis, we must draw on both.
Shared by Post-Pandemic University
Growing Drumbeat for Patient-Reported Outcomes in Heart Failure
medpagetoday.com
Disconnect reported between NYHA, KCCQ scores
Shared by Javed Butler



Recommended Statistics Journal Articles
hbiostat.org
Articles selected on the basis of likely applicability to biomedical research.
Articles selected from 14 statistics journals and mainly published online in the past 60 days.
Shared by Frank Harrell

[Citizen Science and Mapping the Microbiome]
medscape.com
[Jessica Richman, CEO of uBiome, talks about her passion for citizen science, her interest in the microbiome, and how uBiome aims to serve public health.]
Shared by [Boback Ziaeian]

Opinion

wsj.com

Echoes of chronic fatigue in the effort to blame the coronavirus for a host of questionable symptoms.

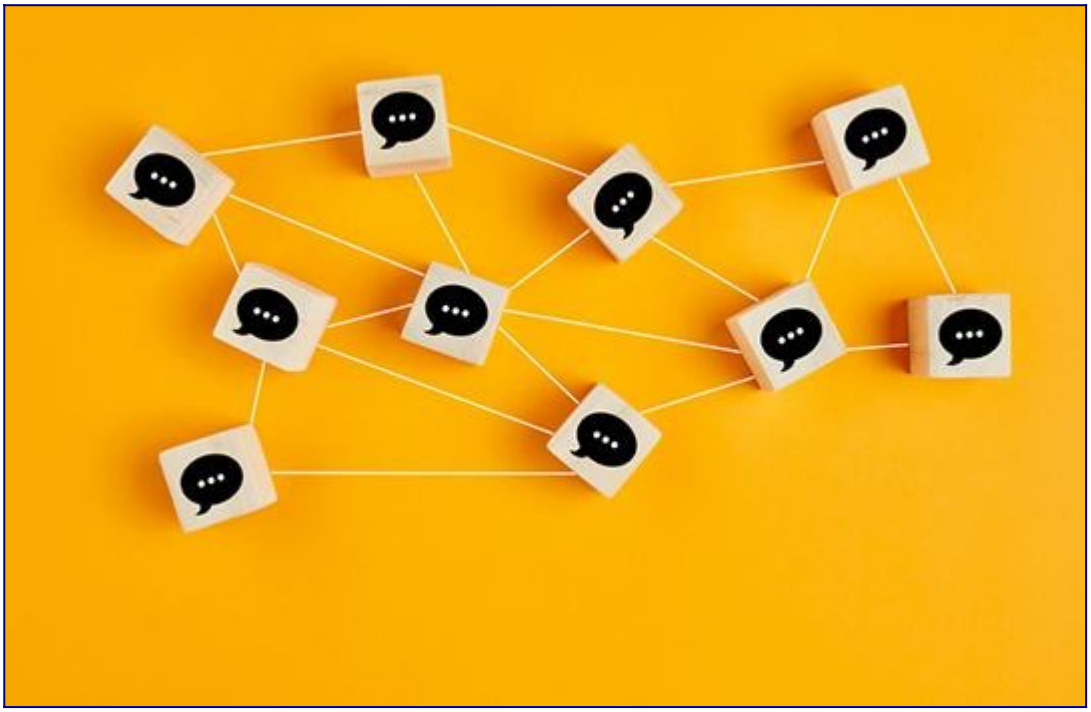
Shared by Vinay Prasad MD MPH

An open plea for dignity and respect in science - The BMJ

blogs.bmj.com

Only by listening to and debating honest opposing views, can we hope to overcome the fear and anger that surrounds us, say these authors Covid-19 has generated multiple and multifarious [...]More...

Shared by Vinay Prasad MD MPH



3 Ways the Pandemic Has Made the World Better

theatlantic.com

COVID-19 has inflicted devastating losses. It has also delivered certain blessings.

Shared by Retraction Watch

Dismissive reviews: A cancer on the body of knowledge
retractionwatch.com
Richard P. Phelps Observers describe the quantity of research information now produced variously as “torrent,” “overload,” “proliferation,” or the like. Technological advances in computing and teleco…
Shared by Retraction Watch

Can Sophisticated Study Designs With Regression Analyses of Observational Data Provide Causal Inferences?
jamanetwork.com
This Viewpoint presents considerations for assessing evidence for causal inference when using sophisticated study designs with regression analyses of longitudinal observational data. A view is someti…
Shared by JAMA Psychiatry
Why There Is So Much Confusion About the AstraZeneca Vaccine
newyorker.com
A lack of transparency seems to have caused unnecessary doubts—and damaged the effort to build public trust in vaccines.
Shared by Boback Ziaeian

No RCT for Masks? No Problem

medpagetoday.com

Other forms of evidence are available to judge effectiveness of this and other interventions

Shared by Jonny 5yve

[Reflections on five years of the janitor R package]
samfirke.com
[One thing led to another. In early 2016, I was participating in discussions on the #rstats Twitter hashtag, a community for users of the R programming language. There, Andrew Martin and I met and rea…]
Shared by [Darren Dahly]

[General Multi-parameter Transformations in rms]
hbiostat.org

As of version 6.2-0 of the rms package appearing in CRAN on 2021-03-18, there is a new transformation function gTrans, joining lsp, rcs, pol, etc. gTrans takes a numeric, character, or factor vector …

Shared by Frank Harrell

It works: 0 deaths, only 4 severe cases among 523,000 fully vaccinated Israelis

timesofisrael.com

HMO data a week after 2nd dose shows 93% effectiveness, ‘unequivocally' proving vaccine's success and leaving ‘no doubt’ it’s saved many Israeli lives, says Maccabi official

Shared by VolleMelk

The Continuing Path of Data Science Maturity

opendatascience.com

In the ever-evolving growth of data science maturity in 2021, these are a few key trends to consider, such as curating data, ethics, and ModelOps.

Shared by Marinela Profi

Good Science
sites.google.com
Innovation and solutions in evidence science
Shared by Frank Harrell

Epistemologies of predictive policing: Mathematical social science, social physics and machine learning
journals.sagepub.com
Predictive policing has become a new panacea for crime prevention. However, we still know too little about the performance of computational methods in the context of predictive policing. The paper ...
Shared by Big Data & Society

Underlying cardiovascular risk and major adverse cardiovascular events after acute respiratory infection: a population-based cohort study of over 4.2…
medrxiv.org
medRxiv - The Preprint Server for Health Sciences
Shared by medRxiv

What is GRADE?
bestpractice.bmj.com
GRADE (Grading of Recommendations, Assessment, Development and Evaluations) is a transparent framework for developing and presenting summaries of evidence and provides a systematic approach for makin…
Shared by Texutree

Potential value and impact of data mining and machine learning in clinical diagnostics
pubmed.ncbi.nlm.nih.gov
Data mining involves the use of mathematical sciences, statistics, artificial intelligence, and machine learning to determine the relationships between variables from a large sample of data. It has p…
Shared by Oncology & Machine Learning

[The global burden of yellow fever.]
europepmc.org
[Europe PMC is an archive of life sciences journal literature.]
Shared by [MRC publications]
[Unmet Need for Equipment to Help With Bathing and Toileting Among Older US Adults]
jamanetwork.com
[Question How many older adults who need equipment to help with bathing and toileting do not have it?
Findings In this nationally representative cohort study of 2614 adults 65 years or older, an estim…]

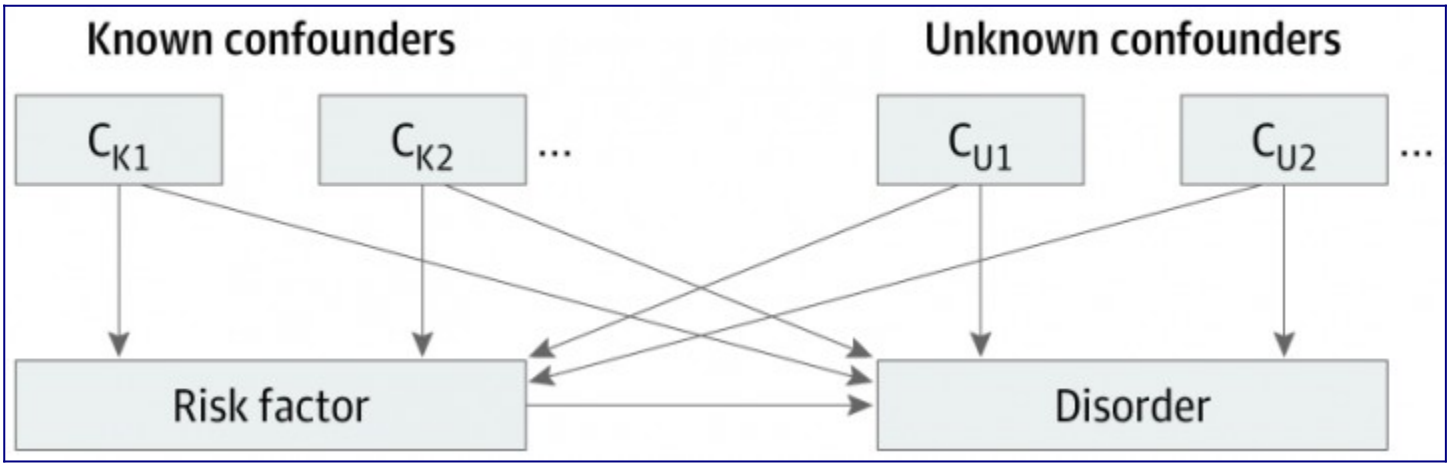
Shared by JAMA Internal Medicine

Causal Inference in Psychiatric Epidemiology

jamanetwork.com

This review examines approaches to causal inference in psychiatric epidemiology.

