# Cognitive biases, errors, and questionable research practices with a focus on clinician-researchers or clinical research in general

UNIVERSITÄT BERN

# Outline

- The current clinical research (trial) landscape
  - Some issues with quality
  - Our current academic reward system

- Being a researcher
  - (Cognitive) biases
  - Human errors
  - Research misconduct
  - Questionable research practices

- Possible solutions for clinical research

# Disclosure

## Biases and conflicts of interests

# My biases

## with regard to the topic of this talk

UNIVERSITÄT BERN

- Not (so) young anymore (you can usually guess age by the level of logorrhea – so I feel very old …)
- I am male
- Father of two kids in an age where they can stay at home alone even in the evening
- Settled in Schmitten, FR
- I have an unlimited contract at the University of Bern
- Well-established professional network
- I am trained as physician but I am actually a methodologist

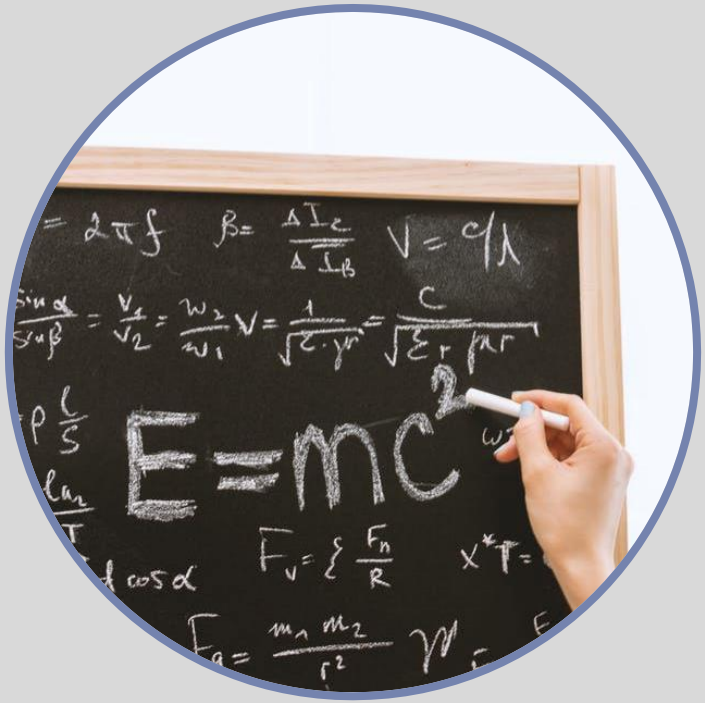
# My conflict(s) of interests

UNIVERSITÄT BERN

- CTU Bern depends on clients and people who seek help
- The more I stress specialization and expertise the better
- The more I downplay the role of the clinician-researcher the better for us

# Am I qualified?

## to talk about this topic



# Am I qualified?

## My academic career

UNIVERSITÄT BERN

- Average-talented (so far, nobody was honest/courageous enough to tell me, therefore: a self-reflection)
- I applied three times for a professorship (at least one serious application), all unsuccessful (including one in Bern)
- I work in academia for 15 years now (3 years in a clinical department)
- I have consulted on >500 clinical research projects, >50 grant applications, and have peer reviewed many

