

GENETIQUE • ASTRONOMIE • PHILOSOPHIE • ANTHROPOLOGIE • ECOLOGIE • CHIMIE • BIOLOGIE • SOCIOLOGIE • MEDECINE • NEUROSCIENCES • INFORMATIQUE • HISTOIRE • **WWW.RECHERCHE-REPRODUCTIBLE.FR** • HISTOIRE • MECANIQUE • GENETIQUE • ASTRONOMIE • CHIMIE • BIOLOGIE • SOCIOLOGIE • PHYSIQUE • MEDECINE • NEUROSCIENCES • INFORMATIQUE • HISTOIRE • DROIT • ECONOMIE • MATHÉMATIQUES

Journées du Réseau Français de Recherche Reproductible

Grenoble

26-28 Mars 2024

SCIENCE • ROBUSTE • REPLICABLE • GENERALISABLE • EXPERIENCE • CODE • OBSERVATION • STATISTIQUES • ANALYSE • DONNEES • METHODE • ECHANTILLON • RESULTAT • RECHERCHE • SCIENCE • REPLICABLE • ROBUSTE • RECHERCHE • GENERALISABLE • EXPERIENCE • CODE • OBSERVATION



Organization of these days

The team !

- Pierre-Antoine Bouttier, GRICAD
- Juana Dos Santos, team assistant, LJK
- Violaine Louvet, LJK
- Antoine Orlandi, GRICAD
- Céline Acary-Robert, LJK/GRICAD/MAIMOSINE

with the support of the **steering committee of the network**

⇒ 2025 ?

Outline of the presentation

- Why this network about reproducibility ?
 - Context
 - Motivations
- Overview of the 2023 days
- Setting up the current structure
- The program !

Reproducibility crisis

- In the past decade: Lots of debates around reproducibility
- The term "reproducibility crisis" gained in popularity, and usually refers to*:
 - 1 absence of replication studies,
 - 2 failure to reproduce results of published studies,
 - 3 publication bias towards positive results,
 - 4 bad research practices,
 - 5 lack of transparency/completeness in the methods, data and analysis.

* taken from <https://plato.stanford.edu/entries/scientific-reproducibility/>

[//plato.stanford.edu/entries/scientific-reproducibility/](https://plato.stanford.edu/entries/scientific-reproducibility/)

History

Some fields that were particularly affected: Social sciences, Psychology, Clinical research;

RESEARCH ARTICLE

Estimating the reproducibility of psychological science

Open Science Collaboration^{*,†}

• See all authors and affiliations

Science 28 Aug 2015;
Vol. 349, Issue 6251, aac4716
DOI: 10.1126/science.aac4716

[...]

RESULTS

We conducted replications of 100 experimental and correlational studies published in three psychology journals using high-powered designs and original materials when available. There is no single standard for evaluating replication success. Here, we evaluated reproducibility using significance and *P* values, effect sizes, subjective assessments of replication teams, and meta-analysis of effect sizes. The mean effect size (*r*) of the replication effects ($M_r = 0.197$, $SD = 0.257$) was half the magnitude of the mean effect size of the original effects ($M_o = 0.403$, $SD = 0.188$), representing a substantial decline. Ninety-seven percent of original studies had significant results ($P < .05$). Thirty-six percent of replications had significant results; 47% of original effect sizes were in the 95% confidence interval of the replication effect size; 39% of effects were subjectively rated to have replicated the original result; and if no bias in original results is assumed, combining original and replication results left 68% with statistically significant effects. Correlational tests suggest that replication success was better predicted by the strength of original evidence than by characteristics of the original and replication teams.

Preclinical research (Freedman et al. 2015)

The Economics of Reproducibility in Preclinical Research

Leonard P. Freedman , Iain M. Cockburn, Timothy S. Simcoe

Published: June 9, 2015 • <https://doi.org/10.1371/journal.pbio.1002165>

Article	Authors	Metrics	Comments	Media Coverage
---------	---------	---------	----------	----------------

Correction

Abstract

Introduction

Defining Reproducibility

Analysis of Four

Categories of

Irreproducibility

Economic Impact of

Irreproducibility

The Role of Best

Practices and Standards

Conclusions

Supporting Information

Correction

10 Apr 2018: The PLOS Biology Staff (2018) Correction: The Economics of Reproducibility in Preclinical Research. PLOS Biology 16(4): e1002626. <https://doi.org/10.1371/journal.pbio.1002626> | [View correction](#)

Abstract

Low reproducibility rates within life science research undermine cumulative knowledge production and contribute to both delays and costs of therapeutic drug development. An analysis of past studies indicates that the cumulative (total) prevalence of irreproducible preclinical research exceeds 50% resulting in approximately US\$28,000,000,000 (US\$28B)/year spent on preclinical research that is not reproducible—in the United States alone. We outline a framework for solutions and a plan for long-term improvements in reproducibility rates that will help to accelerate the discovery of life-saving therapies and cures.