Shared by JAMA Psychiatry



Impact of the timing and adherence to face mask use on the course of the COVID-19 epidemic in France

academic.oup.com

Global spread of coronavirus disease 2019 (COVID-19) has created an unprecedented infectious disease crisis worldwide.1 To date, non-pharmaceutical intervention

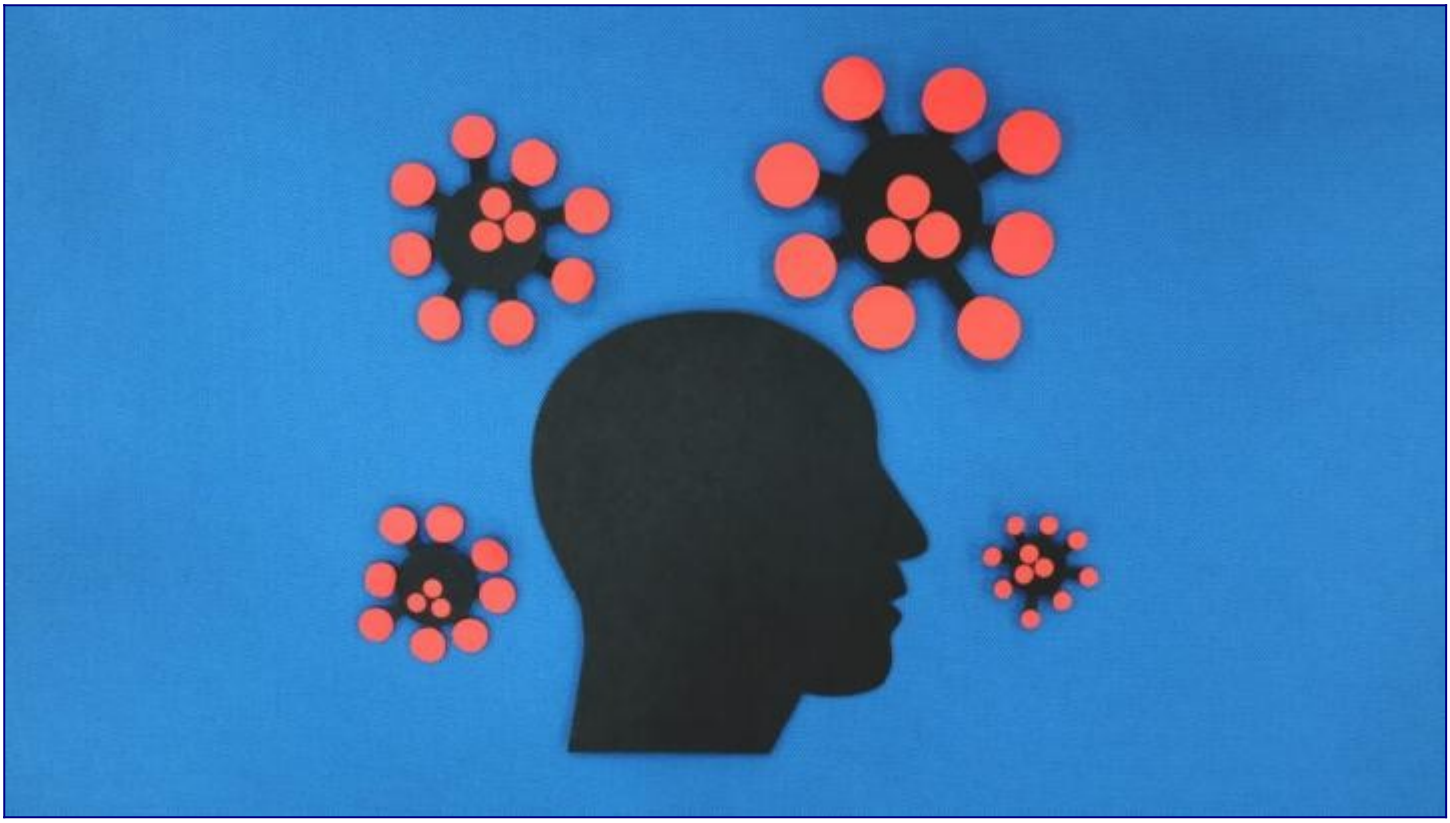
Shared by Romain Brunel

We need to start thinking more critically about long Covid - STAT

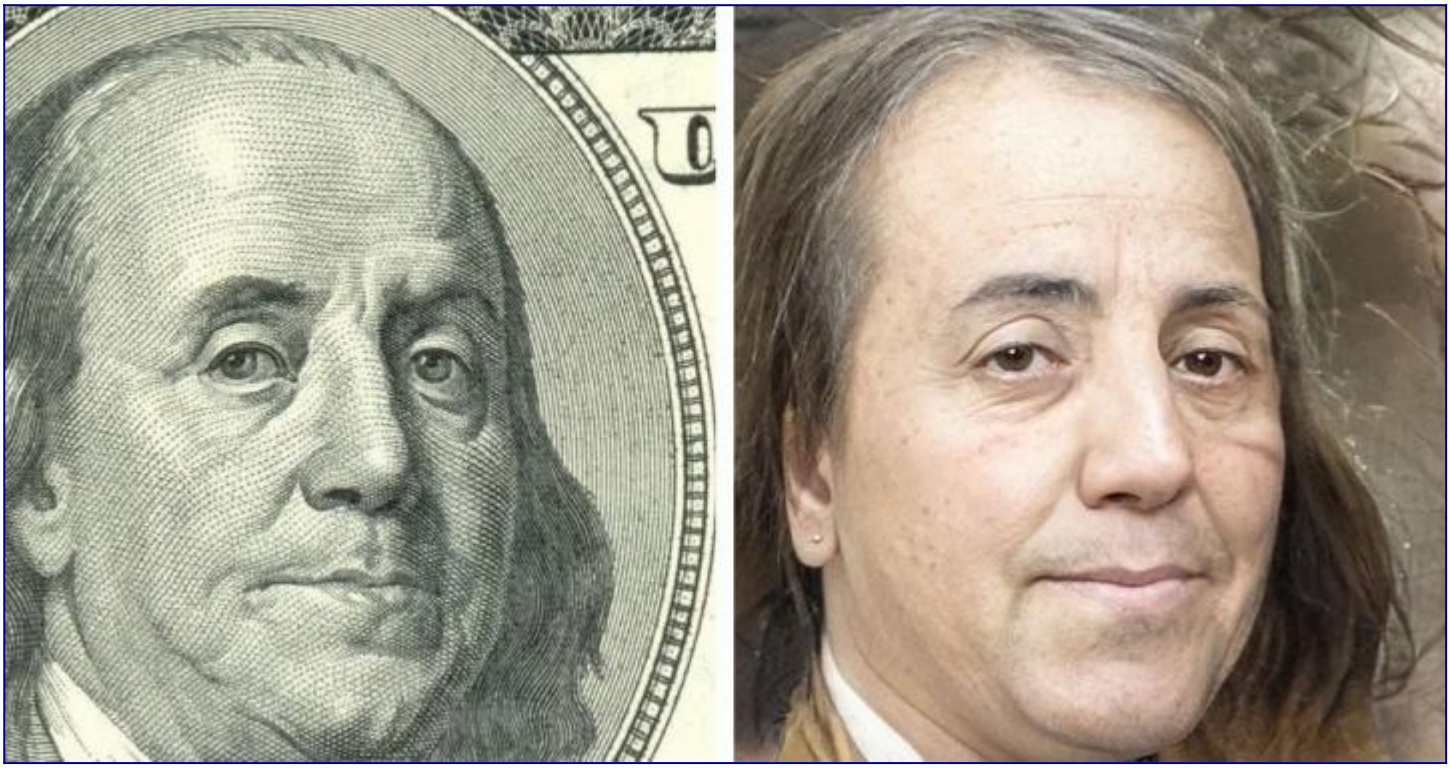
statnews.com

The narrative about long Covid, a disease linked to the SARS-CoV-2 virus, and about long-haulers may be getting ahead of the evidence.

