

UNIVERSITY OF GOTHENBURG **KREFTING RESEARCH CENTRE** 

# Thinking <u>REPRODUCIBILITY</u> in Your Research Work

13<sup>th</sup> August 2020 Bright Nwaru, Associate Professor University of Gothenburg, Sweden



# Aims of talk

- Motivate you into field of reproducible research
- Emphasize importance of thinking reproducibility throughout research
- Highlight consequences of irreproducibility
- Point to tools and platforms to enhance reproducibility in your research



### First things First: What is science/research?

- **SCIENCE:** "Any system of knowledge that is concerned with the physical world and its phenomena and that entails unbiased observations and systematic experimentation. In general, a science involves a pursuit of knowledge covering general truths or the operations of fundamental laws" (Encyclopedia Britannica: https://www.britannica.com/)
- **RESEARCH:** "A process of systematic inquiry that entails collection of data; documentation of critical information; and analysis and interpretation of that data/information, in accordance with suitable methodologies set by specific professional fields and academic disciplines" (Hampshire College: <u>https://www.hampshire.edu/dof/what-is-research</u>)

Why do scientists do research? What is the end goal of my research? These are fundamental questions Often not considered by many scientists



### Science's greatest motivation: a safer society (UNESCO)

### • Science: greatest collective endeavor. Contributes to:

- Longer and healthier life: monitors our health, provides medicine to cure our diseases, alleviates aches and pains
- Provision of our basic needs: water, food, energy, clean air
- Making our lives more fun: sports, music, entertainment and communication technology
- Creating new knowledge to improve education and quality of our lives
- Science generates solutions for everyday life
- Science must respond to societal needs and global challenges.



UNIVERSITY OF GOTHENBURG **KREFTING RESEARCH CENTRE** 

# But does (all) science fulfil its mandate?



"Most research findings are false for most research designs and for most fields"

Ioannidis 2005; PLoS Med 2(8): e124.



UNIVERSITY OF GOTHENBURG **KREFTING RESEARCH CENTRE** 

### Some case studies of bad science

### Pressure: publish or perish

### naturenews

contain faked data.

When colleagues called the work of

good to be true, they meant it as a

compliment. But a preliminary investigative report

phrase, detailing years of data

the prominent Tilburg University

Ewen Callaway

researcher.

news archive specials opinion features news blog

<u>comments</u> on this <u>story</u>

nature news home

Published online 1 November 2011 | Nature 479, 15 (2011) | doi:10.1038/479015a Updated online: 1 November 2011

Investigation claims dozens of social-psychology papers

Stories by subject **Report finds massive fraud at Dutch** . Brain and behaviour universities Lab life

Stories by keywords

- Deiderik Stapel
- Tilburg University
- Academic fraud <u>Retractions</u>

Social psychology

This article elsewhere

😫 Blogs linking to this article

Add to Connotea









Two papers published in Nature in January 2014 promised to revolutionize the way stem cells are made by showing that simply putting differentiated cells under stress can 'reprogram' them and make them pluripotent — able to develop into any type of tissue in the body. But soon, errors were found in the papers, and attempts to replicate the experiments failed. Haruko Obokata, the lead author, was found guilty of misconduct, the papers were retracted and the RIKEN centre where she worked was restructured. The aftermath of the episode has been felt by scientists across Japan, in the form of new anti-misconduct policies.

#### Stem-Cell Scandal



Child Co-Authors in Korea



UNIVERSITY OF GOTHENBURG krefting research centre

# Retracti@n Watch

https://retractionwatch.com/

www.gu.se



#### UNIVERSITY OF GOTHENBURG **krefting research centre**

# Replication

 Duplicating results of prior study using <u>independent</u> investigators, methods, data, equipment, and protocols

(Peng et al. Am J Epidemiol 2006; 163: 783-789)

- <u>Gold standard</u> of scientific investigation
- Credibility of scientific claim depends on its replicability
  - Not on authority of investigators



how science should be

- Replication often impossible
  - Differences between original and replicating study
    - Time
    - Setting
    - Investigators
    - Measurements
    - Protocol execution
    - Costs



# Reproducibility

 The ability of an investigator to duplicate the results of a prior study using the same methods as were used by the original investigator

Goodman et al. *Sci Transl Med* 2016; 8: 341ps12

 <u>Minimum standard</u> in place of replication



• <u>Synonyms</u>: Transparency, validation Verification. Processing trail, Open science, Corroboration



## IF A JOB IS WORTH DOING. IT IS WORTH DOING TWICE



Russell. Nature 2013; 496:7



UNIVERSITY OF GOTHENBURG **KREFTING RESEARCH CENTRE** 

# Reproducibility crisis?

#### Why Most Published Research Findings Are False

John P. A. Ioannidis

Ioannidis 2005; PLoS Med 2(8): e124.





