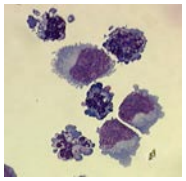


Scientific fraud/sloppiness in data handling/data reproducibilty

Quality control in science

The university perspective

Measures to ensure good scientific practice

Klaus-Michael Debatin
University Children‘s Hospital Ulm
Member DFG Senate

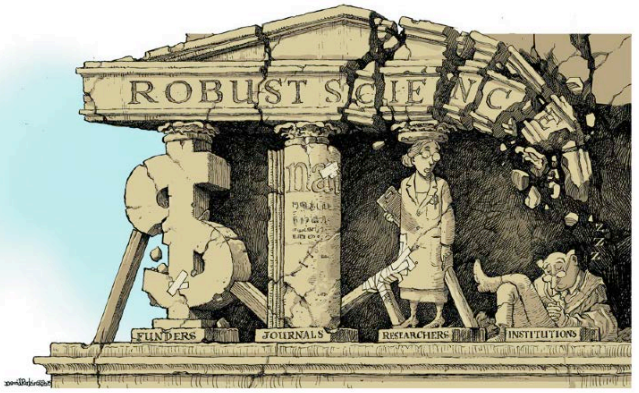
THE LANCET

Research: increasing value, reducing waste · January, 2014 www.thelancet.com

"By ensuring that efforts are infused with rigour from start to finish, the research community might protect itself from the sophistry of politicians, disentangle the conflicted motivations of capital and science, and secure real value for money for charitable givers and taxpayers through increased value and reduced waste."

Research: increasing value, reducing waste

"...the preclinical strategies to evaluate novel agents are suboptimal ..."



Institutions must do their part for reproducibility

Begley, Buchan, Dirnagl Nature 2015

Home › Research Waste/EQUATOR Conference

Research Waste/EQUATOR Conference

REWARD

REduce research WAste and Reward Diligence

http://researchwaste.net/

equator network

Enhancing the QUAlity and Transparency Of health Research

www.equator-network.org/

Increasing value and reducing waste in biomedical research conference

28th – 30th September 2015, Edinburgh

Nature. 2014 January 30; 505(7485): 612–613.

NIH plans to enhance reproducibility

Collins and Tabak

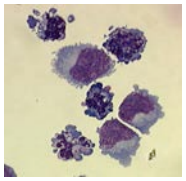
How to Make More Published Research True

John P. A. Ioannidis[1,2,3,4]*

PLOS Medicine 2014

High drug attrition rates—where are we going wrong?

Hutchinson, Kirk; *Nat Rev Clin Onc 2011*



Scientific fraud/sloppiness in data handling/data reproducibilty

Quality control in science

The university perspective

Measures to ensure good scientific practice

- Cases
- Problem(s)
- Measures to ensure quality control in science

Cell

Volume 45, Issue 2, 25 April 1986, Pages 247–259

Article

Altered repertoire of endogenous immunoglobulin gene expression in transgenic mice containing a rearranged Mu heavy chain gene

David Weaver*, †, Moema H. Reis†, §, Christopher Albanese†, §, Frank Costantini‡, David Baltimore*, †, Thereza Imanishi-Kari†, §

- Unexpected finding, reproducibility of experiments (never done ?) ?
- Retracted 1991 by all authors except I-K
- Case running 1986 – 1996
- Sloppyness in data handling vs. faking data

Cell

Volume 64, Issue 6, 22 March 1991, Pages 1103–1110

Article

Oct-3 is a maternal factor required for the first mouse embryonic division

Mitchell H. Rosner*, †, †, Ronald J. De Santo‡, †, Heinz Arnheiter‡, †, Louis M. Staudt*

Science 1991

Oct-3 and the Beginning of Mammalian Development

MITCHELL H. ROSNER, M. ALESSANDRA VIGANO,
PETER W. J. RIGBY, HEINZ ARNHEITER,
LOUIS M. STAUDT*

- Highly exciting data on first steps of mammalian development
- Retracted in 1992
- Data entirely fabricated
- Dr. Staudt took immediate action to directly uncover the fraud
- ORI at NIH already in place