Shared by Vinay Prasad MD MPH



[Artist Uses AI To Recreate How Famous Historical Figures Would Look Like in Real Life](mymodernmet.com)
mymodernmet.com
[San Francisco-based artist Nathan Shipley uses AI technology to create realistic portraits of famous historical figures.](mymodernmet.com)
Shared by C. Michael Gibson MD



Cautionary Tales – Catching a Killer Doctor
timharford.com
A doctor who killed hundreds of patients left us clues… but why couldn't we see them? Family doctor Harold Shipman got away with murdering his patients for decades. He was one of the mo…
Shared by David Spiegelhalter

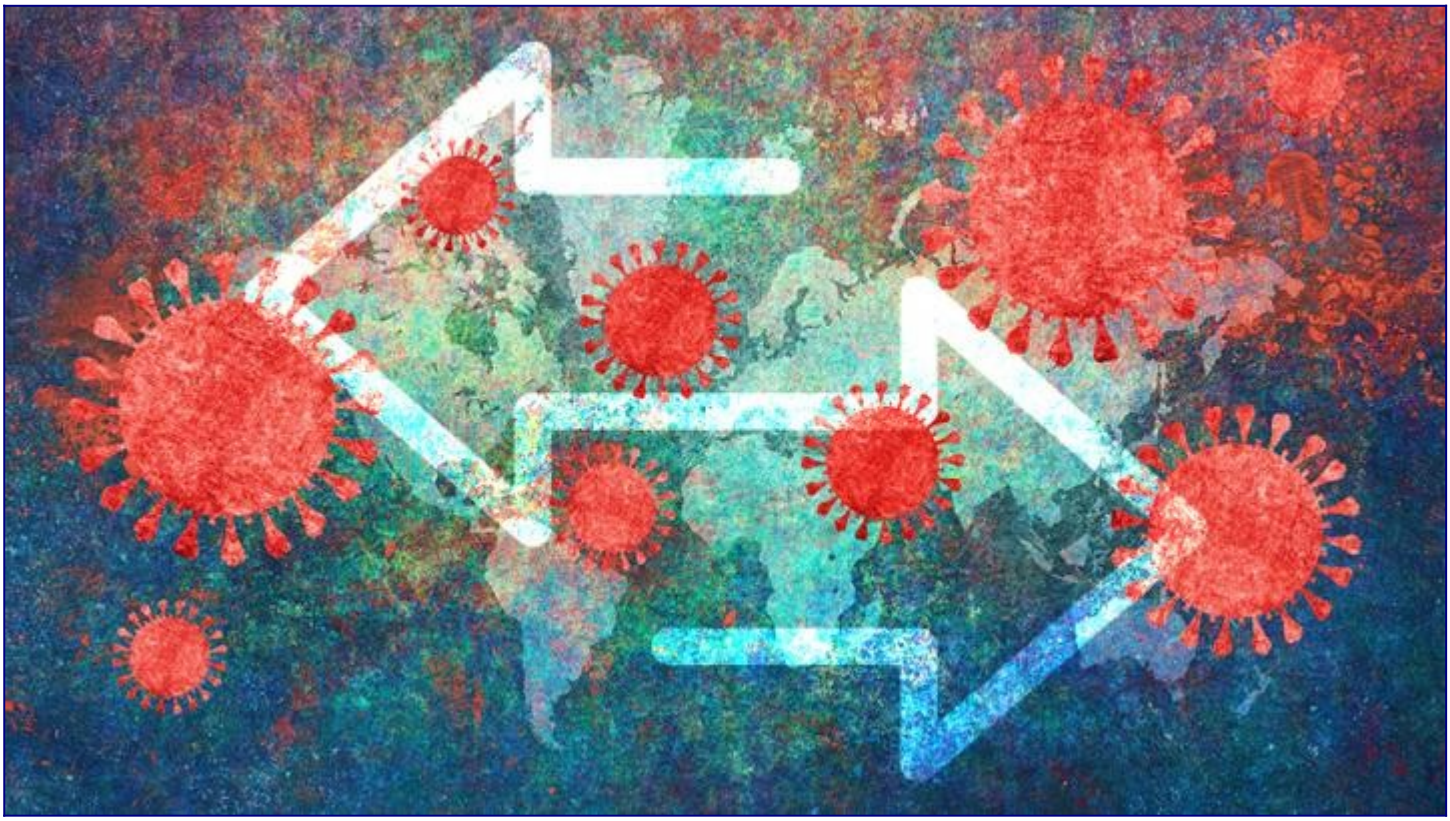
Let's hear scientists with different Covid-19 views, not attack them - STAT
statnews.com
Some scholars and academics wrongly ascribe scientists' legitimate disagreements about Covid-19 to ignorance or questionable political or other motivations.
Shared by Vinay Prasad MD MPH

Op-Ed: #ZeroCOVID or #HarmReduction?

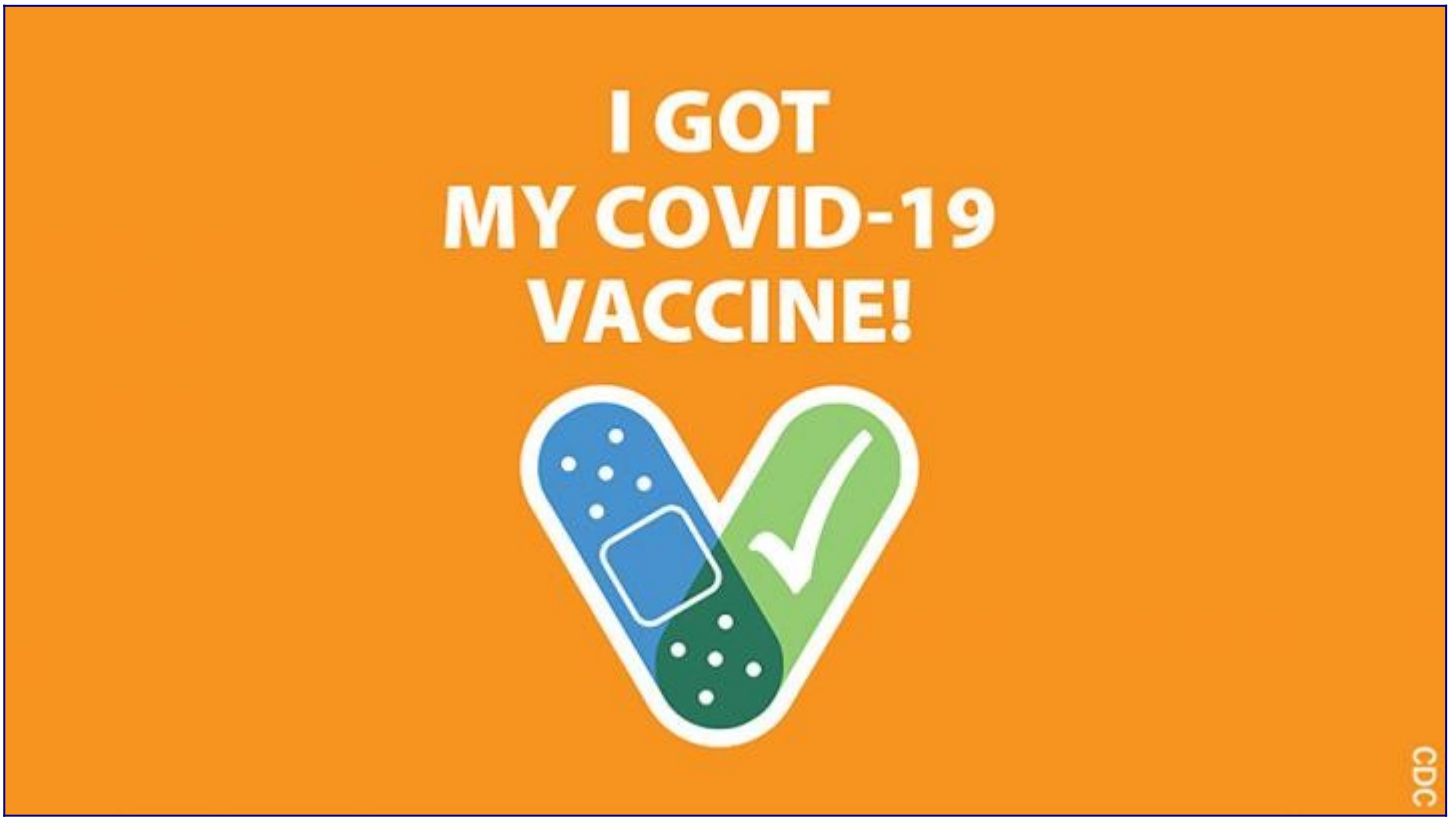
medpagetoday.com

Vinay Prasad weighs the two competing targets

Shared by Vinay Prasad MD MPH



Fully Vaccinated and Time to Party, If You Are 70
nytimes.com
Bobby Stuckey flipped through receipts this month, surprised to see a huge increase in cocktail sales, the highest in the 17-year history of his restaurant, even though the bar section...
Shared by Vinay Prasad MD MPH
Op-Ed: Tribalism and Virtue Signaling in Post-COVID Vax Messaging
medpagetoday.com
No one is chasing a zero-risk life
Shared by Vinay Prasad MD MPH