# A warning at the beginning

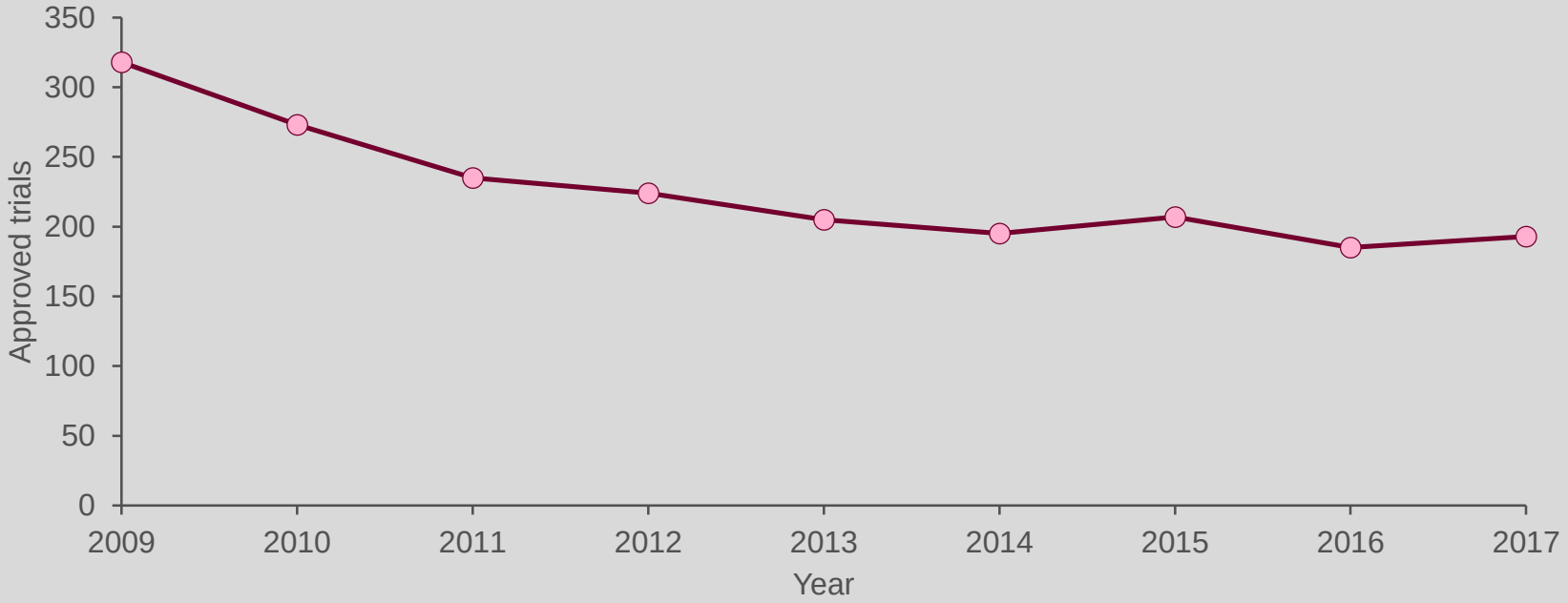
UNIVERSITÄT BERN

# The clinical research (trial) landscape

UNIVERSITÄT BERN

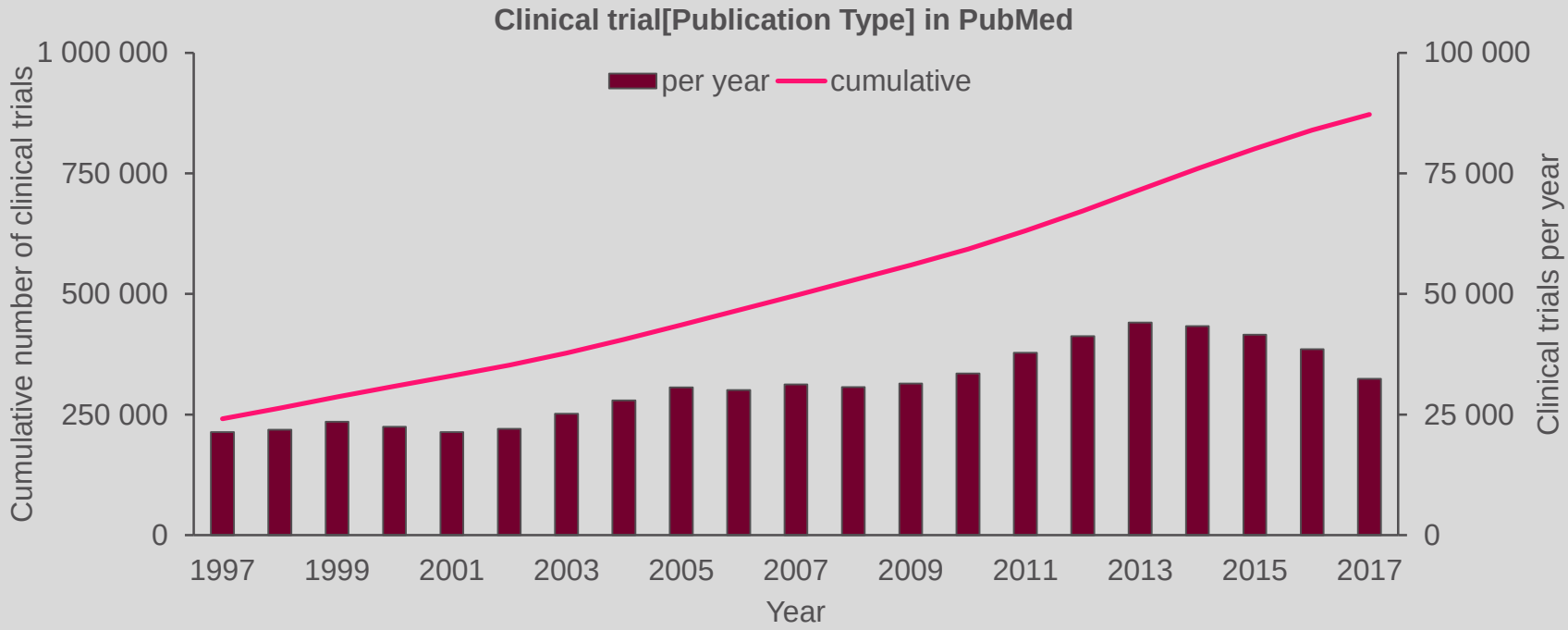
# Clinical trials approved by Swissmedic

## Pharmaceutical products and devices



# Global clinical trials

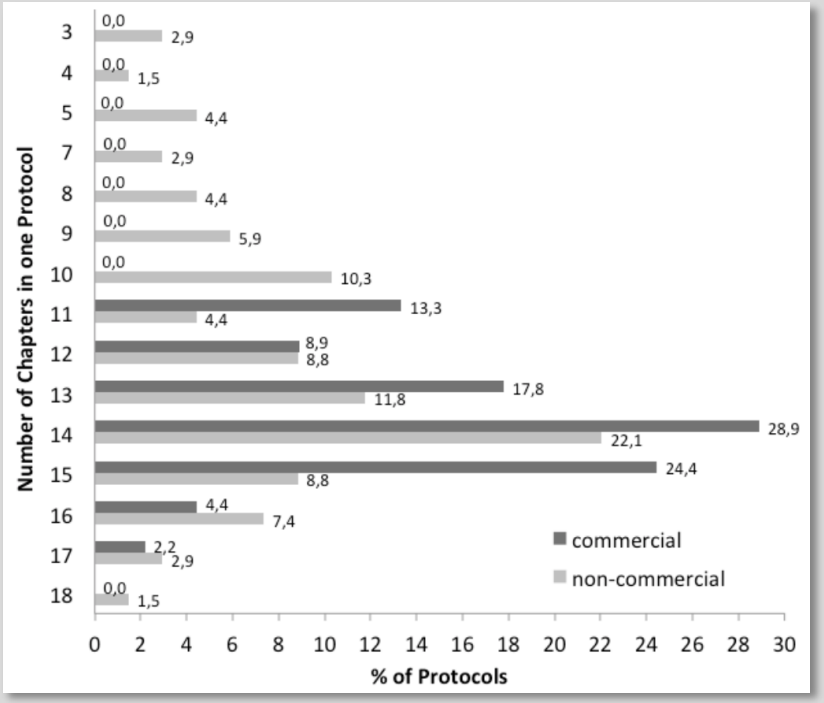
## Clinical trials published in MEDLINE



# Completeness of 113 approved protocols by Swiss ethics committees 2010-12

UNIVERSITÄT BERN

- Title
- Table of contents/ index
- Funding source for the complete study
- Background and rationale
- Objectives or hypothesis
- Trial or study design
- Trial or study setting
- Intervention/-s
- Outcome or endpoint/-s measures
- Sample size
- Randomization or assignment of intervention
- Statistical methods
- Safety or harms
- Ethical consideration
- Data protection
- Confidentiality
- Access to data
- Dissemination policy




*Weidmann R 2018*

# Concordance protocol-registry

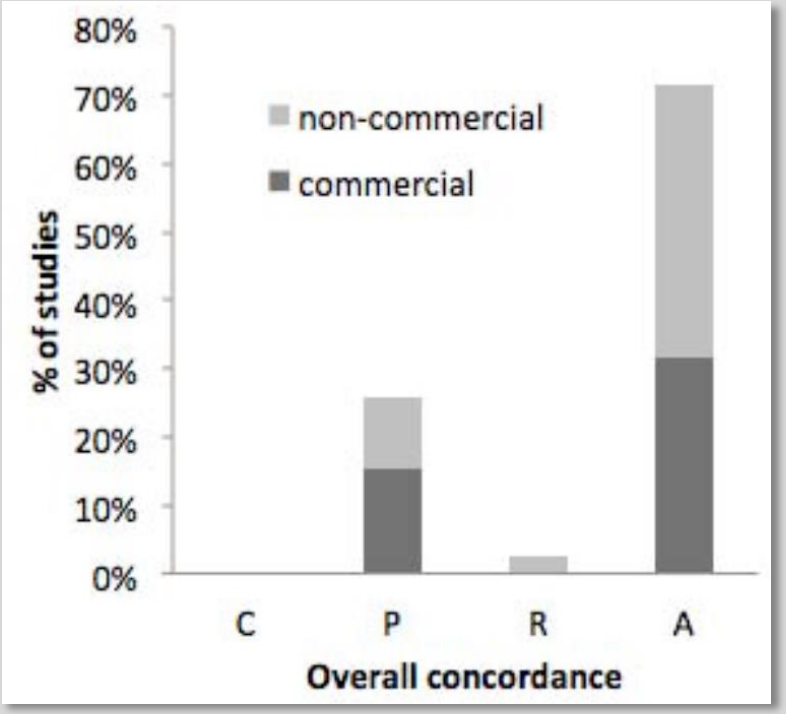
## 113 approved protocols

UNIVERSITÄT BERN

**Assessment**

Five domains (inclusion and exclusion criteria, intervention, sample size, primary outcome)

- **C**oncordant
- **P**rotocol contains more information
- **R**egistry contains more information
- non **A**lignment between the two




*Weidmann R 2018*

# Discrepant reporting of outcomes

## Cohort of Swiss clinical trials

UNIVERSITÄT BERN

- Omission from publication (non-reporting)
- Addition of an outcome
- Change from primary to secondary (or vice versa)

- 452 protocols approved by Swiss ethics committees (1988-98) screened
- 227 with at least 1 publication
- 333 publications used to extract outcomes (comparison with protocol)