#### THE PSYCHOLOGY BEHIND EXAGGERATED & FALSE RESEARCH

We often think that scientists are the most honest people around, and assume that scientific findings are reliable and true. But several new studies have revealed that an enormous number of researchers cut corners, cook data, and lie about results when conducting experiments. This is the world of bad science.



#### UNIVERSITY OF GOTHENBURG krefting research centre

### IS THERE A REPRODUCIBILITY CRISIS?

7% Don't know

> 3% No, there is no crisis

> > 38% Yes, a slight

A Nature survey lifts the lid on how researchers view the 'crisis' rocking science and what they think will help.

**BY MONYA BAKER** 

52% Yes, a significant crisis



#### NATURE | VOL 533 | 26 MAY 2016

#### HAVE YOU EVER TRIED TO PUBLISH A REPRODUCTION ATTEMPT?

Although only a small proportion of respondents tried to publish replication attempts, many had their papers accepted.

Published Failed to publish



#### HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?

Most scientists have experienced failure to reproduce results.





UNIVERSITY OF GOTHENBURG **krefting research centre** 

### HAVE YOU ESTABLISHED PROCEDURES FOR REPRODUCIBILITY?

Among the most popular strategies was having different lab members redo experiments.



NATURE | VOL 533 | 26 MAY 2016



### Prevalence of irreproducibility





### Economics of irreproducibility



www.gu.se



### Impacts of irreproducibility

- Undermines cumulative knowledge production
- False alarm false positive findings
- Contributes to delay in development of therapy
- Increases costs of therapeutic development
- Meta-analyses of studies may lead to wrong conclusion;
- Negative effects on medical guidelines
- Clinical trials based on faked data can lead to:
  - Harm
  - Wastage of resources and time
- Integrity of science
- Undermines public trust of science
- Career of perpetrators and co-authors



## What can be done?

**"NOT EVERY PAPER** NEEDS TO BE MEDICALLY RELEVANT BUT THEY SHOULD ALL BE REPRODUCIBLE.

Russell. Nature 2013; 496:7

"REPRODUCIBILITY IS LIKE BRUSHING YOUR TEETH. ONCE YOU LEARN IT, IT BECOMES A HABIT."

Baker. Nature 2016; 533: 452-454



### How to enhance reproducibility

•Reproducibility affects all research stages



### How to Make More Published Research True

#### John P. A. Ioannidis<sup>1,2,3,4</sup>\*

PLoS Med 2014; 11(10): e1001747

**1** Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, California, United States of America, **2** Department of Medicine, Stanford Prevention Research Center, Stanford, California, United States of America, **3** Department of Health Research and Policy, Stanford University School of Medicine, Stanford, California, United States of America, **4** Department of Statistics, Stanford University School of Humanities and Sciences, Stanford, California, United States of America

### Box 1. Some Research Practices that May Help Increase the Proportion of True Research Findings

- Large-scale collaborative research
- Adoption of replication culture
- Registration (of studies, protocols, analysis codes, datasets, raw data, and results)
- Sharing (of data, protocols, materials, software, and other tools)
- Reproducibility practices
- Containment of conflicted sponsors and authors
- More appropriate statistical methods
- Standardization of definitions and analyses
- More stringent thresholds for claiming discoveries or "successes"
- Improvement of study design standards
- Improvements in peer review, reporting, and dissemination of research
- Better training of scientific workforce in methods and statistical literacy



#### UNIVERSITY OF GOTHENBURG krefting research centre



Reproducibility and reliability of biomedical research: improving research practice

UK Academy of Medical Sciences Symposium report, October 2015 https://acmedsci.ac.uk/filedownload/38189-56531416e2949.pdf



### Open science initiatives





## Study design: register protocol



Develop, register, publish protocol prior to study conduct

Preregistration makes your science better by increasing the credibility of your results Helps prevents Publication bias



# Study design: register protocol

# ClinicalTrials.gov

https://www.clinicaltrialsregister.eu/

### EU Clinical Trials Register

https://clinicaltrials.gov/

EUROPEAN Network of Centres for Pharmacoepidemiology and Pharmacovigilance

http://www.encepp.eu/encepp/ studiesDatabase.jsp

https://www.crd.york.ac.uk/ prospero/ PROSPERO

International prospective register of systematic reviews

www.gu.se



# Why register study protocol?

- Open science practice
- Promotes transparency and reliability of results
- •Helps prevent certain research practices:
  - Flexible analyses
  - Post-hoc analyses
  - Selective outcome reporting, p-hacking, cherry picking
  - HARKing hypothesizing after the results are known
- Informs peers of ongoing research and enhances collaboration
- Increased likelihood of publication