It can happen to everyone

<https://retractionwatch.com/2022/09/03/nobel-prize-winner-gregg-semenza-retracts-four-papers>

Nobel Prize winner Gregg Semenza retracts four papers

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 concerns about images in the articles.



Gregg Semenza

Gregg Semenza is "one of today's preeminent researchers on the molecular mechanisms of oxygen regulation," the work for which he shared the 2019 Nobel, according to Hopkins. But even before that, the pseudonymous Claire Francis began pointing out potential image duplications and other manipulations in Semenza's work on PubPeer, as described in October 2020 by Leonid Schneider.

The four papers retracted yesterday are:

[Continue reading →](#)

SHARE THIS:



Microbiome and cancer

Steven Salzberg ❤️👍👏 @StevenSalzberg1 · 1 août

Major, fatal errors found in the data and methods of a 2020 paper in @Nature, including millions of reads mis-identified as bacteria. The "cancer microbiome" in this study was simply not there. @abrahamghawi @elapertea @YuchenGe1 @JenniferLu717



bioRxiv
THE PREPRINT SERVER FOR BIOLOGY

Major data analysis errors invalidate cancer micro...
We re-analyzed the data from a recent large-scale study that reported strong correlations between ...

🗨️ 35 🔄 582 ❤️ 1502 📄 564,8 k ⬇️

Steven Salzberg ❤️👍👏 @StevenSalzberg1 · 3 août

New story in @statnews by @angrchen and @matthewherper about the big, big problems we discovered in a @Nature paper that reported finding a microbiome associated with 32 cancer types



STAT @statnews · 3 août

Computational biologist @StevenSalzberg1 says the problems with a Nature paper about a microbiome cancer diagnostic are serious. trib.al/N3M72f0

🗨️ 1 🔄 10 ❤️ 46 📄 28,6 k ⬇️

Steven Salzberg ❤️👍👏 @StevenSalzberg1 · 3 août

and once again, in the quotes from Knight et al., they don't address any of the problems in their study, instead just claiming that "other work" supports it. That doesn't fix the problems

🗨️ 2 🔄 9 ❤️ 9 📄 2 999 ⬇️

Steven Salzberg ❤️👍👏 @StevenSalzberg1

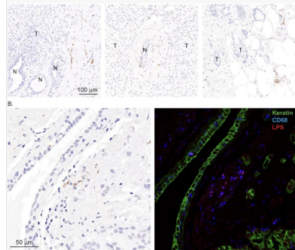
Yet another major blow to the hypothesis that a microbiome of cancer exists. TLDR: the main results from a 2020 @ScienceMagazine paper claiming to find bacteria in breast cancer simply doesn't hold up. Well done @NFdeMiranda, Jacques Neefjes, et al

Traduire le post

Noel F. de Miranda @NFdeMiranda · 29 août

🔍 In a bid to replicate a prior study, we couldn't confirm LPS presence within breast cancer cells. We did spot it around ducts & in macrophages, aligning with its biology.

#ResearchReplication #Cancer #Microbiome
[biorxiv.org/content/10.1101...](https://www.biorxiv.org/content/10.1101...)



Legend Figure 1 – A - Representative examples of LPS immunodetection with typical granular pattern. LPS expression never co-localized with cancer cells. B - Left: LPS detection of a breast cancer section

2:17 PM · 29 août 2023 · 96,5 k vues

Microbiome and cancer

Article | [Published: 11 March 2020](#)

Microbiome analyses of blood and tissues suggest cancer diagnostic approach

[Gregory D. Poore](#), [Evguenia Kopylova](#), [Qiyun Zhu](#), [Carolina Carpenter](#), [Serena Fraraccio](#), [Stephen Wandro](#), [Tomasz Kosciolk](#), [Stefan Janssen](#), [Jessica Metcalf](#), [Se Jin Song](#), [Jad Kanbar](#), [Sandrine Miller-Montgomery](#), [Robert Heaton](#), [Rana Mckay](#), [Sandip Pravin Patel](#), [Austin D. Swafford](#) & [Rob Knight](#) 