*Redmond S et al. 2013*

# Extend of discrepant reporting

UNIVERSITÄT BERN

| | Primary outcome ($n$ = 274) | | |
|---|---|---|---|
| **Outcomes defined in protocols as** | **Classification[a]** | ***N*** | **Prevalence (95% CI)** |
| Reporting in published article | | | |
| Primary outcome | Concordant (1-1) | 165 | 60.2% (54.2, 66.1) |
| Secondary outcome | Discrepant (1-2) | 90 | 32.8% (27.3, 38.8) |
| Not reported | Discrepant (1-0) | 19 | 6.9% (4.2, 10.6) |
| | **Primary outcome ($n$ = 288)** | | |
| **Outcomes reported in publications as** | **Classification[a]** | ***N*** | **Prevalence (95% CI)** |
| Definition in study protocol | | | |
| Primary outcome | Concordant (1-1) | 175 | 60.8% (54.9, 66.4) |
| Secondary outcome | Discrepant (2-1) | 83 | 28.8% (23.7, 34.4) |
| Not defined | Discrepant (0-1) | 30 | 10.4% (7.1, 14.5) |

*Redmond S et al. 2013*

# COMPARE initiative

## www.compare-trials.org

UNIVERSITÄT BERN

### Tracking switched outcomes in clinical trials

Outcome switching in clinical trials is a serious problem. Between October 2015 and January 2016, the COMPare team systematically checked every trial published in the top five medical journals, to see if they misreported their findings:

1. We compared each clinical trial report with its protocol or registry entry. Some trials reported their outcomes perfectly. For the others, we counted how many of the outcomes pre-specified in the protocol or registry were never reported. We also counted how many new outcomes were silently added.
2. When we detected unreported or added outcomes, we wrote a letter to the journal pointing them out. We tracked which journals published our letters – and which did not.

Here's what we found.

| 67 | 9 | 300 | 357 |
|---|---|---|---|
| TRIALS CHECKED | TRIALS WERE PERFECT | OUTCOMES NOT REPORTED | NEW OUTCOMES SILENTLY ADDED |

On average, each trial reported just 62.1% of its specified outcomes. And on average, each trial silently added 5.3 new outcomes.

# Our (rotten?) academic reward system

UNIVERSITÄT BERN

- Research assessment exercise in Bern (up to 2017*)/habilitation
  - (Only) Articles published in high-impact journals are valued
  - (Only) first and last authorship really counts
  - Content and quality per se does not matter**
  - True collaboration is not valued (as their is no easy indicator?)
- Individual achievements count and ruthlessness is not a bad idea
- Ownership versus openness

* Although it has changed it is still in our minds and culture (only personal experience, though)

** I have nothing against quantitative indicators but they need to be sensible and appropriately interpreted (knowing what exactly they measure and their limitations)

# A possible consequence

## A recent example (N Engl Med 2018; 379: 1313-21)

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

# Phase 2 Trial of Selective Tyrosine Kinase 2 Inhibition in Psoriasis

Kim Papp, M.D., Ph.D., Kenneth Gordon, M.D., Diamant Thaçi, M.D., Ph.D., Akimichi Morita, M.D., Ph.D., Melinda Gooderham, M.D., Peter Foley, M.D., Ihab G. Girgis, Ph.D., Sudeep Kundu, Ph.D., and Subhashis Banerjee, M.D.

ABSTRACT

# Trial oversight

p. 1314f

UNIVERSITÄT BERN

*… The trial was sponsored by Bristol-Myers Squibb, which designed the trial, provided the trial drug and placebo, conducted blinded safety monitoring, developed the analysis plan, analyzed the results, and funded professional writing assistance. A contract research organization (ICON, Dublin) conducted the trial under the direction of the sponsor, and medical writers paid by the sponsor wrote the first draft of the manuscript. All the authors had full access to the trial data, reviewed and approved the manuscript before submission, and vouch for the adherence of the trial to the protocol, the completeness and accuracy of the data and analyses, and the reporting of adverse events. There were confidentiality agreements between the authors and the sponsor.*

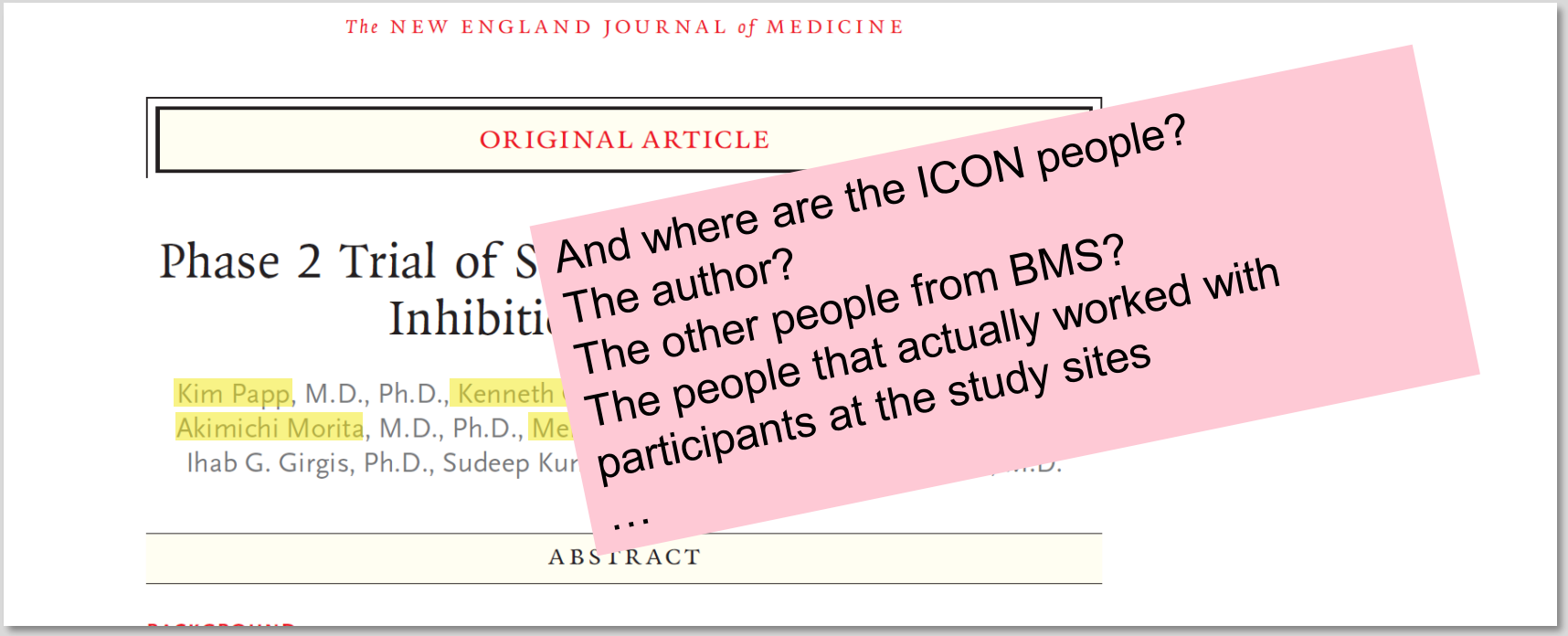
# International Committee of Medical Journal Editors authorship criteria

UNIVERSITÄT BERN

- Substantial contributions to conception and design, or acquisition of data, or analysis, or interpretation of data
- AND drafting the article or revising it critically for important intellectual content
- AND final approval of the version to be published
- AND agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved

# Let's remind us of the list of authors

## Non-industry authors yellow



# Being a researcher

# *characteristics good researcher*

## www.google.com

UNIVERSITÄT BERN

- 5 Qualities of a Good Researcher
  - manifests thirst for new information
  - keen sense of things around him
  - likes to reflect or think about the things he (sic!) encounters
  - must be intelligent enough to express his (sic!) ideas
  - applies a systematic approach in assessing situations

- (sic!) → implictly: must be a male

*https://simplyeducate.me/2012/10/24/5-qualities-of-a-good-researcher/*

# Good ≠ Success

# What do I think makes a good researcher?

Not necessarily a  successful, though!