Understanding Covid-19, a 'supernova in human history' - STAT
statnews.com
Understanding the full impact of Covid-19 will take the work of economists, epidemiologists, historians, philosophers, sociologists, and many others.
Shared by Vinay Prasad MD MPH

[Facebook: A Worthy Judge of Medical Info?]
[medpagetoday.com]
[The social media giant's fact-checkers are plucked from a constellation of Twitter stars]
Shared by [Vinay Prasad MD MPH]

Op-Ed: Throw Away Your Mask After COVID Vaccination?
medpagetoday.com
Vinay Prasad says it's probably safe
Shared by Vinay Prasad MD MPH

Scientists Should Admit They Bring Personal Values to Their Work
scientificamerican.com
Scientific American is the essential guide to the most awe-inspiring advances in science and technology, explaining how they change our understanding of the world and shape our lives.
Shared by Retraction Watch

[Columnist Issues Lengthy Apology For Tweets Defaming Journalist as Pedophile-Worshipper](#)
mediaite.com
Julie Burchill, a columnist for the UK's Sunday Telegraph, posted a lengthy and fairly humiliating retraction letter after smearing a fellow journalist as a pedophile-worshipping Islamist.
Shared by Retraction Watch

The Absurdity of University Rankings
blogs.lse.ac.uk
University rankings are imbued with great significance by university staff and leadership teams and the outcomes of their ranking systems can have significant material consequences. Drawing on a cu…
Shared by Retraction Watch

In French woods, rivals take aim at senator's WWI research

apnews.com

HARRISBURG, Pa. (AP) — The World War I exploits of Sgt. Alvin C. York netted Gary Cooper a best actor Academy Award and Pennsylvania state Sen. Doug Mastriano a degree, a book deal — and academic...

Shared by Retraction Watch

Indonesia should stop pushing its academics to chase empty indicators
asia.nikkei.com
Prodita Sabarini is executive editor of The Conversation Indonesia, a nonprofit online media that brings together academics and journalists to produce
Shared by Retraction Watch

[“Riddled with errors”: Study of cell phones and breast cancer retracted](#)
retractionwatch.com
[via Wikimedia A journal has retracted a study that said exposure to radiofrequency radiation increased the risk of breast cancer after an epidemiologist found that some of the papers it relied upon d…](#)
Shared by [Retraction Watch](#)

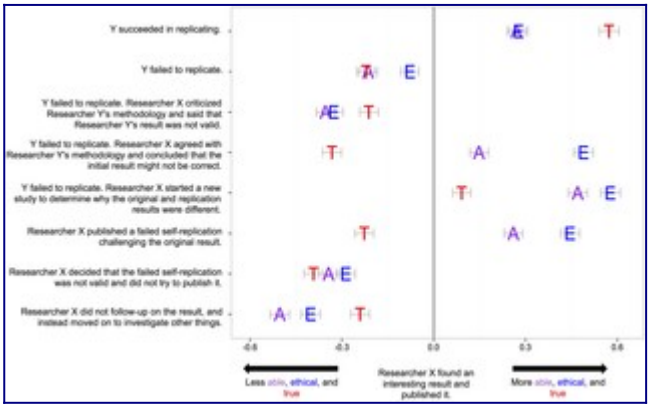
[Opinion: Preprints in the Public Eye](#)
the-scientist.com
[ASAPbio has developed resources for preprint servers, institutions, scientists, and journalists to promote the responsible reporting of research in the media.](#)
Shared by [Retraction Watch](#)

Meet the medical resident who had his wife peer review five of his papers
retractionwatch.com
Husband and Wife road (Beauty and Fashion) wedding,husband,wife,couple,marriage The pantheon of husband-wife teams in science includes Marie and Pierre Curie, Gerty and Carl Cori, even Ugur Sahin and…
Shared by Retraction Watch

Watch John Entwistle's Bass-Playing Genius on Display in Isolated Tracks for "Won't Be Fooled Again" and "Baba O'Reilly"
openculture.com
I guess it's easy to be "The Quiet One" in The Who when surrounded by a preening singer with golden locks, a guitarist with a windmill arm who smashes his equipment, and a completely insane drummer (…
Shared by Arthur Charpentier

Can artificial intelligence combat wildfires? Sonoma County tests new technology
latimes.com
Sonoma County is adding artificial intelligence to its wildfire-fighting arsenal
Shared by Arthur Charpentier

The importance of stupidity in scientific research
jcs.biologists.org
I recently saw an old friend for the first time in many years. We had been Ph.D. students at the same time, both studying science, although in different areas. She later dropped out of graduate schoo…
Shared by Brian Nosek
Scientists' Reputations Are Based on Getting It Right, Not Being Right
journals.plos.org
How are scientists evaluated? Members of the general public, undergraduate students, and active researchers all agree that the research process, not the results, determines a scientist's reputation.
Shared by Brian Nosek



Adaptive sample size determination for the development of clinical prediction models
diagnprognres.biomedcentral.com
We suggest an adaptive sample size calculation method for developing clinical prediction models, in which model performance is monitored sequentially as new data comes in. We illustrate the approach …
Shared by Maarten van Smeden

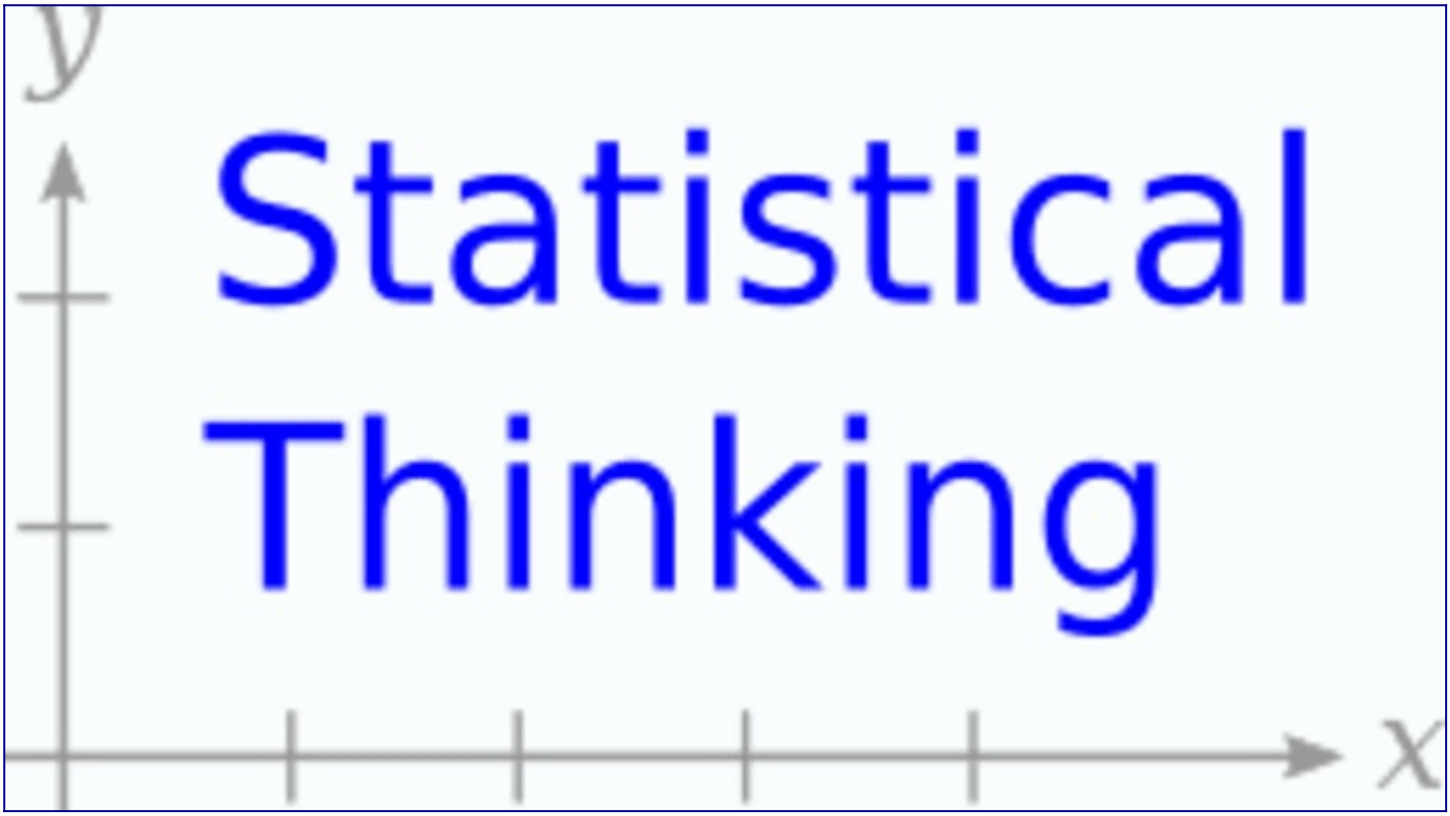
Eating one rasher of bacon a day 'increases risk of dementia by 44%'
dailymail.co.uk

