How health science is done: the good, the bad and the ugly

Nerd Nite Tokyo — 2023年12月8日

Catherine BEAUCHEMIN

カトゥリン ・ ボシゥメン

Negative impressions about health research

- experts disagree / keep changing their minds

- experts have an agenda (profit, politics, social engineering, self-promotion, etc.)

Key issues with health research

- health research has for-profit components (pharma co., private hospitals/insurance)

- urgency to cure a disease conflicts with understanding how it works

- poor mathematical/statistical training in health sciences
  - disregarding data in favour of one's own expectation
  - studies designed to prove rather than disprove hypothesis
  - routinely major flaws in study design or analysis

International standards for clinical trials

In 2008, the World Medical Association's (WMA) Declaration of Helsinki stated:

- Clinical trials must be registered in public database before recruiting 1st subject

- All studies should be published, regardless of statistical significance of outcome
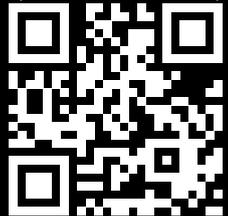
In 2015, World Health Organization (WHO) added:

- Main findings of clinical trials are to be published within 2 years of completion, key outcomes publicly available within one year in result section of clinical trial registry

Adherence to these 3 recommendations is **voluntary**.

ClinicalTrials.gov
(2015–2020)

	completed	w results	funded by industry
United States	20,102	7,462 (37%)	8,677 (43%)
Canada	3,247	1,055 (32%)	1,471 (45%)
United Kingdom*	3,746	1,071 (29%)	1,959 (52%)
Japan*	1,025	547 (53%)	945 (92%)

$\rightarrow$ **Lack of transparency reduces future participation & trust in results.**

Irregularities in registered clinical trials

When registering a clinical trial, you need to identify:

- therapy or intervention to be evaluated (e.g., drug)
- target population (e.g., sex, age, race or genetic traits)
- inclusion/exclusion criteria (e.g., people with other underlying diseases)
- **primary outcome** indicating success (e.g., reduction in # of pimples)
- (optionally) secondary outcomes (e.g., make hair more curly, lower itchiness score)

Moving the goalposts

A study of trials conducted in Germany, 2009–2017 [doi:10.1371/journal.pmed.1004306]:

- **8%** changed **primary outcome** bwn registration & latest entry.
- **18%** changed **primary outcome** bwn latest entry & paper (**1%** said so in paper).

Changing 'measure of success' = bias reporting + statistical need to account for multiple hypothesis testing → misconduct (can be caught by **peer-reviewer** thanks to registry)

Research misconduct → paper retractions

To get grants/job/promotion, researchers need to publish peer-reviewed papers
→ extra points for: more papers, fancier journals, more citations

Rank. Name	# retracted	Field	Country
01. J Boldt	194	**Medicine**	Germany
02. Y Fuji	172	**Medicine**	Japan
03. H Ueshima	124	**Medicine**	Japan
04. Y Sato	113	**Medicine**	Japan
05. A Nazari	100	Engineering	Iran
06. J Iwamoto	88	**Medicine**	Japan
07. D Stapel	58	Psychology	Netherlands
08. Y Saitoh	56	**Medicine**	Japan
09. A Maxim	48	Engineering	USA
10. AS Elahi	44	Engineering	Iran
11. CY Chen	43	Engineering	Taiwan
12. F Sarkar	41	**Medicine**	USA
13. S Shamshirband	41	Comp. Sci.	Malaysia
14. H Zhong	41	**Medicine**	China
15. S Kato	40	**Medicine**	Japan

From

Retraction Watch
Leaderboard

9/15 (60%) in Medicine! Sadly 6/15 (40%) from Japan, all Medicine.

Issues with the peer-review system

[Fauci] *So that [hydroxychloroquine] study is a flawed study and I think anyone who examines it carefully sees that **it is not a randomized, placebo-controlled trial**.*

[Luetkemeyer] *It's been peer-reviewed...*

[Fauci] *It doesn't matter, you can peer-review something that's a bad study.*

[Wait... Why?!?]