# Cognitive biases

www.yourbias.is & www.yourlogicalfallacies.com

- Bystander effect
  - Someone else will take care …
- Framing
  - Undue influenced by context and delivery
- Availability heuristic
  - Undue influenced of what springs most easily to mind

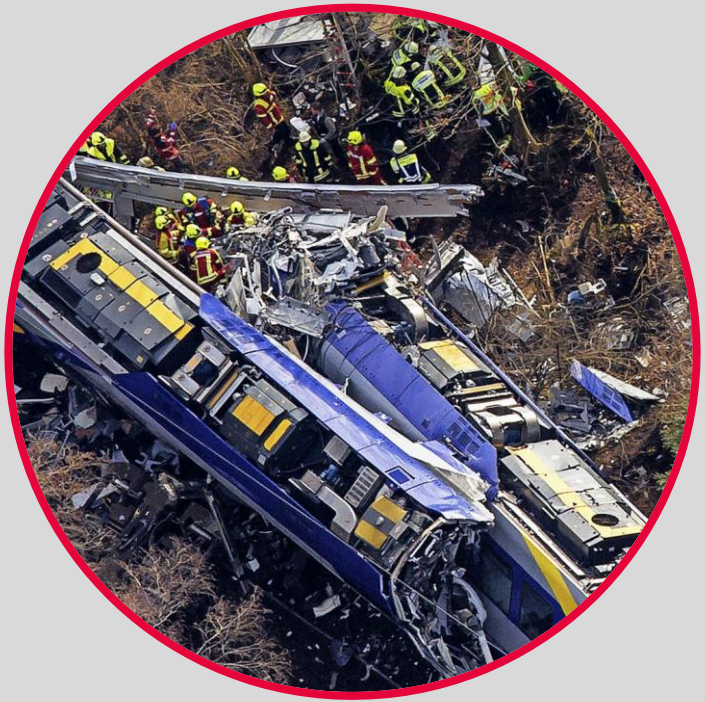
# Human errors

## A particular type of error

UNIVERSITÄT BERN

**What they are not**

- Not incompetence
  → Training!
- Not misunderstanding
  → Communication!



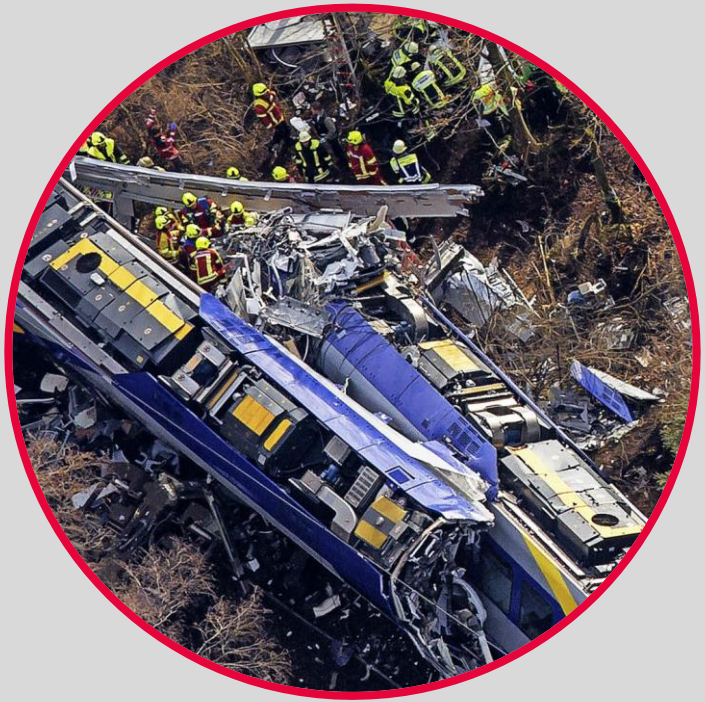
# Human errors

## A particular type of error

UNIVERSITÄT BERN

**Types**

- Random errors
  → Risk assessment
  → Quality Control?
- Systematic errors
  → Quality assurance:
  Screen – analyze – change

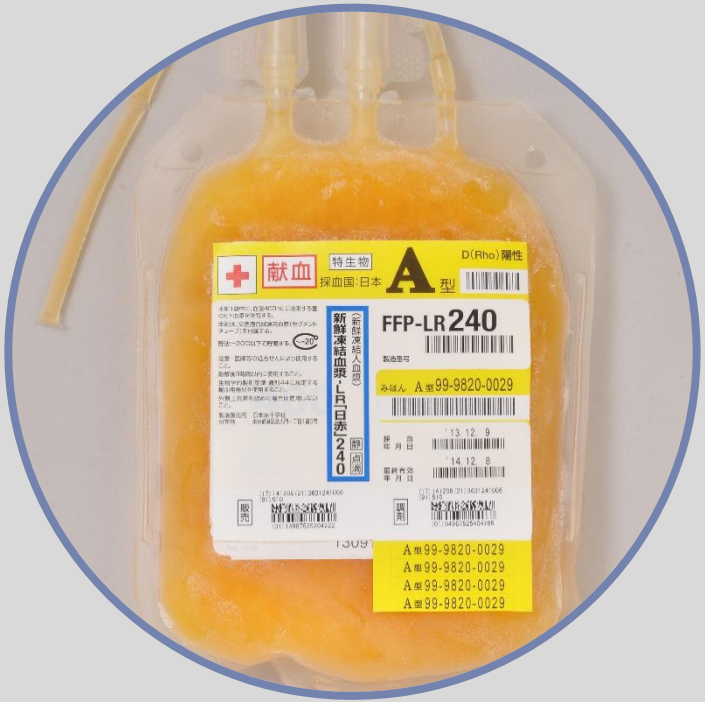


# Research misconduct

UNIVERSITÄT BERN

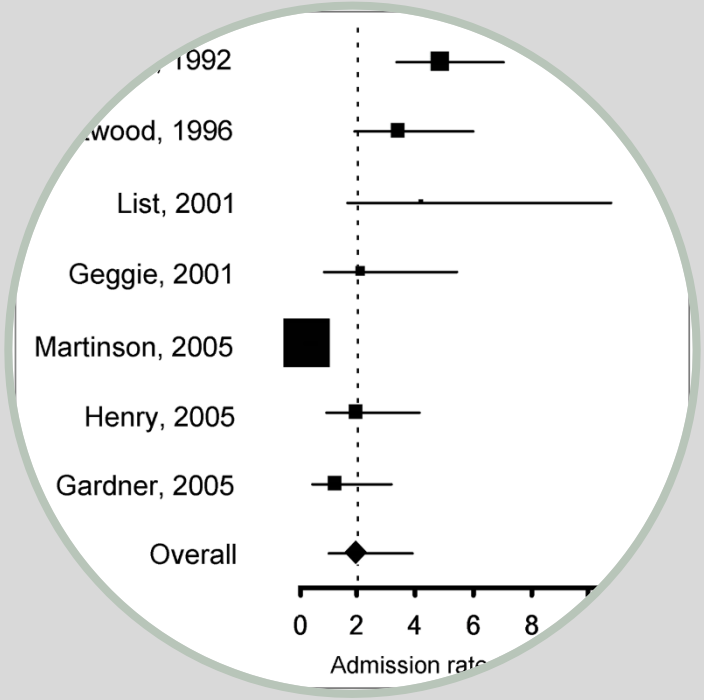
## FFP

- Fabrication
- Falsification
- Plagiarism



# Fabrication and falsification

– *… a significant effect was found for surveys targeted at **medical and clinical researchers**, who reported higher percentages of misconduct than respondents in biomedical research and other fields.*