Study suggests that eating just one rasher of bacon a day could increase the chances of getting dementia by 44 per cent. Leeds University research explored the potential link in data of 500,000 peopl…
Shared by Cecile Janssens

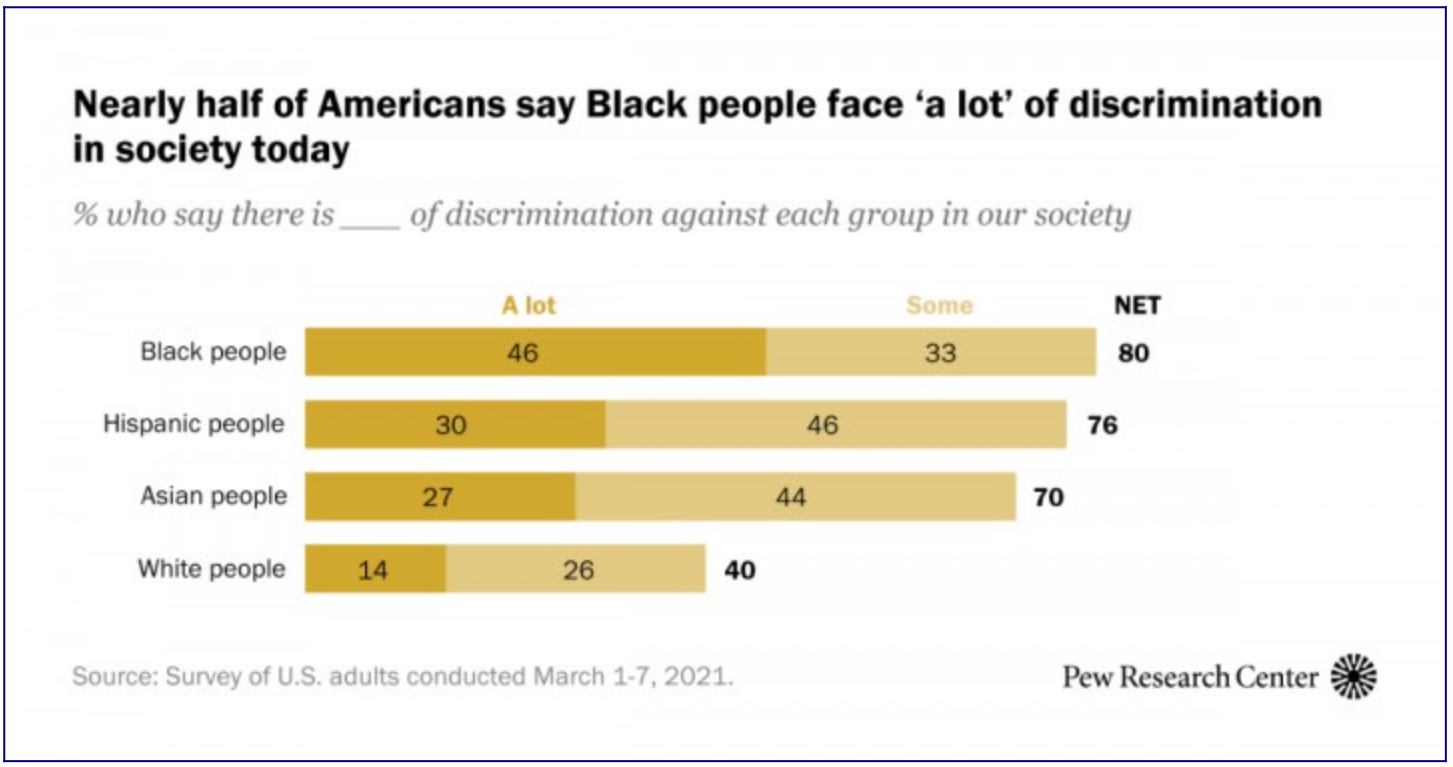
Prevalence of Central Nervous System–Active Polypharmacy Among Older Adults With Dementia in the US
jamanetwork.com
This pharmacoepidemiology study uses Medicare claims data to describe the prevalence of central nervous system–active polypharmacy among community-dwelling olde
Shared by WesElyMD

Recommended Recent Articles in Statistics Journals
hbiostat.org
These are articles that apply to biomedical research, from 14 statistical journals, primarily published online in the past 60 days.
Most of the articles are selected because they appear to be applicable to real problems. A few are selected to critique.
Shared by Frank Harrell

Statistical Thinking Blog
fharrell.com
A blog devoted to better use of statistical methods to extract more information from data, and to critique the status quo.
Shared by Frank Harrell



First CANSSI-NISS Health Data Science Workshop - FULL PROGRAM
niss.org
THURSDAY, MAY 6 [Please Note: All times are Eastern Time] [Link back to Event page] Short Course Registration $35 registration per short course. Select the registration options on the right-hand side…
Shared by NISS
Majorities of Americans see at least some discrimination against Black, Hispanic and Asian people in the U.S.
pewresearch.org
Americans are much less likely to say there is discrimination against White people: 40% say White people face at least some discrimination.
Shared by Boback Ziaeian



[Whiskey and Cream: Episode XIII - Ari Shapiro]
arishapiro.ca
[Whiskey & Cream for March 20th, 2021. Host: Ari Shapiro 0:45-16:12: "Lincoln would be proud" By now you've probably heard about that scrappily intrepid American political action co…]
Shared by [Allen Frances]

[They Died Saving Others From Covid. Will Anyone Count Them?]
nytimes.com
[Medical workers are called heroes. But there hasn't been a national reckoning over the many thousands lost to Covid. Here are a few of the people who gave their lives while on the front lines of the …]
Shared by [Edward Tufte]

Who is Camille Noûs, the fictitious French researcher with nearly 200 papers?
sciencemag.org
Group invents character as form of protest, but ethicists say campaign is misguided
Shared by Retraction Watch

After 35 years, philosophy journal corrects article…by a cat
retractionwatch.com
In 1982, Bruce Le Catt wrote a response to a paper in the Australasian Journal of Philosophy critiquing an earlier article about prosthetic vision. But Le Catt was no ordinary author. No, he was a ca…
Shared by Retraction Watch

MiR-1290 promotes myoblast differentiation and protects against myotube atrophy via Akt/p70/FoxO3 pathway regulation

skeletalmusclejournal.biomedcentral.com

Sarcopenia is a common skeletal disease related to myogenic disorders and muscle atrophy. Current clinical management has limited effectiveness. We sought to investigate the role of miR-1290 in myobl…