Retraction: Stimulus-triggered fate conversion of somatic cells into pluripotency

Haruko Obokata, Teruhiko Wakayama, Yoshiki Sasai, Koji Kojima, Martin P. Vacanti, Hitoshi Niwa, Masayuki Yamato & Charles A. Vacanti

Nature **511**, 112 (03 July 2014) | doi:10.1038/nature13598
Published online 02 July 2014 | Corrected online **23 July 2014**
Article (January, 2014)
Correction (July, 2014)

Retraction: Bidirectional developmental potential in reprogrammed cells with acquired pluripotency

Haruko Obokata, Yoshiki Sasai, Hitoshi Niwa, Mitsutaka Kadota, Munazah Andrabi, Nozomu Takata, Mikiko Tokoro, Yukari Terashita, Shigenobu Yonemura, Charles A. Vacanti & Teruhiko Wakayama

Nature **511**, 112 (03 July 2014) | doi:10.1038/nature13599
Published online 02 July 2014 | Corrected online **23 July 2014**

- Highly exciting data in a top fashion field (stem cell reprograming)
- “Too good to be true“: adding acid or stress reprograms differentiated cells
- Data mostly fabricated or experiments shown not done
- Retraction within 6 months

Retraction: Kiehntopf, M., F. Herrmann, and M.A. Brach. 1995. Functional NF-IL6/CCAAT enhancer-binding protein is required for tumor necrosis factor α-inducible expression of the granulocyte colony-stimulating factor (CSF), but not the granulocyte/macrophage CSF or interleukin 6 gene in human fibroblasts. *J. Exp. Med.* 181:793–798.

The Editorial Board of *The Journal of Experimental Medicine* regrets to inform its readership that due to scientific misconduct, the results of the above named paper are not valid. A letter dated April 30, 1997 was received from Guido Adler, M.D., the Dean of the School of Medicine at the University of Ulm. In his letter, Dr. Adler states that "fraud was uncovered in the laboratory of Friedhelm Herrmann and Marion Brach. In a series of papers major results were fabricated [including]: . . .

Kiehntopf M, Herrmann F, Brach M. *J. Exp. Med.* 181, 793, 1995."

- The major german scandal in biomedicine leading to the so far largest investigation of scientific fraud in Germany with subsequent action by the DFG and universities.
- Of 347 papers 1985-1996 published by the Hermann group, only 132 were cleared of any suspicion of fraud. 94 papers definitely contained manipulated data.
- Most of the work was on cytokines, at that time a fashionable field with multiple photo shop data, copy and paste, reuse of data etc. All "in line" with data or concepts already published by others.
- A highly dedicated "follower of fashion". All fakes escaped the reviewers attention. Scientific fraud was uncovered by whistleblowing.

Common aspects in (some) of these cases ?

- Productivity and/ or results too good to be true
- Recognition that missing links would represent a breakthrough
- Knowing the answer to their research question
- Knowing that individual experiments may not be exactly reproducible

- Outstanding intellect, well informed in the field
- High pressure on career and/or reputation
- Hubris

- Institution/mentor/supervisor/environment
  that favors “stars“ (also to increase their own reputation)
- Journals like superpositive ”exciting“ data

Scientific fraud and normal science

Ruud Abma

Faculty of Social and Behavioral Sciences / Descartes Centre

Utrecht University

Science in transition - Workshop *Quality and Corruption*

Until recently, the incidence of scientific fraud was unknown. Hard evidence is still lacking, but we do know more: estimates range from 2% (Fanelli, 2009) tot 10% (John, 2011) of people admitting to having falsified data themselves and about 14% having observed in it colleagues.

Fraud is usually not detected by reviewers or referees. Almost always it is co-workers who blow the whistle (Stroebe et al, 2012). Why? Because they are in close contact with the perpetrator and his or her daily routine and also have an overview of the total of his research activities.