[Nature](#) **579**, 567–574 (2020) | [Cite this article](#)

78k Accesses | **481** Citations | **919** Altmetric | [Metrics](#)

> [bioRxiv](#). 2023 Jul 31;2023.07.28.550993. doi: 10.1101/2023.07.28.550993. Preprint

Major data analysis errors invalidate cancer microbiome findings

[Abraham Gihawi](#), [Yuchen Ge](#), [Jennifer Lu](#), [Daniela Puiu](#), [Amanda Xu](#), [Colin S Cooper](#), [Daniel S Brewer](#), [Mihaela Pertea](#), [Steven L Salzberg](#)

PMID: 37577699 | PMID: PMC10418105 | DOI: 10.1101/2023.07.28.550993

[Free PMC article](#)

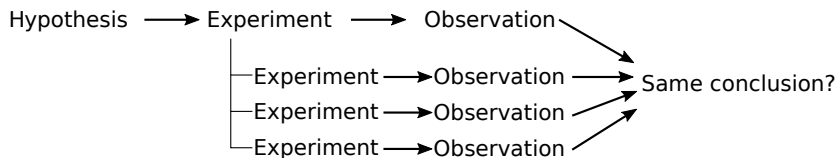
- Peer review is not perfect
- Importance of data and code availability
- Allows further discussions + careful observation of past studies

Many kinds of reproducibility

- **Empirical reproducibility** : non-computational empirical scientific experiments and observations enabled by making data freely available, as well as details of how the data was collected.
- **Statistical reproducibility**: choice of statistical tests, experimental design, model parameters, threshold values, etc.
- **Computational reproducibility**: detailed information about code, software, hardware and implementation details
 - Goal: document how data has been produced
 - In bioinformatics : many challenges described $\dot{\iota}$ 10 years ago

Experimental variability: consequences?

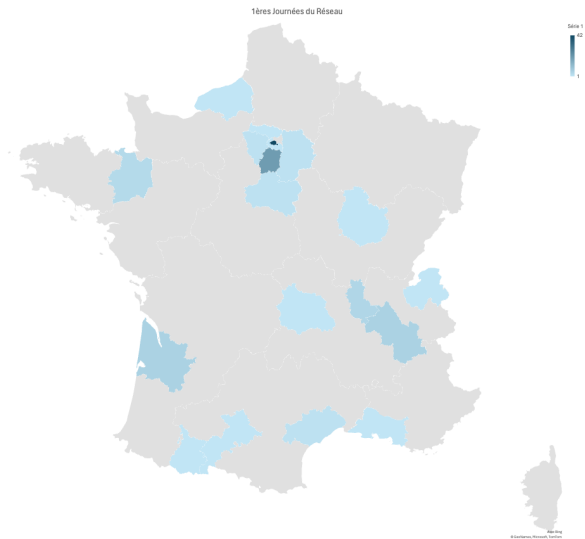
- The same experiment gives different results
- The same experiment leads to the same scientific interpretation (hopefully)



2023 network days: state of the art

- Institut Pasteur 8-9 march 2023
- Hybrid mode (srikes)
- 110 registrations
- ~ 20 talks with different points of view:
 - Observationnal reproductibility
 - Computational reproductibility
 - Statistical reproductibility
 - Experimental reproductibility
 - Teaching / Training
 - Europe / International

2023 network days: participants



2023 network days: topics

Aims: Determine the origin of irreproducibility across various domains

1 Observational

- History of sciences
- Humanities and Social Sciences
- Physics

2 Computational

- Molecular biophysics
- Software archival
- Neuro-imaging
- Machine learning

3 Statistical

- Bio-physics, biochemistry and evolution
- Clinical trials

4 Experimental

- Preparation of biological sample
- Meta-research
- Psychology
- Characterization of biological materials

5 Train students and professionals

- Reprohackathons
- ReproducibiliTea
- MOOCs
- University courses

6 European and international initiatives

- UKRN
- ITRN
- COSO

2023 network days: outcome

- Definition of the boundaries of reproducibility
- Identify topics and working groups
- Identify an active community (+mailing list, etc.)