US House Oversight & Reform Select Subcommittee
on Coronavirus Crisis, July 31, 2020

Peer-reviewed results inform the standard of care

ORIGINAL ARTICLE

MAYO CLINIC

A Decade of Reversal: An Analysis of 146 Contradicted Medical Practices

Vinay Prasad, MD; Andrae Vandross, MD; Caitlin Toomey, MD; Michael Cheung, MD; Jason Rho, MD; Steven Quinn, MD; Satish Jacob Chacko, MD; Durga Borkar, MD; Victor Gall, MD; Senthil Selvaraj, MD; Nancy Ho, MD; and Adam Cifu, MD

Of the 363 articles testing standard of care, 146 (40.2%) reversed that practice, whereas 138 (38.0%) reaffirmed it.

therapy. This study was conducted from August 1, 2011, through October 31, 2012.

Results: We reviewed 2044 original articles, 1344 of which concerned a medical practice. Of these, 981 articles (73.0%) examined a new medical practice, whereas 363 (27.0%) tested an established practice. A total of 947 studies (70.5%) had positive findings, whereas 397 (29.5%) reached a negative conclusion. A total of 756 articles addressing a medical practice constituted replacement, 165 were back to the drawing board, 146 were medical reversals, 138 were reaffirmations, and 139 were inconclusive. Of the 363 articles testing standard of care, 146 (40.2%) reversed that practice, whereas 138 (38.0%) reaffirmed it.

Conclusion: The reversal of established medical practice is common and occurs across all classes of medical practice. This investigation sheds light on low-value practices and patterns of medical research.

Published by Elsevier Inc on behalf of Mayo Foundation for Medical Education and Research ■ Mayo Clin Proc. 2013;88(8):790-798

Standard of care
guidelines generally accepted by medical community in treatment of disease or condition.

Failure to meet it =
    medical negligence

Peer-reviewed results inform the standard of care

ORIGINAL ARTICLE

MAYO CLINIC

A Decade of Reversal: An Analysis of 146

E.g. The practice of implanting Gentamicin-collagen sponge to prevent infection following colorectal surgery, used in millions of patients worldwide since 1985...

A single-centre, randomized trial found a 70% decrease in surgical site infection with this practice.
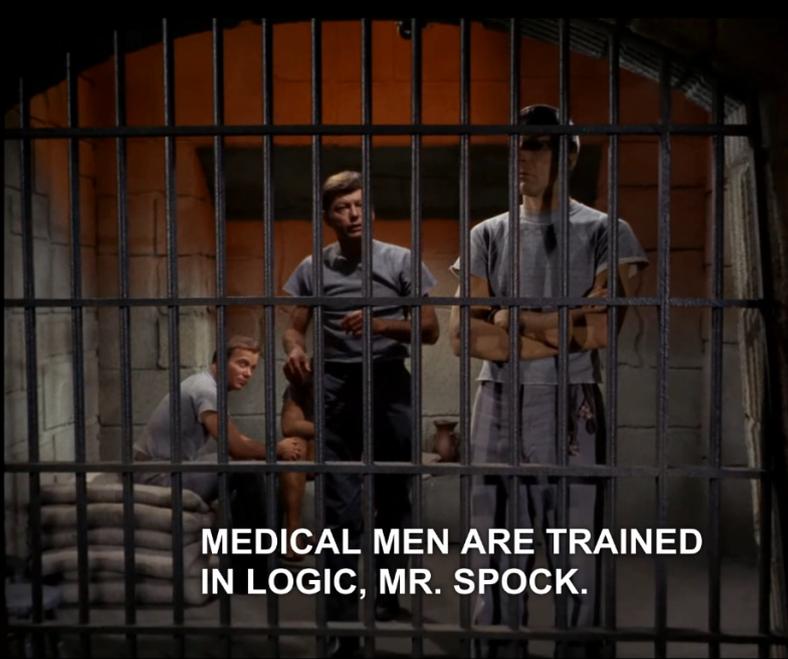
In a larger, multi-centre, phase 3 trial it resulted in significantly more infections, more visits to emergency departments, and more hospitalization for resulting infection.

Standard of care
guidelines generally accepted by medical community in treatment of disease or condition.

Failure to meet it =
    medical negligence

Or in the words of Mr. Spock to Dr. McCoy

MEDICAL MEN ARE TRAINED IN LOGIC, MR. SPOCK.

WATCHING YOU, I ASSUMED IT WAS TRIAL AND ERROR.

[Star Trek (Original Series), Season 2, Epis. 25 *"Bread and Circuses"* © 1968, Desilu Prod.]