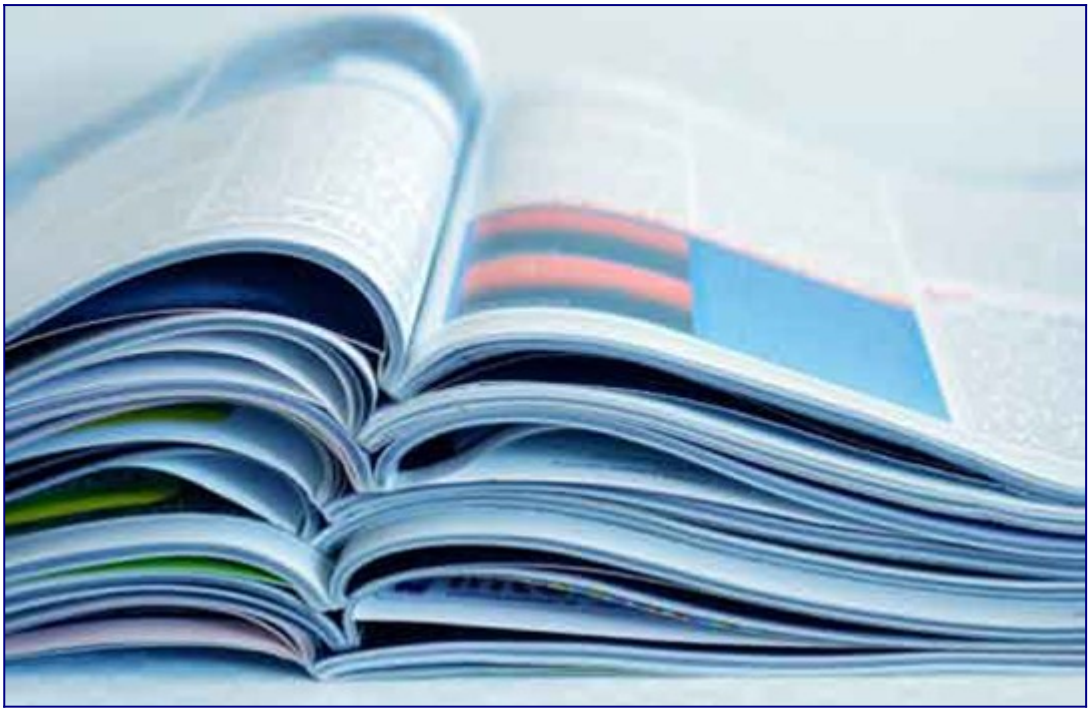
Shared by David J Glass MD

Alex Nowbar’s journal reviews—19 March 2021 - The BMJ

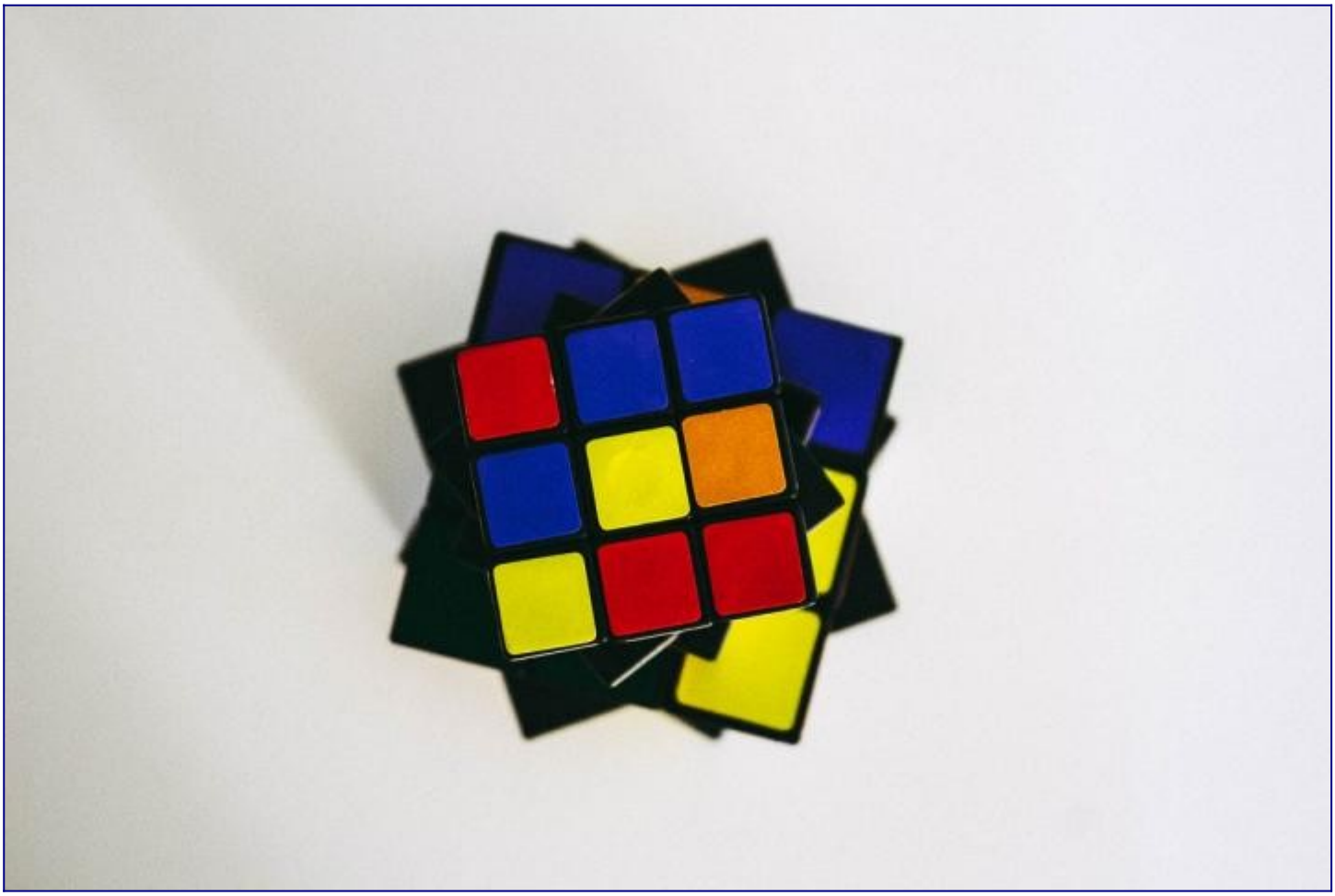
blogs.bmj.com

Alex Nowbar reviews the latest research from the top medical journals Thou shalt have a fishy Cardiovascular disease is less than desirable, but it does get tiresome hearing about one [...]More...

Shared by Alex Nowbar

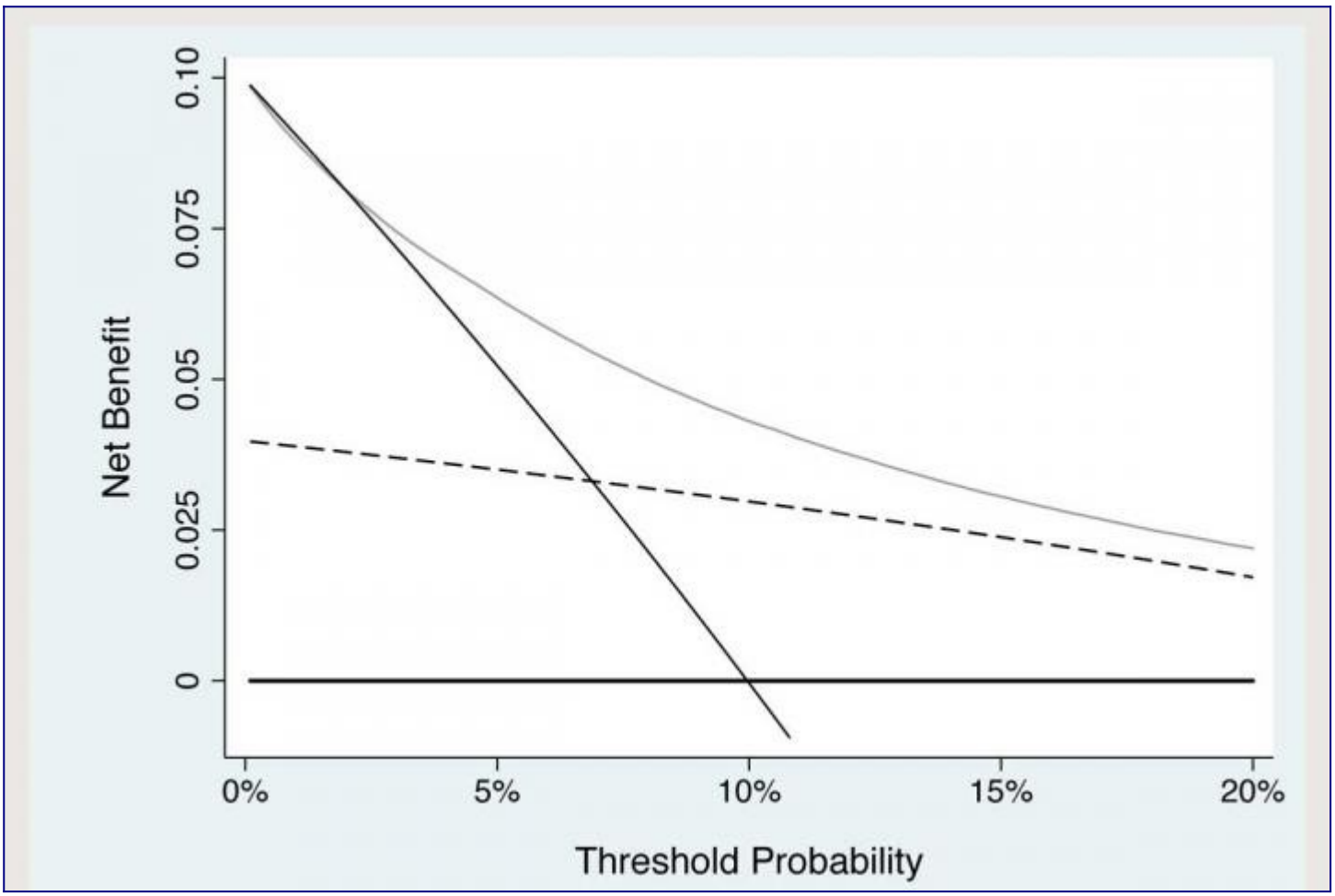


[How can Machine Learning algorithms include better Causality?]
medium.com
In recent years, machine learning algorithms have known a great success. Thanks to the availability of important amount of data and the increase of computation speed, they have outperformed usual…
Shared by 7wdata