French Reproducibility Network

Objectives



- Federate a community interested in the topic
- Support actions aimed at improving reproducibility
- Provide training materials to the community
- Promote strategies and policies in favour of reproducible research
- Integrate european and international networks

Presentation of the reproducible research network

Current state:

- Steering committee: 5 members
 - Céline Acary-Robert
 - Sarah Cohen-Boulakia
 - Arnaud Legrand
 - Frédéric Lemoine
 - Nicolas Rougier

- Network
 - Mailing-list (more than 150 subscribers !)
 - Website
 - Forum

We wish to **reinforce and enlarge the structure**

Presentation of the reproducible research network

Structure

We have proposed a possible **structuration with 3 main components**:

- Steering committee
- 3 transverse colleges
- Working groups on specific topics

Main objectives of this structure:

- Promote the sharing of good practises
- Encourage exchanges
- Suggest methods to improve reproducibility (role of WGs)
- Facilitate bibliographic work

Presentation of the reproducible research network

Setting up the structure

- Main steps
 - Finalization of the colleges
 - Policy of the steering committee constitution
 - New working groups: when necessary
- The dynamism of this network depends on members !

Discussion Thursday morning

- Sharing your opinions and ideas
- Call for volunteers (colleges and WGs)
- You are encouraged to share and note your contributions in this pad:
<https://semestriel.framapad.org/p/jrfr2024-a6iy>

Presentation of the reproducible research network

Ministry support

Ministry of higher education and research strongly supports our actions:

- Organization of these **network days**
- Open position (until March 31st) of a **Project Manager**

Position profile

- Mission: support the network's growth
- Activities :
 - Structuring the network
 - Identification of key players (national, european, global)
 - Coordination with other networks, universities, etc ...
 - Scientific watch
 - Contribution to trainings and scientific animation

Program

Mardi 26 Mars, 14h–17h : Contexte français et européen

- Science ouverte : de la transparence à la reproductibilité, Isabelle Blanc, MSER
- From local to national initiatives: How to make reproducible research the norm, Ulf Toechl, QUEST Center Berlin Institute of Health
- The Center for Reproducible Science at the University of Zurich, Eva Furrer, Center for Reproducible Science Managing Director, University of Zurich

Mercredi 27 Mars, 9h–12h : Reproductibilité expérimentale

- Réplicabilité et reproductibilité en psychologie expérimentale, Dominique Muller, LIP/PC2S, UGA
- Reproductibilité expérimentale en métabolomique, Estelle Pujos-Guillot, directrice scientifique de la Plate-Forme “Exploration du Métabolisme”, MetaboHUB, INRAE
- Ré-utilisabilité des données : l'exemple de la sismologie dans Epos-France, Jonathan Schaeffer, Isterre, OSUG
- Reproducibility in Photon Science, Andrew Goetz, ESRF

Program

Mercredi 27 Mars, 13h30–18h : Reproductibilité computationnelle

- **Café Guix Live** : Introduction à la reproductibilité des environnements de calcul : construction de paquets et liens avec Software Heritage, Ludovic Courtès, Inria / Pierre-Antoine Bouttier, Gricad
- **Software Heritage and IPOL**, a fruitful collaboration toward reproducible research, Miguel Colom-Barco, ENS Paris Saclay
- **Tutoriel** : Dépôt d'un package R sur Software Heritage et référencement sur HAL, Florent Chuffart, IAB (INSERM/CNRS), RIS (MITI/CNRS), CoSO (MESR)
- **Table ronde** : problématiques autour de la reproductibilité en IA, Alexis Arnaud (Gricad, UGA), Ninon Burgos (Aramis, CNRS), Ambre Davat (IPhiG et chaire "éthique&IA"), Arnaud Legrand (LIG, CNRS), Jean-Luc Parouty (SIMAP, CNRS), Didier Schwab (LIG, UGA)

Program

Jeudi 28 Mars, 8h45–12h : Formation et gouvernance

- Restitution du GT formation, Frédéric Lemoine, Institut Pasteur
- Restitution du GT notebooks, Raphaëlle Krummeich, Université de Rouen, Sébastien Rey-Coyrehourcq, Université de Rouen
- Retour d'expérience sur le MOOC recherche reproductible, Konrad Hinsén, CBM
- Discussion autour des GTs : fonctionnement, besoins, attentes
- Discussions sur la structuration et la gouvernance du réseau