Bar graphs are sneaky & p-values are misleading
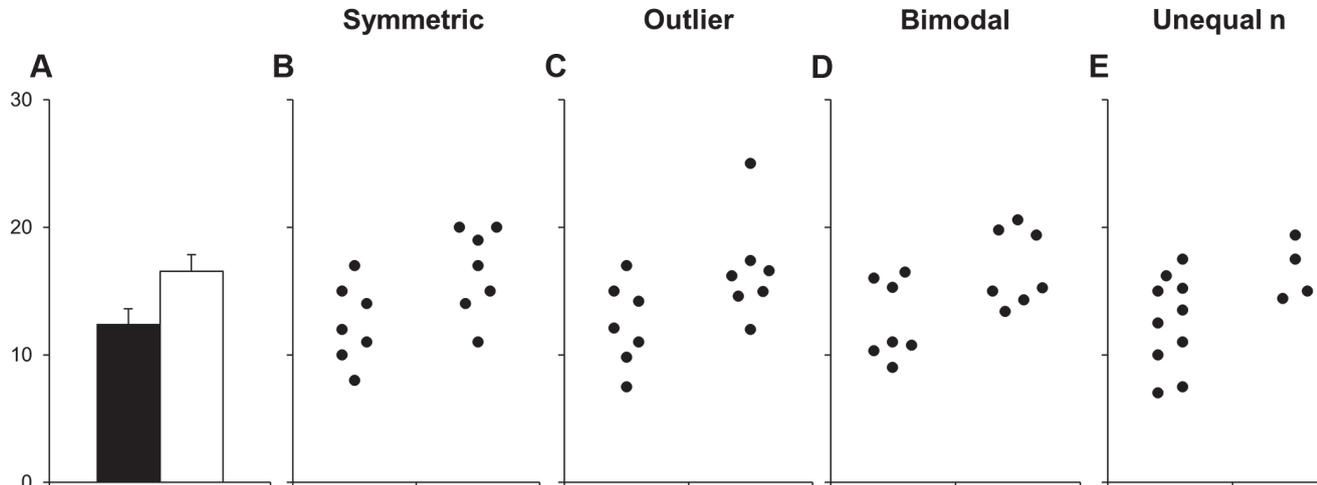
doi:10.1371/journal.pbio.1002128.g001

Bad graphs: when a picture is not worth 1,000 words

Figure from peer-reviewed article published in Nature Medicine (top 10 medical journal)
Primary endpoint: increase in [y-axis value] from Baseline to 12 months.
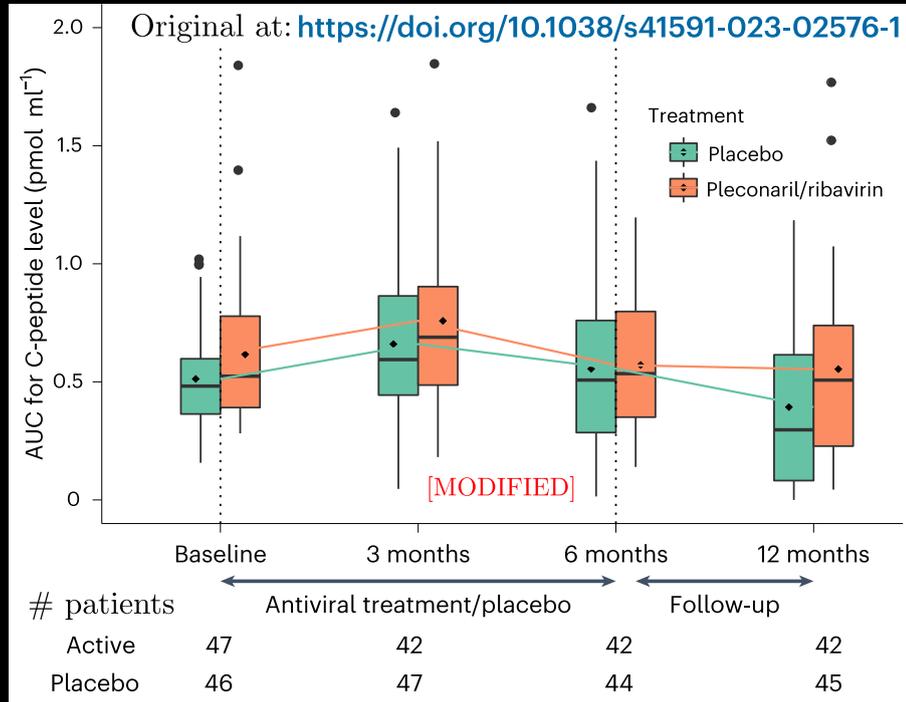
Key finding:
'[...]levels after 12 months were higher in [treatment] group compared with [placebo] (average marginal effect (AME)=0.057, p=0.037)'

Why is AME used rather than simply increase from baseline?