Between brackets the number of fraudulent publications:

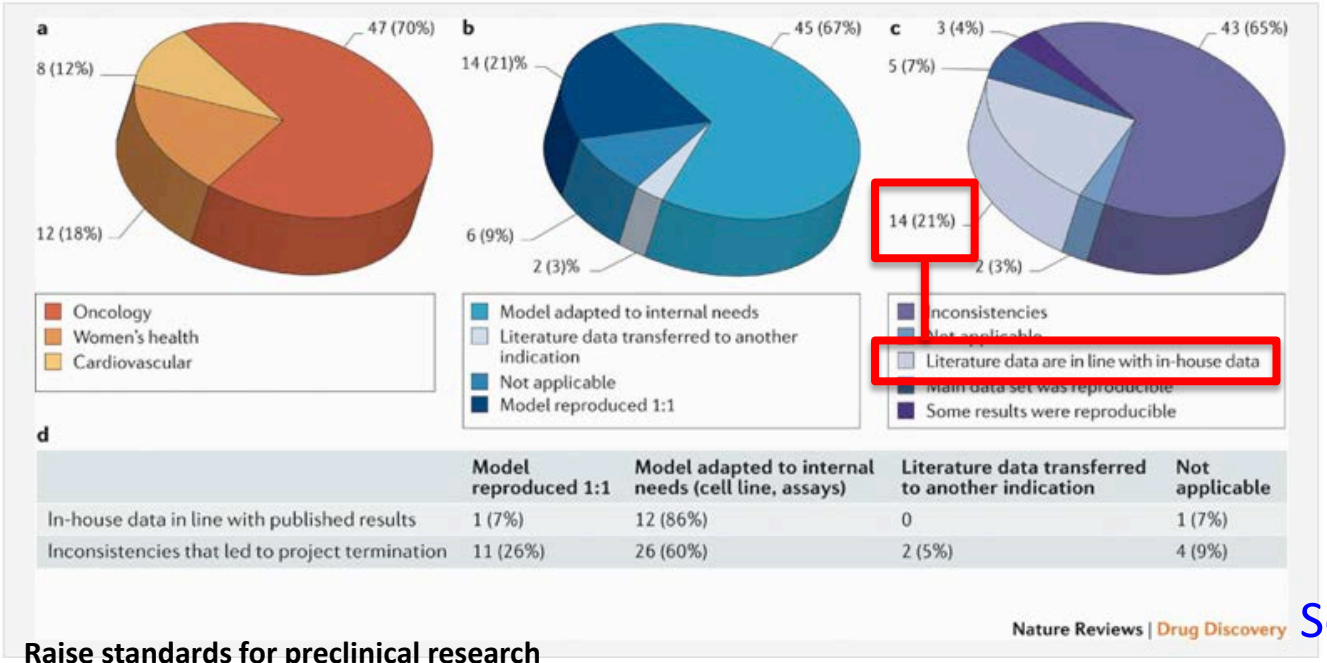
1. Yoshitaka Fuji, anaesthesiologist, Japan (172)
2. Dipak K. Das, heart surgeon, USA (145)
3. John Darsee, physician, VS, 1966-1981 (82 tot 104).
4. Friedhelm Hermann en Marion Brach, physicians, Duitsland, 1994-1997 (94)
5. Diederik Stapel, social psychologist, Nederland, 1997-2011 (69)
6. Jan-Hendrik Schön, physicist,VS/Duitsland, 1997-2002 (27 tot 35)
7. Scott Reuben, physician, VS, 1996-2009 (21)
8. Alirio Mendelez, immunologist, Singapore (21)
9. Stephen Breuning, physician, VS, 1975-1988 (20)
10. John Sudbø, dentist and oncologist, Noorwegen, 1993-2005 (15)
11. Roger Poisson, physician, Canada, 1977-1990 (14)
12. Luk van Parijs, biologist, België/VS, 2000-2004 (11)
13. Eric Poehlman, physician, VS, 1992-2002 (10)
14. Marc D. Hauser, primatologist, VS, 1995-2010 (9)

The other problem: reproducibility

FIGURE 1 | Analysis of the reproducibility of published data in 67 in-house projects.