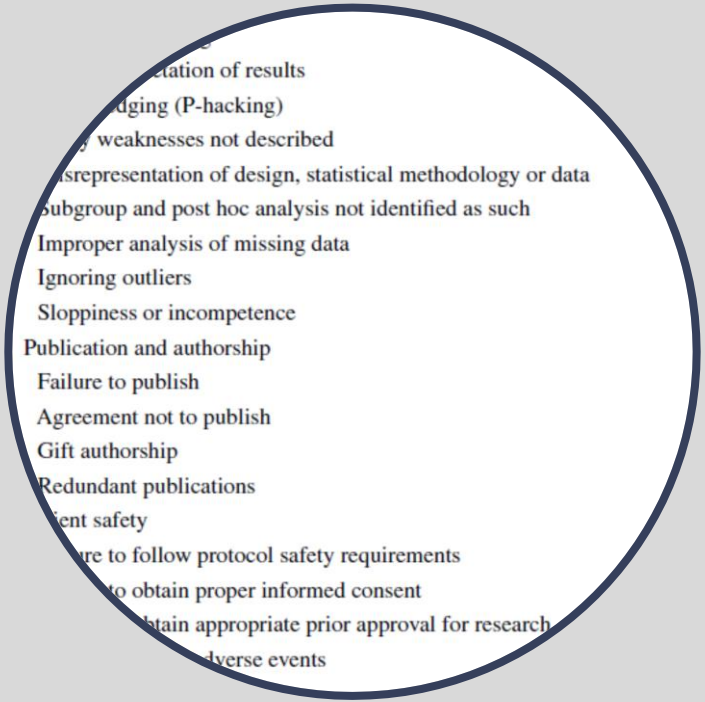

*Fanelli D 2009; Fiedler 2015*

# Questionable research practices

UNIVERSITÄT BERN

**To mention a few**

- Sloppiness
- Failure to follow protocol
- Over-interpretation of results
- Failure to publish, selective reporting
- Conflicts of interest
- …

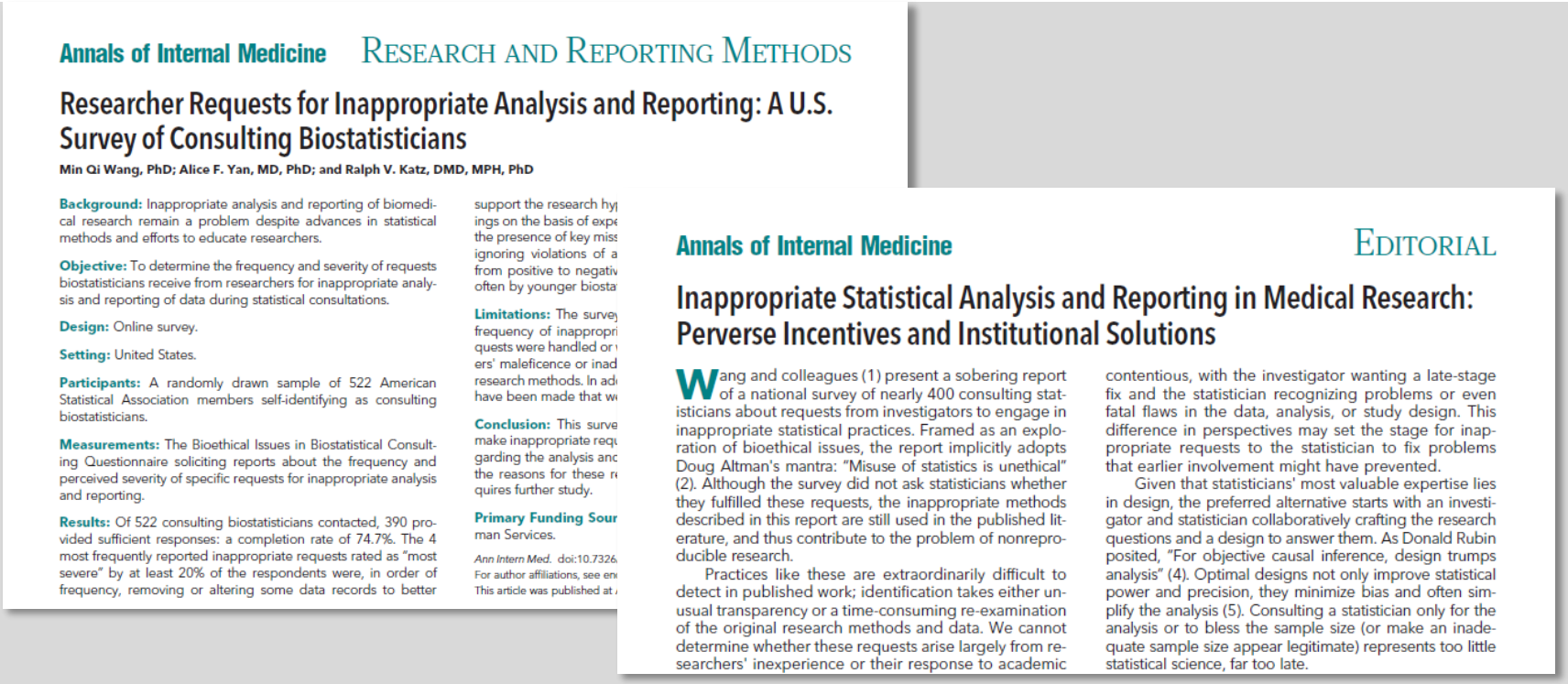



*George SL 2016; John 2012*

# Annals of Internal Medicine

## N=400 (of 522) ASA statisticians





Annals of Internal Medicine RESEARCH AND REPORTING METHODS

**Researcher Requests for Inappropriate Analysis and Reporting: A U.S. Survey of Consulting Biostatisticians**

Min Qi Wang, PhD; Alice F. Yan, MD, PhD; and Ralph V. Katz, DMD, MPH, PhD

**Background:** Inappropriate analysis and reporting of biomedical research remain a problem despite advances in statistical methods and efforts to educate researchers.

**Objective:** To determine the frequency and severity of requests biostatisticians receive from researchers for inappropriate analysis and reporting of data during statistical consultations.

**Design:** Online survey.

**Setting:** United States.

**Participants:** A randomly drawn sample of 522 American Statistical Association members self-identifying as consulting biostatisticians.

**Measurements:** The Bioethical Issues in Biostatistical Consulting Questionnaire soliciting reports about the frequency and perceived severity of specific requests for inappropriate analysis and reporting.

**Results:** Of 522 consulting biostatisticians contacted, 390 provided sufficient responses: a completion rate of 74.7%. The 4 most frequently reported inappropriate requests rated as "most severe" by at least 20% of the respondents were, in order of frequency, removing or altering some data records to better

support the research hy
ings on the basis of exp
the presence of key miss
ignoring violations of a
from positive to negativ
often by younger biosta

**Limitations:** The survey
frequency of inappropri
quests were handled or
ers' maleficence or inad
research methods. In ad
have been made that w

**Conclusion:** This surve
make inappropriate requ
garding the analysis and
the reasons for these r
quires further study.

**Primary Funding Sour**
man Services.

*Ann Intern Med.* doi:10.7326
For author affiliations, see en
This article was published at

Annals of Internal Medicine EDITORIAL

**Inappropriate Statistical Analysis and Reporting in Medical Research: Perverse Incentives and Institutional Solutions**

Wang and colleagues (1) present a sobering report of a national survey of nearly 400 consulting statisticians about requests from investigators to engage in inappropriate statistical practices. Framed as an exploration of bioethical issues, the report implicitly adopts Doug Altman's mantra: "Misuse of statistics is unethical" (2). Although the survey did not ask statisticians whether they fulfilled these requests, the inappropriate methods described in this report are still used in the published literature, and thus contribute to the problem of nonreproducible research.