Do we expect this marginal change to result in significant health improvement?

Ongoing efforts to improve practices

Some encouraging signs of improvement
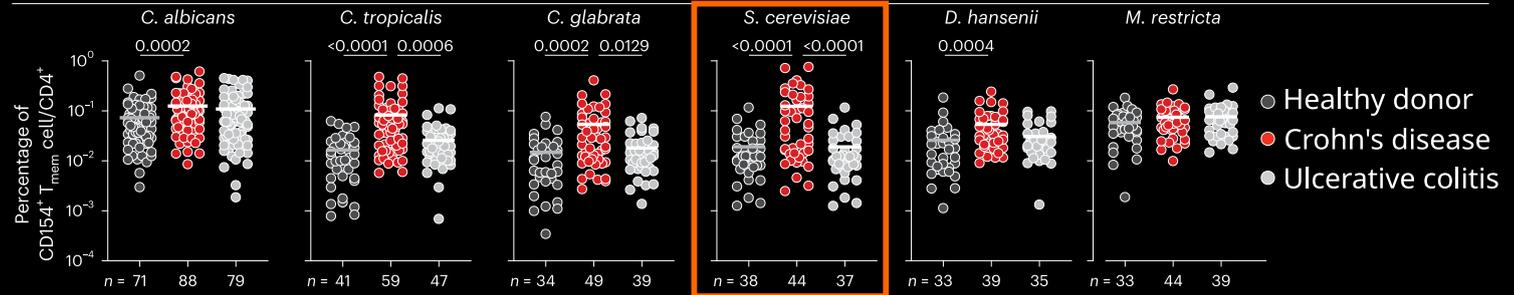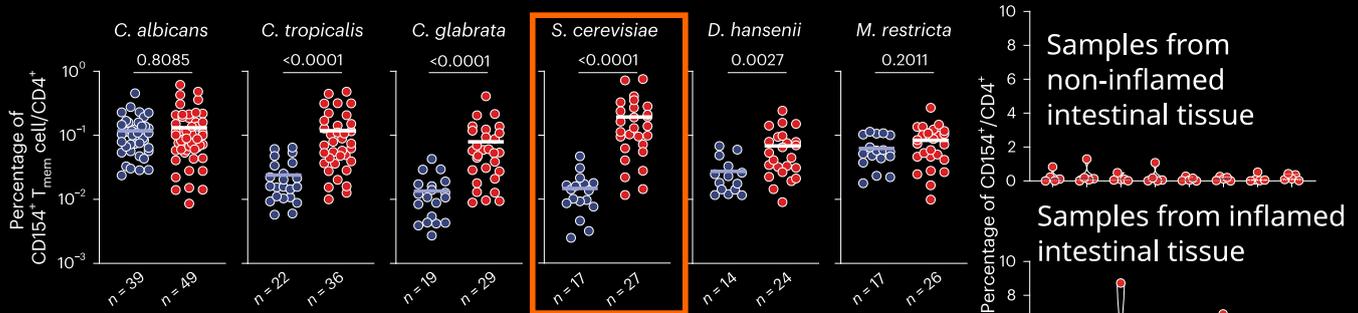
Selection of cross-reactive T cells by commensal and food-derived yeasts drives cytotoxic T_H1 cell responses in Crohn's disease

Main take home message

- Health research has issues: profit motive, reporting bias, misconduct, innumeracy

- I have seen encouraging progress over the last 5–7 years
  bar→dot plots, less p-value obsession, more honest titles/conclusions, **open data**

- There is still a long way to go...

Where does that leave you?

- Avoid thinking in absolute terms (vaccines are good/bad), be more moderate (some vaccines are essentials, others so-so)

- Avoid trusting results just because they agree with your expectations

The END.

Catherine Beauchemin (`cbeau@riken.jp`)
`https://phymbie.physics.ryerson.ca`