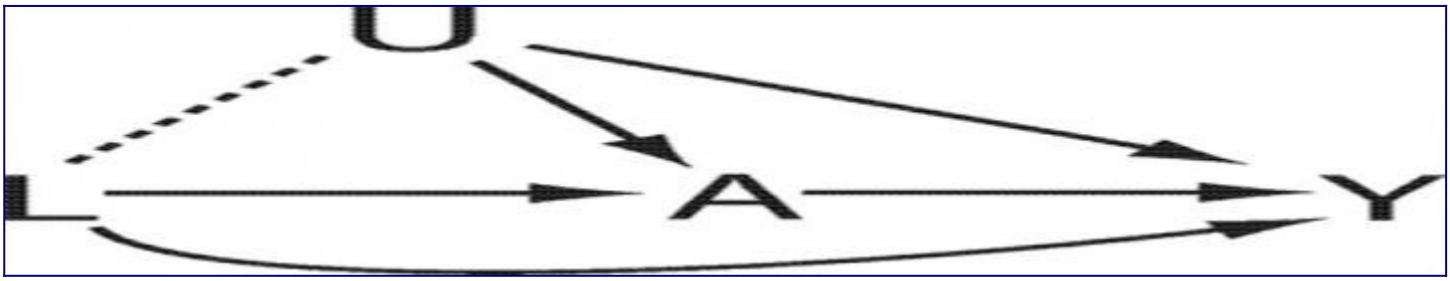
[Proptech Lets CRE Companies Fly Virtually Through Data-Decorated Properties]
bisnow.com
Use of AI and GIS technologies like Liquid Galaxy, Placer AI, Place Intelligence and PowerLeads AI have become a strategy for efficient CRE decision-making.
Shared by Nareit

UAB's Learning Health System Month - School of Medicine - Center for Outcomes and Effectiveness Research and Education
uab.edu
Russell Rothman, MD, MPP Professor, Internal Medicine, Pediatrics, & Health Policy Ingram Chair in Integrative and Population Health Senior Vice President, Population and Public Health Director, Inst…
Shared by UABCOERE
Decision Curve Analysis
mskcc.org
View the Biostatistics page for Decision Curve Analysis.
Shared by Matt Tenan, PhD ATC



Causal Inference from Observational Data
hsph.harvard.edu
Try explaining to your extended family that you are considered an expert in causal inference. That's why, when people ask, I just say that my job is to learn what works for the prevention and…
Shared by Christoph Kronenberg

[Negative Controls: A Tool for Detecting Confounding and... : Epidemiology]
journals.lww.com
[Abstract: Noncausal associations between exposures and outcomes are a threat to validity of causa]
Shared by [杏林大学医学部高齢医学]

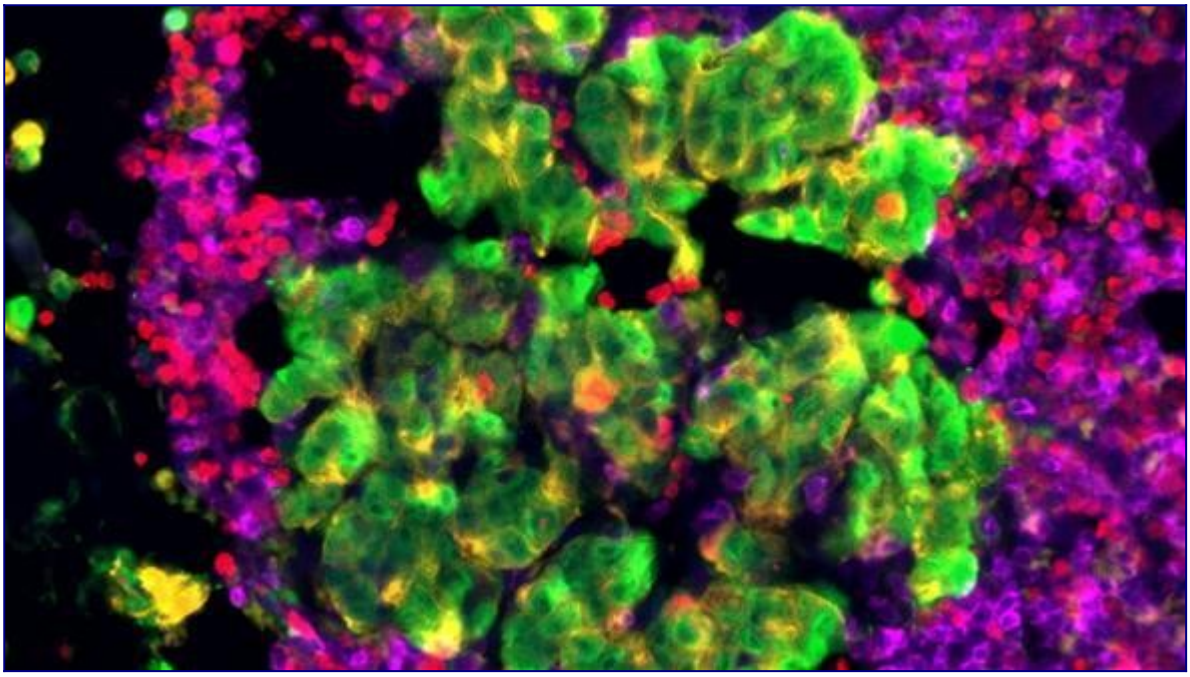


[4 unusual ways to explore data with your classes (sponsored)]
tes.com
[The weather offers a perfect way in to teaching about data, enabling you to cover all the key skills in one tangible topic. The weather forecasts we come across every day rely on gathering, processin…]
Shared by [Met Office Learning]

[Serum Apolipoprotein A-I Predicts Response of Rectal Cancer | CMAR]
dovepress.com
Serum apolipoprotein A-I emerged to negatively correlate with Rectal Cancer response to Neoadjuvant Chemoradiotherapy

Dovepress

[Science Talk - How data modelling will transform prostate cancer trea]
icr.ac.uk
Our researchers have been at the forefront of major clinical trials that have dramatically improved prostate cancer treatment. Henry French explains how we now plan to use statistical modelling metho…



Spotlight Scholar | Research Next
umass.edu
Within biostatistics, Balzer focuses on causal inference, which aims to isolate cause-and-effect from other biasing factors. Says Balzer: “In my first course on causal inference, I saw how statistics…

Some Observations On the Efficacy of Masks in a #COVID19 World
wattsupwiththat.com
Masks mandates are not a risk free intervention. They have a poor effect of civil society, they absorb resources, they possibly carry health risks of their own, and they certainly contribute to mis…

Shared by Diane Lange

Prevalence of RT-PCR-detected SARS-CoV-2 infection at schools: First results from the Austrian School-SARS-CoV-2 Study
medrxiv.org
medRxiv - The Preprint Server for Health Sciences
Shared by Sophie VM
Mapping the human genetic architecture of COVID-19 by worldwide meta-analysis
medrxiv.org
medRxiv - The Preprint Server for Health Sciences
Shared by Nicolas Arning
100 Best Coursera Courses & Specializations 2021 | JA Directives
jadirectives.com
Are you looking Best Free Coursera Courses 2021? Grab the list of Best Coursera Specializations & Certifications online in one place
Shared by Liam Smith
N95 Respirators vs Medical Masks for Preventing Influenza Among Health Care Personnel: A Randomized Clinical Trial
jamanetwork.com
Question Is the use of N95 respirators or medical masks more effective in preventing influenza infection among outpatient health care personnel in close contact with patients with suspected respirato…

Shared by “let Ireland live”

Effect of High-Dose Zinc and Ascorbic Acid Supplementation vs Usual Care on Symptom Length and Reduction Among Ambulatory Patients With SARS-CoV-2 In…

jamanetwork.com

Question Do high-dose zinc, high-dose ascorbic acid, and/or a combination of the 2 reduce the duration of symptoms of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)? Findings In this ra…

Shared by Alp

Research integrity in the COVID-19 era: insights from Retraction Watch co-founder Ivan Oransky

thepublicationplan.com

Read Ivan Oransky’s take on the retraction landscape and his perspectives on how to tackle research integrity issues in the future.