Practices like these are extraordinarily difficult to detect in published work; identification takes either unusual transparency or a time-consuming re-examination of the original research methods and data. We cannot determine whether these requests arise largely from researchers' inexperience or their response to academic contentious, with the investigator wanting a late-stage fix and the statistician recognizing problems or even fatal flaws in the data, analysis, or study design. This difference in perspectives may set the stage for inappropriate requests to the statistician to fix problems that earlier involvement might have prevented.

Given that statisticians' most valuable expertise lies in design, the preferred alternative starts with an investigator and statistician collaboratively crafting the research questions and a design to answer them. As Donald Rubin posited, "For objective causal inference, design trumps analysis" (4). Optimal designs not only improve statistical power and precision, they minimize bias and often simplify the analysis (5). Consulting a statistician only for the analysis or to bless the sample size (or make an inadequate sample size appear legitimate) represents too little statistical science, far too late.

*Wang 2018*

*Table 1.* Biostatician-Reported Frequency and Severity Rating of Requests for Inappropriate Analysis and Reporting ($n$ = 390)*

| Violation Request | Respondents Rating the Item as "Most Severe," %† | Reported Requests During the Past 5 Years, % | | |
|---|---|---|---|---|
| | | 0 | 1-9 | ≥10 |
| Falsify the statistical significance (such as the $P$ value) to support a desired result | 84 | 97 | 2 | 1 |
| Change data to achieve the desired outcome (such as the prevalence rate of cancer or another disease) | 84 | 93 | 7 | - |
| Remove or alter some data records (observations) to better support the research hypothesis | 80 | 76 | 22 | 2 |
| Interpret the statistical findings on the basis of expectations, not the actual results | 68 | 70 | 28 | 2 |
| Do not fully describe the treatment under study because protocol was not exactly followed | 62 | 85 | 15 | - |
| Do not report the presence of key missing data that could bias the results | 68 | 76 | 23 | 1 |
| Ignore violations of assumptions because results may change to negative | 64 | 71 | 28 | 1 |
| Modify a measurement scale to achieve some desired results rather than adhering to the original scale as validated | 55 | 79 | 20 | 1 |
| Report power on the basis of a post hoc calculation, but make it seem like an a priori statement | 54 | 76 | 23 | 2 |
| Request to not properly adjust for multiple testing when "a priori, originally planned secondary outcomes" are shifted to an "a posteriori primary outcome status" | 56 | 80 | 18 | 2 |
| Conduct too many post hoc tests, but purposefully do not adjust $\alpha$ levels to make results look more impressive than they really are | 54 | 60 | 36 | 4 |
| Remove categories of a variable to report more favorable results | 48 | 68 | 31 | 1 |
| Do not mention interim analyses to avoid "too much testing" | 50 | 81 | 18 | 1 |
| Report results before data have been cleaned and validated | 48 | 56 | 39 | 5 |
| Do not discuss the duration of follow-up because it was inconsistent | 45 | 84 | 15 | 1 |
| Stress only the significant findings, but underreport nonsignificant ones | 42 | 45 | 48 | 7 |
| Do not report the model statistics (including effect size in ANOVA or $R^2$ in linear regression) because they seemed too small to indicate any meaningful changes | 42 | 76 | 23 | 1 |
| Do not show plot because it did not show as strong an effect as you had hoped | 33 | 58 | 39 | 3 |

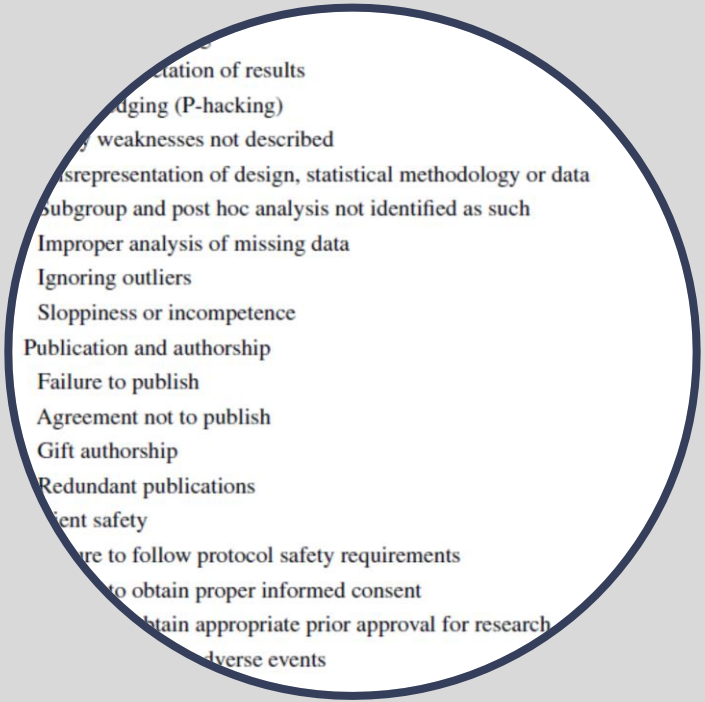
ANOVA = analysis of variance.
* Based on findings from questions 1–18 of the Bioethical Issues in Biostatistical Consulting Questionnaire, which asked biostatisticians "to estimate the number of times—during the past 5 years—that you, personally, have been DIRECTLY asked to do this." Data are presented in decreasing order by the percentage of respondents with a perceived severity score of 4 or 5.
† Items were defined as "most severe" if respondents ranked the severity as 4 or 5 on a scale of 0–5.

# Questionable research practices

## Probably most important

**To mention a few**

- Sloppiness
- Failure to follow protocol
- Over-interpretation of results
- Failure to publish
- Conflicts of interest
- …



# The trial/study protocol

## Purpose

- Helps to do a study in a uniform and reproducible way
- Operating manual for investigators and other study personnel
  → a good protocol helps and is not a barrier
- Safeguard for study participants
- Defines the boundaries of the project (approval, consent, …)

# Extend of problem

## Single-center experience

- 2010-2015
- 499 oncology trials (industry and investigator-initiated)
- Deviation reports filed at Institutional Review Board

