

**Article:**

John P A Ioannidis.

*Reproducibility: Has Cancer Biology Failed beyond Repair?*Clin Chem 2022; 68(8): 1005–7. <https://doi.org/10.1093/clinchem/hvac030>**Guest:** Dr. John Ioannidis from the Meta-Research Innovation Center at Stanford University.

Bob Barrett:

This is a podcast from *Clinical Chemistry*, sponsored by the Department of Laboratory Medicine at Boston Children's Hospital. I am Bob Barrett.

Cancer is the second major cause of death worldwide and its burden of disease increases over time. Discovering, developing, and implementing useful cancer interventions is a top priority and pressure is high on regulatory agencies to accelerate cancer-related approvals. However, 90% to 95% of cancer drugs that enter clinical trials are not approved. Worse, among the drugs licensed in the recent past, the accumulated evidence has been disappointing. These interventions typically improve surrogate endpoints but achieve minimal or no survival improvements.

Cancer biomarkers is a field of research that similarly results in many thousands of publications annually but rarely translates to major tools saving lives. Why are we failing? That is the question posed by Dr. John Ioannidis in a perspective article appearing in the August 2022 issue of *Clinical Chemistry*. He is both Professor of Medicine and Professor of Epidemiology and Population Health at Stanford University, where he is leading the Meta-Research Innovation Center. His work includes improving research methods and practices and the reproducibility of scientific investigation and he is our guest in this podcast.

First of all Dr. Ioannidis, in your perspective article appearing in *Clinical Chemistry*, you cite findings from the Reproducibility Project in Cancer Biology. Can you tell us a bit about that project and what you found surprising in their report?

John Ioannidis:

This was a monumental project that tried over a number of years to reproduce the results of 193 experiments on cancer biology, a frontier field, one of the key core preclinical research areas. And unfortunately, they could do that only for 50 because in the remaining they could not even get the experiments done. It was not possible to understand what the protocol was and what was happening, and among those 50 that they did manage to at least do the experiments, their results were very different compared to the original publications that have been heavily cited and very prestigious

in prestigious journals. On average, 85% of the effect size of the original paper vanished. Only 1/7 of the original effect remained and when it was a black and white or a yes or no type of question, in the majority of cases the original result was not seen again.

Bob Barrett: Is such low reproducibility typical only for preclinical cancer research? What about research in the treatment of other diseases or conditions?

John Ioannidis: We have a number of reproducibility efforts happening over the last ten years that cover very diverse fields in biomedicine ranging from early basic research, preclinical, clinical, but also in many other sciences. Actually, psychology and economics and other fields have taken reproducibility seriously and the results can be seen as half empty or half full but clearly, there's a very high rate of lack of reproducibility. When it comes to preclinical research and biomedical research, the rates of non-reproducibility are very high, so this is not an outlier. This is a result that fits very well. In a frame of many other results, there can be debates about how even exactly to define that something is reproducible or not and what it means but these results are suggestive that we have a generic problem not something that is very specific to one field or subfield.

Bob Barrett: How can reproducibility be improved in cancer research, or in other areas for that matter?

John Ioannidis: There's many scientists who have turned their attention to finding solutions to the problem and of course set up that we have to make sure that we don't throw out the baby with the bathwater. Reproducibility is an indicator, it doesn't mean that an original research result is wrong. It could be that the reproducibility effort is wrong. It could be that both of them are correct but something is different. It could be that both of them are wrong but it's a signal that that we need to do something, and many efforts are focusing on improving transparency, improving openness, improving data sharing, improving sharing of code, having pre-registration of the protocols so that people would know what was promised to be done and what was done, independent efforts of validation, improvement of research methods, improvement of the peer review process and the way that we publish and disseminate research, improvement of independent validity checks including post-publication peer review, how do we correct the literature, how do we improve on having rigor and quality in the scientific literature.