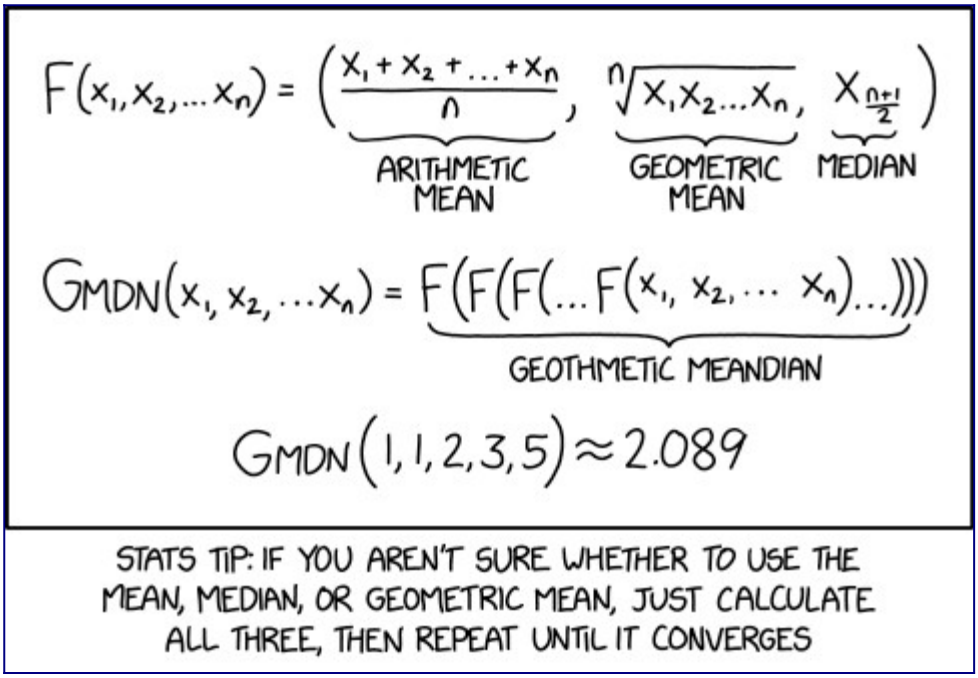
Shared by Retraction Watch

2435: Geothmetic Meandian

explainxkcd.com

Explain xkcd is a wiki dedicated to explaining the webcomic xkcd. Go figure.

Shared by Arthur Charpentier



Access to Care | Missouri Foundation for Health
mffh.org
The inability to access adequate care has profoundly negative consequences on the health and well-being of our state's most vulnerable groups. Access to care is an important consideration in nearly a…
Shared by Dr Ellie Murray

José Baselga, acclaimed drug hunter, physician and AstraZeneca's pioneering cancer R&D chief, has died
endpts.com
José Baselga, the brilliant oncology R&D chief at AstraZeneca and a towering figure in cancer drug development who had earlier been chief medical officer at Memorial Sloan Kettering, has died at the …
Shared by C. Michael Gibson MD

Married 66 years, husband, wife die minutes apart of virus

boston.com

“It is so precious, so wonderful, such a heartwarming feeling to know they went together,” Milewski said, then adding, “I miss them.”

Shared by C. Michael Gibson MD

Driven by the pandemic, applicants flood public health schools - STAT
statnews.com
Despite its challenges, public health and medical professionals see emerging signs that Covid-19 will inspire the next generation.
Shared by Edward Tufte

Is preprint the future of science? A thirty year journey of online preprint services
arxiv.org
Preprint is a version of a scientific paper that is publicly distributed preceding formal peer review. Since the launch of arXiv in 1991, preprints have been increasingly distributed over the Interne…
Shared by Retraction Watch

A way out of the replication crisis
review.chicagobooth.edu
Researchers argue that instead of overhauling the current academic publishing system, journals could improve it by making more explicit the criteria they use to select papers.
Shared by Retraction Watch

JAMA Editor Resigns Over Controversial Podcast
medscape.com
Edward H. Livingston, MD, has resigned as deputy editor of JAMA after he and the journal faced significant backlash over a February podcast that questioned the existence of structural racism.
Shared by Retraction Watch

[Police warn students and universities against using 'the Pirate Bay of science'](#)
news.sky.com
[Sci-Hub has been described as "the Pirate Bay of science", but often receives praise for opening access to research.](#)
Shared by [Retraction Watch](#)

What I learned about scientific misconduct from reading the NSF OIG’s semiannual reports
dynamicecology.wordpress.com
The National Science Foundation's Office of the Inspector General (NSF OIG) investigates misuse of NSF funds by individuals and organizations that receive awards from or conduct business with NSF. Th…
Shared by Retraction Watch

Whassup with the haphazard coronavirus statistics? « Statistical Modeling, Causal Inference, and Social Science
statmodeling.stat.columbia.edu
This piece by Robinson Meyer and Alexis Madrigal on the inadequacy of Covid data is useful but frustrating. I think they could have dispensed with the self-puffery, rhetoric and sweeping generalizati…
Shared by Andrew Gelman

Positive Coronavirus Test? Canadians Worry Their Neighbors Will Find Out
nytimes.com
The fear of public shaming is becoming so prevalent in some Canadian provinces that doctors worry it is driving virus cases underground.
Shared by Arthur Charpentier

COVID Experts: We’re Putting Out Campfires but the Forest Fire Rages
nautil.us
 “There is light at the end of the tunnel,” Wayne Koff, President and CEO of the Human Vaccines Project, an organization…
Shared by Arthur Charpentier

Platform Capitalism, Empire and Authoritarianism: Is There a Way Out?

cigionline.org

The world will face more widespread and intensified surveillance, but this time it could be framed as something for our own good, for the good of humanity.

Shared by Arthur Charpentier

[What's Behind the Ivermectin-for-COVID Buzz?]
[medpagetoday.com]
[Maverick physicians spurn randomized trials while 'people are dying']
Shared by [Arthur Charpentier]

Biden promised change at the border. He's kept Trump's Title 42 policy to close it and cut off asylum
latimes.com
The Biden administration says migrants can still get protection under a Trump-era pandemic policy. But in a year, fewer than 1% have been able to do so.
Shared by Arthur Charpentier

GOP warns HR 1 could be 'absolutely devastating for Republicans'
abcnews.go.com
Republican officials across the country fear the potential of H.R. 1.
Shared by Allen Frances

The Meadows, Edinburgh

en.wikipedia.org

The Meadows is historically common land and although now in the care of the council is technically in the ownership of the community itself. It was used for unhindered common grazing until at least 1…

Shared by Stephen John Senn

Opinion | How to Collect $1.4 Trillion in Unpaid Taxes

nytimes.com

Wealthy Americans are concealing large amounts of income from the I.R.S. There is a straightforward corrective.

Shared by Boback Ziaeian

Spring breakers are flocking to South Florida while some residents worry about another spike in Covid-19

cnn.com

(News Agents Network) – Veronica Pena says she spent the past year carefully navigating the internet Covid-19 pandemic in the mostly unrestricted state of Florida. The Miami Beach resident no longer …

Shared by Mike #MaskUp

Where Europe Went Wrong in Its Vaccine Rollout, and Why

nytimes.com

While Washington went into business with the drug companies, Europe was more fiscally conservative and trusted the free market.

Shared by Boback Ziaeian

Covid-19: Police Shut Down Late-Night Spring Break Parties in Miami
nytimes.com
International spectators have been barred at this summer’s Tokyo Olympics. Greece has reopened its ancient tourist sites, despite a virus surge.
Shared by Boback Ziaeian

Recast as ‘Stimmies,’ Federal Relief Checks Drive a Stock Buying Spree

nytimes.com
The government set out to prop up the economy. It may also be propping up the market.
Shared by Boback Ziaeian

‘It’s not a local issue anymore’: D.C. statehood moves from political fringe to the center of the national Democratic agenda
washingtonpost.com
On the eve of a House hearing, the District appears closer than ever to achieving its once-fanciful goal of becoming the 51st state. Yet statehood advocates still face major obstacles.
Shared by Boback Ziaeian

Statistical Thinking Blog
fharrell.com
A blog devoted to statistical thinking about medical research, data analysis, statistical modeling, and more.
Shared by Frank Harrell
Statistical Science: Some Current Challenges
hdsr.mitpress.mit.edu
A broad review is given of the role of statistical concepts in the design of studies, in various aspects of data collection and definition and especially in the analysis and interpretation of data. O…

Shared by (λ --- --. .. -.-.)³

Would you trust a machine to pick a vaccine?
review.chicagobooth.edu
Machine learning is being tasked with an increasing number of important decisions. But the answers it generates involve a degree of uncertainty.
Shared by रमण मूर्ति - आत्मनिर्भर भारत
Gait Speed and Survival in Older Adults
jamanetwork.com
Context Survival estimates help individualize goals of care for geriatric patients, but life tables fail to account for the great variability in survival. Physical performance measures, such as gait …
Shared by Nathan Stall
Open Rank Faculty Position (Biostatistics) | Center for Research on Health Care
gim-crhc.pitt.edu
POSITION ANNOUNCEMENT Open Rank Faculty Position (Biostatistics)
Shared by Pitt General Internal Medicine
Resources
statsof1.netlify.com
Actigraphy + Accelerometry Bayesian Approaches Causal Inference + Causality Clinical Research, Trials, Practice Community Connectivity Map Digital Phenotyping Dynamic Systems Effect Size and Tau-U (f…

Shared by Eric J. Daza, DrPH (Mask up, thx!)
Web-Based Lectures
amstat.org
Web-based lectures.
Shared by ASA
Target Validity and the Hierarchy of Study Designs
academic.oup.com
Abstract. In recent years, increasing attention has been paid to problems of external validity, specifically to methodological approaches for both quantitative
Shared by Elizabeth Arkema

-