

Brief Opinion

The Epistemological (Not Reproducibility) Crisis

John Park, MD^{a,b,*}

^aDepartment of Radiation Oncology, Kansas City VA Medical Center, Kansas City, Missouri; and ^bDepartment of Radiology, University of Missouri – Kansas City, Kansas City, Missouri

Received 17 January 2020; revised 17 April 2020; accepted 27 July 2020

Abstract

The current reproducibility crisis is fundamentally a crisis of knowledge, thus in reality it is an epistemological crisis. The current reigning paradigm of null hypothesis testing using a P value of $<.05$ has made the medical literature prone to be filled with spurious correlations rather than true knowledge. This article brings attention to 3 foundational issues to help navigate the current crisis: The problem of induction, the concept of epistemological access, and the iatrogenics of information. Scientific reasoning is inductive reasoning and the problem of induction highlights the limitations of such knowledge. The concept of epistemological access is introduced to describe the inability of low-level data to extract true findings. This lack of true knowledge brings with it the iatrogenics of information, where having more data are in fact harmful and can lead to patients receiving ineffective treatments.

© 2020 Published by Elsevier Inc. on behalf of American Society for Radiation Oncology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

In 2005, John Ioannidis published his seminal article titled “*Why Most Research Findings are False*,” which was the catalyst that sparked the current reproducibility crisis.¹ This led to the claim that up to 90% of all research findings are false.² What exactly is the problem with irreproducibility? The problem with irreproducible findings of research are that the data contain just spurious correlations, not representative of true knowledge, and thus cannot be reproduced and verified. As highly trained physicians with strong scientific backgrounds, we incorporate new information based on clinical trial data. The issue is thus being certain of one’s knowledge based on the published findings of new research. The study of certainty, justified beliefs, and knowledge fall under the field of epistemology. In other words, what we have is not fundamentally a crisis of reproducibility, but of knowledge, that is an epistemological crisis.

Now why is this termed a “scientific crisis”? The term “scientific crisis” was made famous by Thomas Kuhn, in his landmark book *The Structure of Scientific*

Revolutions.³ Kuhn gave prescient insight into what science is and an incredible description of what happens during scientific revolutions (Fig 1). In light of all this, our main epistemological issue is the prolific wrong use of null hypothesis statistical testing with P values of $<.05$ as our reigning paradigm. The origin of the P value cutoff is from R.A. Fisher, who noted, “It is usual and convenient for experimenters to take 5% as a standard level of significance, in the sense that they are prepared to ignore all results which fail to reach this standard.”⁴ Gerd Gigerenzer notes, “Fisher’s assertion that 5% (in some cases, 1%) is a convention that is adopted by all experimenters and in all experiments, and nonsignificant results are to be ignored, became part of the institutionalized hybrid logic.”⁵ This has led to the American Statistical Association to declare 2 strong warning statements in 2016 and 2019, about “statistical significance” based on this cutoff.^{6,7} Many are gaming the system with methods such as “ P -hacking” or “data-dredging” on data sets, and this, coupled with medical journal inflation, there is now much more noise than signal, especially in the setting of “publish or perish.”⁸ To help guide us out of this crisis there are 3 foundational issues that must be understood:

Sources of support: none.

Disclosures: Dr Park is a paid speaker for Plexus Communications.

* Corresponding author: John Park, MD; E-mail: john.park@va.gov

<https://doi.org/10.1016/j.adro.2020.07.019>

2452-1094/© 2020 Published by Elsevier Inc. on behalf of American Society for Radiation Oncology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

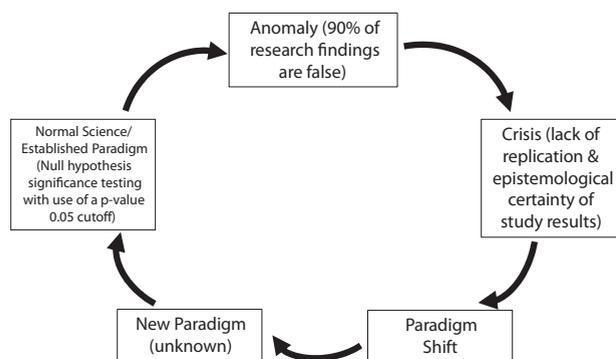


Figure 1 Kuhn's Cycle of Scientific Revolutions.

The problem of induction, the concept of epistemological access, and the iatrogenics of information.

The Problem of Induction

What are we actually doing with scientific studies and how we can trust their results? Philosophers of science have struggled mightily with these kinds of questions for the past 250 years. Are only verifiable statements true? What is line of demarcation between science and pseudoscience? Do the theories we hold affect how we interpret data? Unfortunately, these topics are not given a moment's notice in most medical school curricula. Einstein forewarned us saying, "The man of science is a poor philosopher."⁹ The fundamental issue with any scientific endeavor is how can we make universal generalizations from particular experimental data. How do we take data from the small number of patients in clinical studies and generalize them to all patients with the same condition? This is known as the problem of induction made famous by the empiricist philosopher David Hume. So great is the problem that in his 1946 lectures on Hume, W.V.O. Quine, arguably the most influential philosopher of the 20th century, noted that the "Humean predicament [the problem of induction] is the human predicament."¹⁰ This is the first hurdle we need to understand and overcome—the limitation of scientific knowledge. The classic example of the problem of induction relates to the discovery of black swans. If you see 1 white swan, 2 white swans, 3 white swans, *ad infinitum*, you cannot logically conclude that all swans are white. It took only 1 single observation of a black swan first seen by Willem de Vlamingh in 1697 during his voyage to Australia to dismiss the notion that all swans are white despite countless empirical observations of white swans in the West.¹¹ David Hume, applied this to sunrises where he noted that after seeing multiple sunrises, he could not logically conclude that the sun would rise tomorrow. Philosopher Stephen Law of the University of London notes, "Hume admits that ... why we believe the sun will