**Table 1** Details of protocol deviations and responsible factors

| Type of deviation | Number | Institution | Subject | Sponsor | Schedule management | National holiday | Disease condition | Others |
|---|---|---|---|---|---|---|---|---|
| Visit | 189 (0) | 12 (0) | 2 (0) | 0 | 64 (0) | 68 (0) | 29 (0) | 14 (0) |
| Examination | 446 (16) | | | | | | | |
| Vital signs | 31 (0) | 28 (0) | 1 (0) | 0 | 1 (0) | 0 | 1 (0) | 0 |
| Clinical examination | 398 (10) | 291 (7) | 28 (0) | 12 (2) | 18 (0) | 5 (0) | 29 (1) | 15 (0) |
| Management of samples | 17 (6) | 12 (1) | 0 | 5 (5) | 0 | 0 | 0 | 0 |
| Treatment | 275 (72) | | | | | | | |
| Administration criteria | 29 (15) | 24 (11) | 1 (1) | 0 | 0 | 0 | 3 (3) | 1 (0) |
| Overdose | 36 (19) | 22 (10) | 12 (8) | 2 (1) | 0 | 0 | 0 | 0 |
| Underdose | 37 (6) | 5 (2) | 24 (0) | 2 (2) | 0 | 0 | 5 (2) | 1 (0) |
| Therapeutic method | 123 (14) | 55 (9) | 19 (0) | 2 (0) | 4 (0) | 1 (0) | 13 (5) | 29 (0) |
| Concomitant medication/activity | 39 (12) | 14 (2) | 4 (1) | 0 | 0 | 0 | 15 (8) | 6 (1) |
| Management of investigational products | 11 (6) | 7 (4) | 2 (0) | 2 (2) | 0 | 0 | 0 | 0 |
| Others | 57 (37) | | | | | | | |
| Eligibility criteria | 16 (11) | 12 (8) | 0 | 1 (1) | 0 | 0 | 0 | 3 (2) |
| Stratification | 4 (3) | 3 (3) | 0 | 1 (0) | 0 | 0 | 0 | 0 |
| Informed consent | 18 (13) | 18 (13) | 0 | 0 | 0 | 0 | 0 | 0 |
| Protection of personal information | 4 (4) | 3 (3) | 0 | 1 (1) | 0 | 0 | 0 | 0 |
| Reporting | 10 (2) | 10 (2) | 0 | 0 | 0 | 0 | 0 | 0 |
| Case report form | 1 (1) | 1 (1) | 0 | 0 | 0 | 0 | 0 | 0 |
| Others | 4 (3) | 3 (3) | 0 | 0 | 1 (0) | 0 | 0 | 0 |
| Total | 967 (125) | 520 (79) | 93 (10) | 28 (14) | 88 (0) | 74 (0) | 95 (19) | 69 (3) |

The numbers in parentheses denote serious deviations

*Sasada 2017*

**Table 1** Details of protocol deviations and responsible factors

| Type of deviation | Number | Institution | Subject | Sponsor | Schedule management | National holiday | Disease condition | Others |
|---|---|---|---|---|---|---|---|---|
| Visit | 189 (0) | 12 (0) | 2 (0) | 0 | 64 (0) | 68 (0) | 29 (0) | 14 (0) |
| Examination | 446 (16) | | | | | | | |
| Vital signs | 31 (0) | 28 (0) | 1 (0) | 0 | 1 (0) | 0 | 1 (0) | 0 |
| Clinical examination | 398 (10) | 291 (7) | 28 (0) | 12 (2) | 18 (0) | 5 (0) | 29 (1) | 15 (0) |
| Management of samples | 17 (6) | 12 (1) | 0 | 5 (5) | 0 | 0 | 0 | 0 |
| Treatment | 275 (72) | | | | | | | |
| Administration criteria | 29 (15) | 24 (11) | 1 (1) | 0 | 0 | 0 | 3 (3) | 1 (0) |
| Overdose | 36 (19) | 22 (10) | 12 (8) | 2 (1) | 0 | 0 | 0 | 0 |
| Underdose | 37 (6) | 5 (2) | 24 (0) | 2 (2) | 0 | 0 | 5 (2) | 1 (0) |
| Therapeutic method | 123 (14) | 55 (9) | 19 (0) | 2 (0) | 4 (0) | 1 (0) | 13 (5) | 29 (0) |
| Concomitant medication/activity | 39 (12) | 14 (2) | 4 (1) | 0 | 0 | 0 | 15 (8) | 6 (1) |
| Management of investigational products | 11 (6) | 7 (4) | 2 (0) | 2 (2) | 0 | 0 | 0 | 0 |
| Others | 57 (37) | | | | | | | |
| Eligibility criteria | 16 (11) | 12 (8) | 0 | 1 (1) | 0 | 0 | 0 | 3 (2) |
| Stratification | 4 (3) | 3 (3) | 0 | 1 (0) | 0 | 0 | 0 | 0 |
| Informed consent | 18 (13) | 18 (13) | 0 | 0 | 0 | 0 | 0 | 0 |
| Protection of personal information | 4 (4) | 3 (3) | 0 | 1 (1) | 0 | 0 | 0 | 0 |
| Reporting | 10 (2) | 10 (2) | 0 | 0 | 0 | 0 | 0 | 0 |
| Case report form | 1 (1) | 1 (1) | 0 | 0 | 0 | 0 | 0 | 0 |
| Others | 4 (3) | 3 (3) | 0 | 0 | 1 (0) | 0 | 0 | 0 |
| Total | 967 (125) | 520 (79) | 93 (10) | 28 (14) | 88 (0) | 74 (0) | 95 (19) | 69 (3) |

The numbers in parentheses denote serious deviations

# Protocol = ethics application?

## A common missunderstanding

UNIVERSITÄT BERN

*Q: “Do you already have a study protocol or at least an outline?”*
*A: “Yes, I am currently working on the ethics application!”*

# Protocol = ethics application?

## A common missunderstanding

– People often talk about the protocol as being “the ethics application”

→ Although the protocol has to be submitted to the ethics committee and despite the fact that study protocols should usually be written using the Swissethics protocol templates, it has **nothing** to do with the **ethics** application **per se**.

# Main issue (root cause) in practice

## Physician versus investigator (researcher)

# Causal factors in general

UNIVERSITÄT BERN

## On different levels

- Individual traits
- Institutional issues
- Structural problems in science itself

## Smith R 2016

> *The answer that researchers love is `pressure to publish', but my preferred answer is `Why wouldn't it happen?' All human activity is associated with misconduct. […] they [researchers] have fooled themselves that science is a wholly objective enterprise unsullied by the usual human subjectivity and imperfections. It is not. It is a human activity.*

# Clinician researcher

## A case of fraud

UNIVERSITÄT BERN

**Swazey JP & Scher SR**

*It is our sense, primarily experiential and impressionistic in nature, that honesty in research work as a fundamental rule is valued more strongly among scientists than among physicians … Physicians tend to evaluate research in terms of harm or benefit to patients rather than in terms of adherence to the rigorous norms of scientific investigation*