### Reproducibility in data analysis

"A study is reproducible if all of the code and data used to generate the numbers and figures in the paper are available and exactly produce the published results"

Leek & Jager. Annu Rev Stat 2017; 4:109-22



# Reproducibility in data analysis

BMJ 2012;344:e4363 doi: 10.1136/bmj.e4363 (Published 26 June 2012)

#### EDITORIALS

Provine 1 and 2

### Open science and reproducible research

New reports call for scientists to share data and publishers to embrace open access

"Scientists should communicate the data they collect and the models they create, to allow free and open access, and in ways that are intelligible, assessable and usable for other specialists.... Where data justify it, scientists should make them available in an appropriate data repository."



### Reproducibility in data analysis

Item	Requirement
Data	Study data set made available, if possible
Code	Code used for results, tables, figures made available in human readable form. Software used also available
Documentation	Sufficient details of the analyses steps and code to help others repeat them
Distribution	Above made available through a platform others can have access to them



### Reproducibility across statistical software

- Good developments in R, Stata, SAS
  - Goal: bringing statistical code, data, and presentation together in a way for others to easily follow
    - Automate the process
    - Reduce human errors
  - Computationally advanced
  - Currently used primarily by those in statistics, engineering, computer sciences

### Reproducible Research with R and RStudio

The R Series

	Ro	ot	
	Research	Project	
Data	Analysis		Presentation
MainData.cov	MainAnalysis.R	Article	Other
Gatherfource	Bessindigues	Paper Street	Sideshow.
Maketile	Figure 1.R	Main.Mb	Südesbow Row
MergeData.R	Figure2.8	Packages.htb	Website
Gather1.8	Figure 3.8	figure	Website Artist
Gather2.N		Figure 1 pdf	
Gather3.8		Figure2.pdf	
		Pigure3.pdf	

#### **Christopher Gandrud**



## Develop an analysis plan

- Detailed outline of your analyses
- A recipe: step-by-step analysis decisions
- If possible, publish a protocol
- If possible, register your protocol
- Consult with investigators / collaborators

#### NH) U.S. National Library of Medicine ClinicalTrials.gov

Find Studies 
About Studies 
Submit Studies

ClinicalTrials.gov is a database of privately and publicly funded clinical studies conducted around the world.





### Data editing

- Clean and code your data before analyses
- Data editing: careful scrutiny of raw data for errors
- Code variables sensibly
- •Code variables in a reproducible way
  - Data coding syntaxes
  - If possible, publish syntaxes alongside paper
- Sensible management of data files crucial
- Sensible choice of categorization and variable transformation



# Reporting (dissemination) and reproducibility

### THE EQUATOR NETWORK:

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

www.gu.se



# Diversifying peer review

- Preprint platforms facilitating easy sharing and discovery of research before publication
  - arXiv physical sciences
  - bioRxiv and PeerJ life sciences
  - engrXiv engineering
  - PsyArXiv psychology
  - SocArXiv social sciences)

- Post-publication peer review platforms
  - PubMed Commons
  - PubPeer



### Ethos of scientific research

- •Robert K. Merton, 1942
- Most influential discourse of values of modern scientific research
- Articulated four core ethos (values)
  - Universalism
  - Communality
  - Disinterestedness
  - Organized skepticism



### Universalism

- Scientific findings are not personal
  - "pre-established impersonal criteria."
- To be valid, should not depend on investigator's
  - Personal
  - Social
  - Political
  - National characteristics/affiliations
- Science is anti-authoritarian: idea acceptance not because of the qualities of its protagonist
  - With right training, anyone can contribute to science



### Communality

### • Evidence generation should be through:

- Active collaboration within the scientific community
- Open exchange of ideas
- Discussion and consensus
- Sharing of evidence
- Secrecy is antithesis of scientific research
- Scientific research is not motivated by commercialization



### Disinterestedness

- Scientific research should be devoid of personal interest
  - Monetary motivations
  - Other interests
- •Overarching motivation should be to uncover the truth



### Organized skepticism

- Scientists should not take findings at face value
- Scrutiny should be strictly part of the process
- Active peer review process indispensable
- Replication/reproducibility indispensable