Show notes

Doing Science Differently

#2 - The Reproducibility Project: Cancer Biology [01:17:57]

Guest: Timothy M. Errington, PhD

November 15th, 2022

In this episode of *Doing Science Differently*, we will hear more about reproducible and open science practices or the lack thereof. Research is both innovative and self-correcting - or maybe not? Are we losing robustness by constantly aiming for novelty and breakthrough results? In 2011, the Center for Open Science set out to estimate the robustness of findings in the field of cancer biology by trying to replicate 50 landmark papers published between 2011-2012. In the second episode of our podcast, we will have a conversation with Timothy Errington, Head of Research at the Center for Open Science, and manager of the Reproducibility Project: Cancer Biology. We will learn more about replicability and factors that influence it, challenges encountered during the project, and what researchers can do to change the lack of replicability in scientific publications and their attitude towards errors.

In this episode, you will hear more about:

[02:21] the reasons that got Tim invested in the question of reproducibility in cancer biology

> *"[...], I spent the first six months trying to reproduce somebody else's experiment and the paper that just got published. And that person was in the lab, I had access to like all their notes, to the person and could talk to the person. Couldn't figure it out."*

[07:18] the findings of the project, and some of the roadblocks and surprises Tim and his team encountered

> *"[...] yeah, we could see things in the same direction, we could see effects, but most of them were not the large effects that we got excited for in the first place."*

[30:17] the idea behind the pre-registration of the study

> *"[...] please tell us how to make it the best before we invest any research into it."*

[33:14] examples of challenges encountered, and why informative and complete method sections are important

> *"So that's a good example of where we just ran into challenges. Couldn't figure it out."*

[40:16] context-dependent versus robust findings, and the power of replication

> *"You and I should be able to change things that we don't think that matter. And it doesn't matter, I get a similar finding. So that's the beauty of replication."*

[44:38] how original authors reacted differently towards the replication project

> *"[...] you know what, I'd help you if you weren't doing a replication."*

[52:54] approaches researchers could try to improve the situation

> *"[...] summarize your methods in the paper, and then put a link to the actual protocol."*

[01:03:56] the need to embrace the culture of correcting

> *"[...] oops, I made a mistake. Let me fix it. So, the next person coming along reading it gets the right representation."*

[01:07:43] the need for replication studies and systemic change

> *"[...] that pendulum is probably swung a little too far on high impact innovation, positive shiny results and rush, rush, rush."*

Guest biography

This episode's guest is **Timothy M. Errington**, PhD. Tim successfully finished his studies in biology and chemistry at St. Lawrence University (bachelor's degree) and in molecular and cell biology at UC Berkeley (master's degree). After receiving his PhD in microbiology, immunology, and cancer biology from the University of Virginia in 2013, he started working as Project Manager at the Center for Open Science (COS) that aims to increase the reproducibility, openness, and integrity of scientific research. He continued his career at the Center and became Metascience Manager and later Director of Research. Since January 2022, Tim has been the Senior Director of Research at COS. His work at COS allows him to collaborate with researchers coming from different disciplines and to participate in large scale projects such as the Reproducibility Project: Cancer Biology.

List of resources

[01:00] The Center of Open Science aims towards the increase of openness, integrity, and reproducibility of research. If you want to find out more, visit: https://www.cos.io/?hsLang=en

[02:53] The Reproducibility Project: Cancer Biology aimed at replicating experiments from high-end articles from the field of cancer biology published between 2010-2012. The completion of the project took 8 years, and researchers involved in this effort faced a number of different challenges as outlined by Tim in the interview. Learn more about the project here: https://www.cos.io/rpcb.

[03:21] The Bayer/Amgen studies are reports from the biotechnology company Amgen and the pharmaceutical company Bayer HealthCare in which they tried to replicate research findings from several landmark studies that they were planning to base new products on. However, the majority of investigated studies could not be replicated.