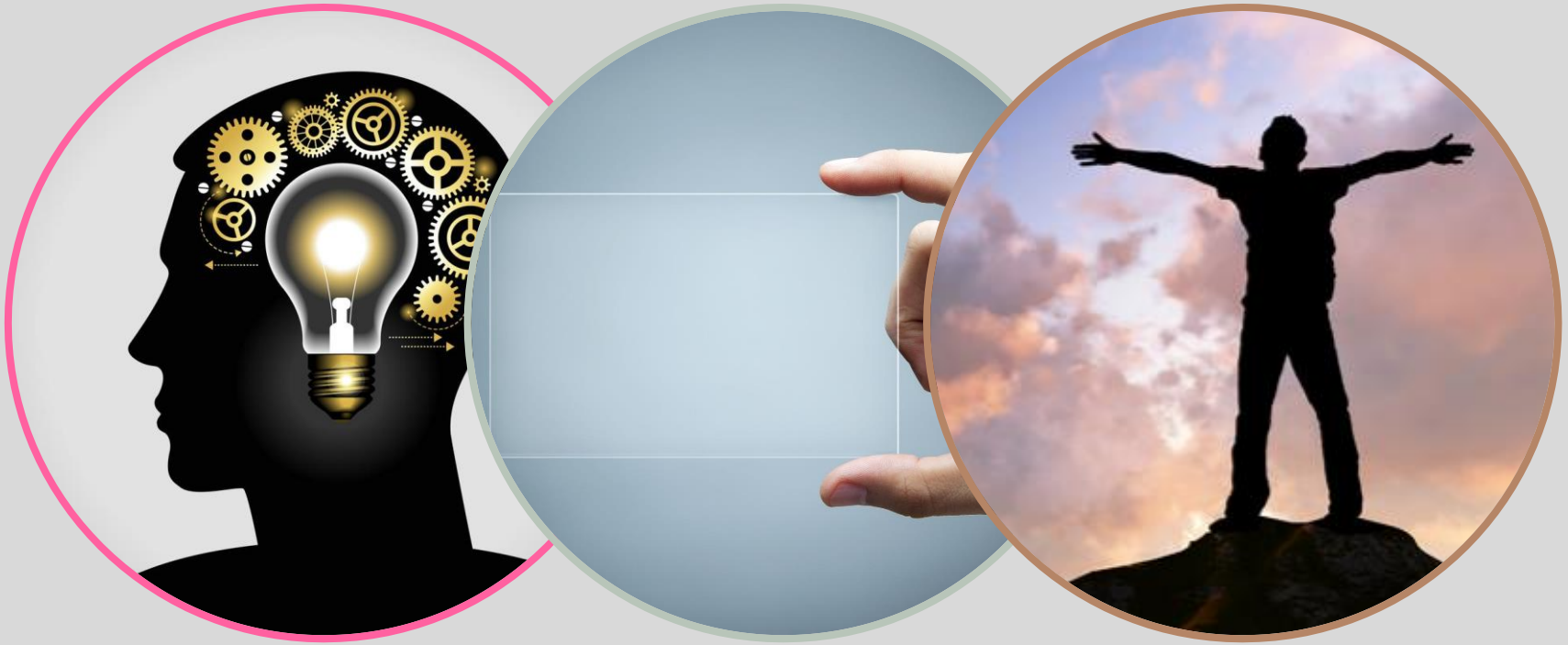
**Poisson R**

*I believed I understood the reasons behind the study rules, and I felt that the rules were meant to be understood as guidelines and not necessarily followed blindly. My sole concern at all times was the health of my patients … Maintaining the proper balance between good clinical care and rigid research methods is not an easy task*

*Cited in George SL 2016*

# What can we do about these issues?

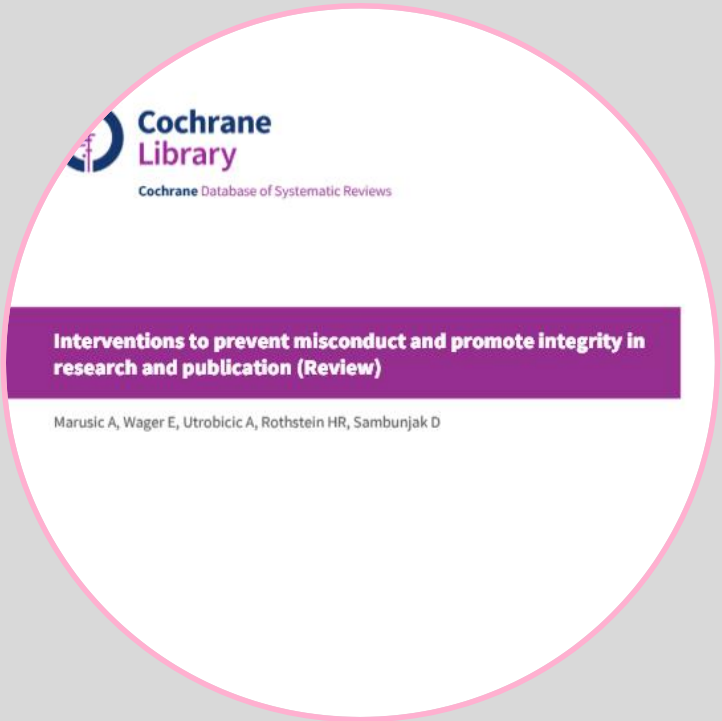




# Prevention

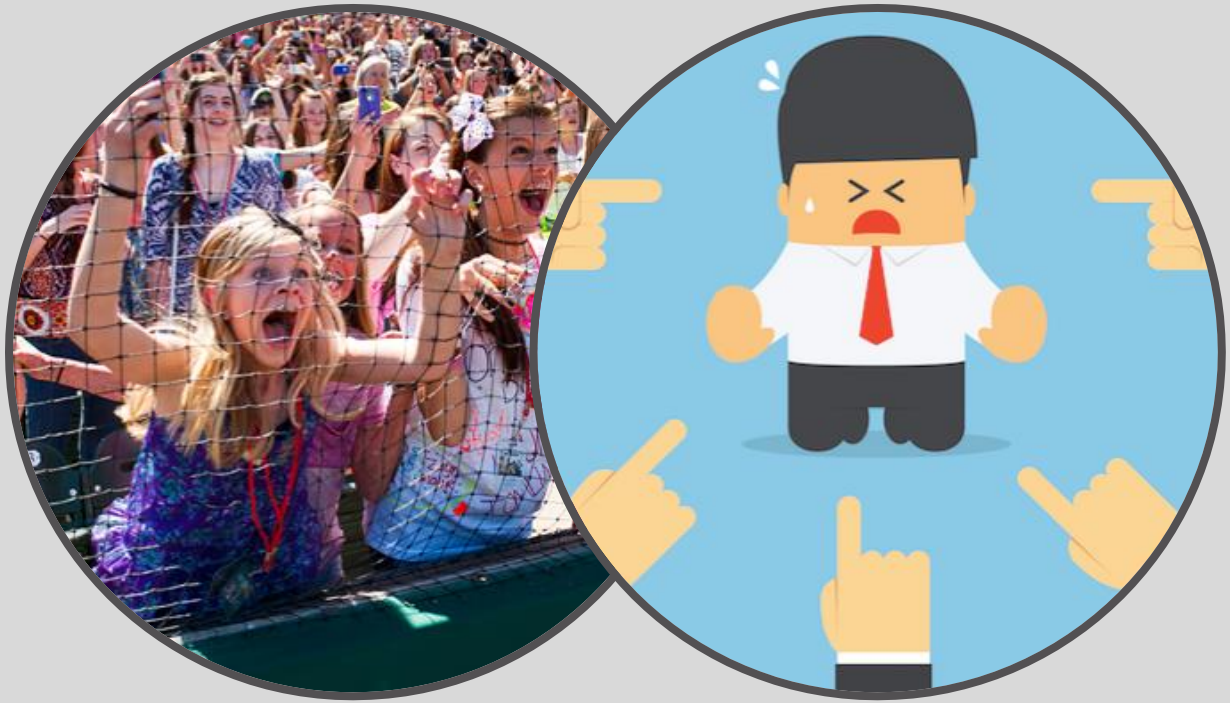
## Best evidence

- The evidence base is incomplete
- Many studies with high risk of bias and inadequately reported
- Even randomized designs were difficult to generalize



# Culture





# The socialization (way of thinking)

## Individual versus general (population)

# Thank you

# for your attention!

**Sven Trelle, CTU Bern**

UNIVERSITÄT BERN

**References**

- Fanelli D 2009. PLoS One; 4: e5738
- Fiedler 2015. Soc Psych Personal Sci; 7: 45.
- George SL 2016. Int J Clin Oncol; 21:15.
- John 2012. Psychol Sci; 23: 524.
- Redmond S et al. 2013: J Clin Epidemiol; 66: 1367.
- Sackett D 2006. In: Haynes BR et al. Clinical Epidemiology 3rd ed. Philadelphia: LWW. P. 415ff
- Sasada 2018. Invest New Drugs; 35: 392.
- Smith R 2006. J R Soc Med 99:232
- Wang 2018. Ann Intern Med; 169: 554.
- Weidmann R 2018: Master thesis. Med Fac Uni Bern.

**Sources**

- Balancing on a string by Leio McLaren (@leiomclaren) on Unsplash
- Female ballerina by David Hofmann on Unsplash
- Bisons by Richard Lee on Unsplash
- Weight by Pau Casals on Unsplash
- www.pexels.com