FROM THE FOLLOWING ARTICLE:
Believe it or not: how much can we rely on published data on potential drug targets?
Florian Prinz, Thomas Schlange & Khusru Asadullah
Nature Reviews Drug Discovery **10**, 712 (September 2011)
doi:10.1038/nrd3439-c1

‹ Back to article | ‹ Back to figures and tables



d

	Model reproduced 1:1	Model adapted to internal needs (cell line, assays)	Literature data transferred to another indication	Not applicable
In-house data in line with published results	1 (7%)	12 (86%)	0	1 (7%)
Inconsistencies that led to project termination	11 (26%)	26 (60%)	2 (5%)	4 (9%)

Nature Reviews | Drug Discovery

Only 20-25%
of published data
in line with
in-house findings

Raise standards for preclinical research
C. Glenn Begley & Lee M. Ellis
Nature **483**, 531–533 (29 March 2012)

Table 1: Reproducibility of research findings

Preclinical research generates many secondary publications, even when results cannot be reproduced.

Journal impact factor	Number of articles	Mean number of citations of non-reproduced articles*	Mean number of citations of reproduced articles
>20	21	248 (range 3–800)	231 (range 82–519)
5–19	32	169 (range 6–1,909)	13 (range 3–24)

Results from ten-year retrospective analysis of experiments performed prospectively. The term 'non-reproduced' was assigned on the basis of findings not being sufficiently robust to drive a drug-development programme.

Scientific findings
(oncology)
were (entirely) confirmed
in only 11% of the cases

BLAME IT ON THE ANTIBODIES

Antibodies are the workhorses of biological experiments, but they are littering the field with false findings. A few evangelists are pushing for change.

BY MONYA BAKER

NATURE, 21 May 2015

Reproducibility of research and preclinical validation: problems and solutions

Lajos Pusztai, Christos Hatzis and Fabrice Andre

Nat.Rev.Clin.Oncol. 10,720-724 (2013)

Box 1 | Contributing factors to poor reproducibility and counter measures

Contributors to poor reproducibility

- Unrecognized experimental variables
- Poor documentation of methodology
- Selective reporting
- Misinterpretation of noise as an indication of a positive finding
- Inappropriate statistics or deceptive analysis
- Data fabrication
- Academic and financial incentives to publish premature results

Possible counter measures

- Standard operating procedures and relevant control experiments
- Detailed protocols and meticulous laboratory log books, publication of software codes and raw data
- Accounting for all experiments and deposition of all raw data
- Publication of experimental protocol and raw data, appropriate statistics, awareness of biases during data analysis
- Prospective statistical planning and analysis blinded to outcome (if possible)
- Laboratory supervision and culture of respect for negative results
- Venues for publication of negative results, more diverse research funding, changes in academic promotion practices

Box 2 | Three-pronged funding schema

Innovative Research Grant (12–36 months)
Discovery oriented, research question driven, high risk for failure

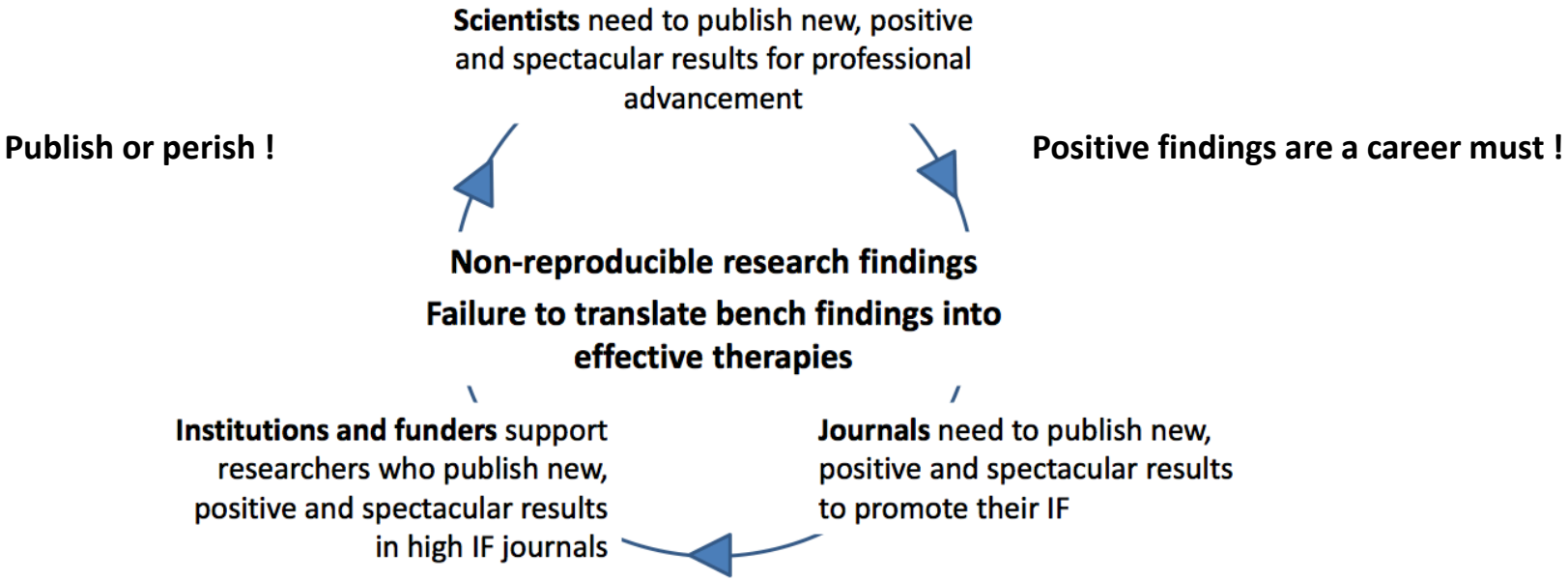
Replication Grant (12–24 months)
Reproduce important published results with minor or no changes in experimental design