Begley and Ellis, Raise standards for preclinical cancer research. *Nature.* 2012; **483**, 531-533. https://doi.org/10.1038/483531a

Prinz, Schlange and Asadullah, Believe it or not: how much can we rely on published data on potential drug targets. *Nature Reviews Drug Discovery*. 2011; **10**, 712. https://doi.org/10.1038/nrd3439-c1

[06:53] Collection of studies and report published as part of the Reproducibility Project: Cancer Biology (https://www.cos.io/rpcb). The project had a closer look at the reproducibility of published findings in the field of cancer biology, and tried to identify factors that influence the success or failure of reproducibility in science more generally. To access the publications, please visit: https://elifesciences.org/collections/9b1e83d1/reproducibility-project-cancer-biology

[11:51] The power debate - small sample sizes overestimate effects: Very briefly, it is difficult to reliably quantify the correct mean and standard deviation of an effect with only few data points. Thus, the estimated size of an effect is not very stable across experiments with a low N. In other words, with little data, individual measurements strongly influence the statistical outcome and the measured effect can easily deviate from the real effect. At the same time, very strong effects are necessary to pass the common significance threshold ($p<0.05$) with low sample size, so reporting based on significance testing effectively selects for high effect sizes. In this combination, strong effects from low-powered studies are more likely to be one-off events with overestimated size. If you want to learn more about this, we recommend the following resources:

Button, Ioannidis et al., Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*. 2013; **14**, 365–376. https://doi.org/10.1038/nrn3475

Krzywinski & Altman, Power and sample size. *Nature Methods*. 2013; **10**, 1139-1140. https://doi.org/10.1038/nmeth.2738

http://daniellakens.blogspot.com/search?q=sesoi

[27:58] Science Exchange is an online R&D platform that provides access to scientific research services from multiple contract research organizations. The Reproducibility Project: Cancer Biology was performed in collaboration between the Center of Open Science and Science Exchange. For more information please visit: https://ww2.scienceexchange.com/s/.

[29:17] DARPA is the Defense Advanced Research Projects Agency, a United States government agency that invests in the development of new technologies for use by the military: https://www.darpa.mil/.

[55:47] Protocols.io is an open access and free repository that can be used to share up to date detailed protocols as well as research methods. This allows researchers to share detailed protocols in addition to the methods sections in publication. To find out more, go to: https://www.protocols.io/. The publisher PLOS also encourages the publishing of protocols: https://plos.org/protocols/.

[56:11] Data repositories, also known as data libraries or data archives, help researchers to organize and manage their data, allow for the long-term storage of data, and can facilitate the discovery of data by fellow researchers. Tim mentions two repositories here, Open Science Framework (OSF;

https://www.cos.io/products/osf) and Figshare (https://figshare.com/). There are many more repositories to deposit reagents, data, or code. These include Addgene (https://www.addgene.org/), ATCC (https://www.atcc.org/), Zenodo (https://zenodo.org/) and GitHub (https://github.com/).

[01:05:39] Executable Research Articles (ERAs) are enriched articles published in the open access journal *eLife*. Authors can enhance their submitted publications by including data, interactive figures as well as live code creating the figures. This aims at increasing the reproducibility and transparency of scientific findings, and fosters the trust of the articles' readers. You can find more information on ERAs here:

https://elifesciences.org/labs/dc5acbde/welcome-to-a-new-era-of-reproducible-publishing

https://elifesciences.org/for-the-press/eb096af1/elife-launches-executable-research-articles-for-publishing-computationally-reproducible-results

[01:10:57] Contract Research Organizations (CROs) provide specific research services to pharmaceutical, biotechnology, and medical device industries. They, for example, plan, conduct, and supervise clinical studies themselves or support pharmaceutical companies that conduct these studies.

[01:13:37] The Brazilian Reproducibility Initiative is a project to systematically assess the reproducibility of findings in Biomedical Science published in Brazil. You can learn more here: https://www.reprodutibilidade.bio.br/home.

Tim also discusses the need for change in the way we conduct experiments and publish findings. The scientific community should establish a replication and pre-registration culture. For further information on pre-registration and registered reports, please visit: https://www.cos.io/initiatives/prereg and https://www.cos.io/initiatives/registered-reports. Further, a correction culture that includes more levels than just rectraction needs to be established. Check out this example of a correction in one of the publications from the Reproducibility Project (https://elifesciences.org/articles/71601#annotations:BDMZslnJEeyMjfNlLwsTUg.

You can find out more about the 'Doing Science Differently' Podcast-team under https://www.bihealth.org/en/quest/projects/spokes/think-tank-members.

Sponsored by: SPOKES Wellcome Trust Funded Think Tank. For more information visit https://www.bihealth.org/en/quest/projects/spokes.

Sound editing and original music from Giorgio Cattaneo (https://linktr.ee/exoplanetrec, giorgio.sauro)

Logo design and social media management from Anita Waltho (@Walthoanita)