As you realize, this is a vast frontier. It's pretty much the whole frontier of science because every aspect of science can be visited and revisited on how we do things, how we

communicate things, how we accept or discard observations and discoveries or validations.

Bob Barrett: Is it fair to expect that we can achieve perfect reproducibility in preclinical research?

John Ioannidis: I think that that would be a utopian goal. Perfection is not something that science can aim for. We struggle with error, we struggle with biases, we struggle with inadequate equipment and tools and theories and observations and measurements, so perfection is not attainable. The issue is can we improve by a certain number of points, if we talk percentage-wise, even 1% improvement. If we can improve science by 1% then that would be an amazing achievement because you realize that there's millions of papers being published and we have about 200 million scholarly works that have been published within the scientific corpus, so 1% improvement is an amazing feat. 10% improvement would be astronomical and I think that some of these interventions that are being discussed and debated and some evidences accruing on them on improving research premises may be achieving some improvement that is worthwhile.

Bob Barrett: Readers of *Clinical Chemistry* would be interested in progress made in developing new cancer biomarkers so that disease might be detected earlier. Have there been reproducibility checks in that area, and if not, should there be?

John Ioannidis: So, the answer to that question is that there is concern about the validity of different biomarkers. To my knowledge, there is no organized reproducibility check that is as systematic as what happened in cancer biology, although some of the cancer biology items that people try to reproduce may be seen as neighboring this type of discipline. There're many studies that try to repeat or investigate independently associations with different biomarkers. There's a vast biomarker literature, so one can see those as efforts to reproduce. Of course, people might say that I do things a little differently and I'm trying to build a parade or puzzle of lots of observations. Much of that empirical work suggests that these biomarkers, when they're originally proposed, they find very strong results in some cases and strong effects and strong ability to tell what will happen, but then the subsequent studies suggest that the predictive ability is much smaller, so this is congruent with what we have seen in other fields, inflated results, excess significance, overpromising observations early on that very often do not really translate into something that is as strong in subsequent efforts, let alone in something that helps improve clinical outcomes, because eventually, this is what we care about.

Bob Barrett: Well finally doctor, the current issue of *Wired* magazine suggests that sloppy use of machine learning is causing a

reproducibility crisis in science as we rely more on artificial intelligence. I would love to hear your opinion on whether you think that situation would actually worsen.

John Ioannidis: Machine learning and artificial intelligence are excellent tools in an area of tools that we have analytically, and our analytic capacity is improving at a very rapid pace in biomedicine, in conjunction with data science, so we should not dismiss them as just leading us down the slope of disaster. I think that the challenge is that many of these new tools are extremely complex, they're very convoluted. In order to be transparent about what exactly is happening, what exactly has happened in processing the data, in analyzing them, in converting them, in presenting them, you need to have extra steps of transparency. It's not as easy as just conveying the results of a two by two table so there's more opportunities for error, for bias, for misrepresentation, for things to go wrong, in a more convoluted path. There's probably greater ability to do more fancy things but these fancy things come at a cost of lack of transparency unless one pays extra attention to transparency, and eventually to reproducibility. Also, there's often a disconnect between some very basic principles of epidemiology that we've known for many, many decades pertaining to what data are we working with? Are they fundamentally sound? Are they okay to start with? Just throwing a very fancy machine learning apparatus or artificial intelligence on sloppy data, on miserably sloppy data much of the time, is not going to help.

It will just validate noise. It will make noise sound as if it is the king of evidence and I think that this is a danger that we see very often materialized when there's disconnect between understanding the data and their strengths, and particularly their weaknesses, and how these might be transmuted, augmented perhaps, upon application of a very powerful analytical tool.

Bob Barrett: That was Dr. John Ioannidis from Stanford University. He has been our guest in this podcast on "Reproducibility: Has Cancer Biology Failed beyond Repair?" He is author of a perspective article with that title that appears in the August 2022 issue of *Clinical Chemistry*. I am Bob Barrett, thanks for listening.