Product Development Grant (12–36 months)
Turn reproducible results into practically useful product

Peter Higgs, Nobel Prize 2013:
„..... Today, I wouldn‘t get an academic job: It‘s as simple as that. I don‘t think I would be regarded as productive enough“....

Randy Schekman, Nobel Prize 2013:
„...These luxury journals are supposed to be the epitome of quality... While they publish many outstanding papers, they do not publish only outstanding papers. Neither are they the only publishers of outstanding papers“...

The vicious cycle of biomedical research



Money (and reputation) for individual scientists and institutions goes with performance.
Performance is measured by total amount of grant money acquired by scientist/institution and total IF

The issues

- Low rate of reproducible preclinical data e.g. on drug targets (20 – 40%) and gene expression profiles (10-40%)

- 60-80% failure rate in phase II – III cinical trials.
  Costs for taking a drug to the clinic = over 800 Mio USD due to a 3-4 in 5 failure rate (costs for one agent around 160 Mio USD)

- 2% of scientists admitted to have fabricated, falsified or modified dat at least once
  33% admitted questionable research practices. Rates observed in collegues
  are 14% and 72% respectively (Fanelli, PLOS one 2009)

- **"Benevolent mistakes"** **vs** **outright fraud:**
  poor experimental design — fabrication
  sloppy data management — falsification
  bias in interpretation of data — plagiarism

- According to Medline 0.03% of 17.8 Mio published papers 1980-2014 were retracted (Gunn, Nature 2014)

The culture

- Progress in science means continual production of positive results, only novelty counts for journals
  and scientists prefer new truth and discovery

- Lack of incentives to report negative results (journals are not interested)

- Lack of incentives to replicate experiments or recognize inconsistancies
  (journals and grant agencies are not interested)

- Impact factor = Impact ?
  Are cumulative IF + aquired grant money the right figures to judge/measure performance ?

Begley CG Nature 2013

Six red flags for suspect work

Six Questions

- Were experiments performed blinded ?
- Were basic experiments repeated ?
- Were all the results presented ?
- Were there positive and negative controls ?
- Were reagents validated ?
- Were statistical tests appropriate ?

The extreme view

Principles and Guidelines for Reporting Preclinical Research

The signatories represent journals that publish preclinical biological research — an area of research that encompasses both exploratory studies and hypothesis-testing studies, with many different designs. The reproducibility of these studies is expected to vary. The journals agree to adhere to the following principles with the aim of facilitating the interpretation and repetition of experiments as they have been conducted in the published study. These measures and principles do not obviate the need for replication and reproduction in subsequent investigations to establish the robustness of published results across multiple biological systems.

Nov 2014; update Feb 2015

A Checklist for Journals

1. Rigorous statistical analysis

2. Transparency in reporting
   - Standards
   - Replicates
   - Statistics
   - Randomization
   - Blinding
   - Sample-size estimation
   - Inclusion and exclusion criteria

3. Data and material sharing

4. Consideration of refutation

5. Establishing best practice guidelines for
   - Image based data
   - Antibodies
   - Cell lines
   - Animals

Measures by the university in general

- Statutes for Safeguarding Good Scientific Practice introduced by the senate see https://www.uni-ulm.de/fileadmin/website_uni_ulm/zuv/zuv.dezIII.abt2u3/3-2oeffentlich/bekanntmachungen/2009/verantwortung_id__wiss_09.pdf

- Ombudspersons of the university
  First port of call for members and staff members of the university wishing to clarify issues concerning good scientific practice or in the event of a suspected case of scientific misconduct.

- Committee on Responsibility in the Conduct of Science
  Monitors the guidelines and investigates allegations of scientific misconduct when requested by the ombudspersons

- Additional consequences
  The presidency and the faculties can – after proven scientific misconduct – decide about additional measures, e.g. deprivation of degrees, dismissal of staff, informing injured parties including journals and funding agencies (e.g. DFG) that decide about additional consequence

Good scientific practice

During PhD Program (International Graduate School in Molecular Medicine)

- Good scientific practice, scientific misconduct and fraud are a special focus during contract signing (an individual conversation between the managing director and every student).
- Good Scientific Practice is a compulsory seminar at the Graduate School since 2008 to be taken at the beginning of the practical work; topics covered are:
  - Process of Developing Scientific Knowledge from developing the scientific question to publication of results
  - Data Management
  - Ownership of data and material
  - Authorship
  - Conflict of Interest
  - Tutorship for young scientists
  - Different scientific cultures
  - Conflicts in cooperation
  - Scientific fraud
  - Basis of Bioethics
- Seminars "Intellectual Property Rights" and "How to cite correctly"
- Random checks of PhD theses for plagiarism
- Thesis advisory by senior scientists compulsorily including an external scientist
- When submitting the thesis, students need to sign a declaration in lieu of oath confirming Good Scientific Practice, especially no plagiarism

Quality Control (Graduate School)

Graduate School, PhD Program and PIs

- Accreditation and renewal of accreditation of PIs
- Student representatives in PhD committee
- Evaluation measures
  - Yearly plenary meetings between students and board of directors in order to identify weak points of the programs and new training modules
  - Two meetings per semester with the student representatives in order to discuss current issues
  - Online evaluation program together with the quality assurance board of the Medical Faculty; systematic evaluation of individual activities
  - Evaluation of the program from the supervisors' viewpoint planned
  - Plenary meetings of the PIs in order to identify requirements from their views and scientific orientation
  - Statistical Monitoring of students (e.g. drop out rate, median time to completion, history of former graduates) and PIs (e.g. publications and grant acquisition)
  - Discussion of evaluation results in the board
- International and interdisciplinary Advisory board (including industry and alumni) meets annually and evaluates activities
- Statement of accounts (financial evaluation)
- Ombudspersons within the graduate school (different from ombudspersons for scientific misconduct)
- Brochure "Good Supervision Practice" for PIs
- Supervision Agreement to be concluded at the beginning

Good scientific practice: Culture and Rules

Labs Universities Institutions	- Internal lab control: regular lab seminar and journal club with critical discussion - Rules for handling and storage of original data - Signed responsibility for authorship - Ombudsman - Constitutional rules for Good Scientific Practice and handling suspicion/claim for scientific misconduct - New measures of performance in career development ?
Journals/Reviewers:	- Reduce “excitement level“ - Transparency in data provided, reagents and procedures - Check for overinterpretation of data - Allow publication of negative results, - Allow confirmatory studies (“the data are not entirely new“)
Grant agencies:	- Reduce “excitement level“ - Reconsideration of rating of grant applications: originality, novelty, excellence, cumulative IF etc. - Confirmatory studies are necessary - Negative data are data !
Scientific community:	- Open discussion on quality of science - Impact is not Impact Factor - WE ARE THE REVIEWERS IN ALL ABOVE !