rise tomorrow does not supply the slightest grounds for supposing that this belief is true ... It is, again, a belief we simply find ourselves 'stuck' with."¹² Why is this pertinent to the field of oncology? Allow me to put it this way: If we treated 1 patient with treatment X, 2 patients with treatment X, 3 patients with treatment X ... *n* patients with treatment X, therefore we know treatment X works. This is a logical fallacy and surely less certain than if the sun will rise tomorrow. The key point here is that scientific reasoning is inductive reasoning and thus the problem of induction looms over all scientific endeavors. In order cross the inductive bridge, we run randomized controlled trials (RCT) grounded in objective measures based on mathematics in the form of statistics. Statistician James Berger notes "[s]cientists hold up objectivity as the ideal of science, but often fail to achieve it in rather spectacular fashion ... it is statistics that the scientists often call upon to provide objective validation of what they do ... they come to statistics in large part because they wish it to provide objective validation of their science."¹³ Unfortunately, physicians in general do not know how to wield this weapon. For instance, how many of us have looked at the combined European Organisation for Research and Treatment of Cancer (EORTC) 22931/Radiation Therapy Oncology Group (RTOG) 9501 extracapsular extension and positive margin analysis and asked ourselves of the validity of this post hoc analysis and if was adjusted for multiplicity?¹⁴ Much more humbling is the notion that we have not and frankly cannot.

Epistemological Access

Early phase data, observational studies, database, and retrospective reviews cannot give us the confidence to overcome the inherent limitations of empirical knowledge. It does not give, what I call, "epistemological access" to the data. To better grasp this concept, we need to make an important distinction between "what is true" (ontology) with the "how do we know it is true" (epistemology). As an example, if we were to travel back into past and pronounce to our past self that concurrent chemoradiation followed by immunotherapy for lung cancer improves survival compared with concurrent chemoradiation alone. This is true, but why should our past self believe it? It is one thing to say (1) concurrent chemotherapy followed by immunotherapy improves survival, and (2) we are able to know concurrent chemotherapy followed immunotherapy improves survival due to the PACIFIC trial data.¹⁵ This inability to know outcomes, whether they are actually true or not, without high quality trial data is the concept of "epistemological access" (Fig 2). The evidence-based pyramid is based on this idea by noting that on higher levels of the pyramid, we have access to higher levels of confidence in our data. A group at the University of Michigan found no correlation between observational

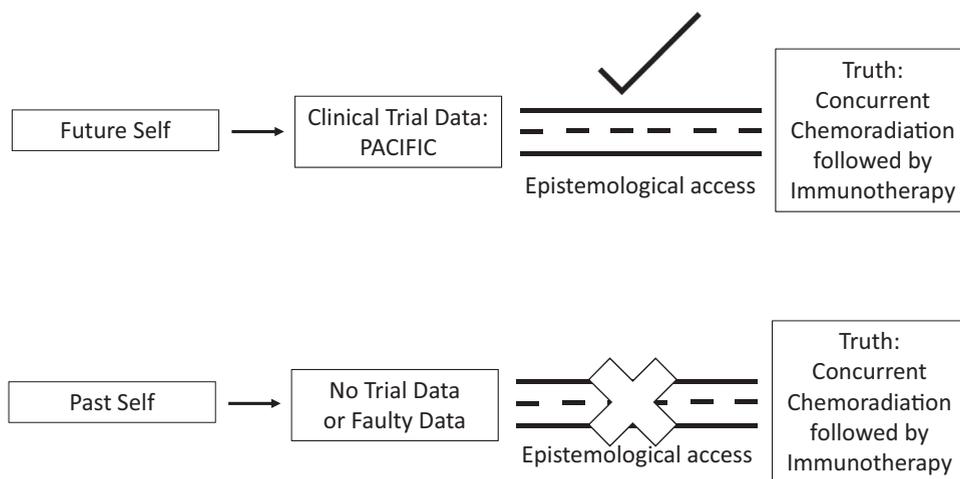


Figure 2 The roads of epistemological access.

studies from the Surveillance, Epidemiology, and End Results (SEER) and National Cancer Database (NCDB) data, regardless of the statistical rigor, with RCT data for the same treatments.¹⁶ The observational studies had no epistemological access to the true effect found in RCTs. Other examples include a study that found 81% of chemotherapeutic regimens in phase III trials had a lower response rate than the phase II studies that preceded them and none of the phase II characteristics could predict for positive phase III studies.¹⁷ Again, we know the benefits of phase III trials (elimination of selection bias, balance of treatment allocation, reduction of unexplained variances, etc), but what I want to emphasize is that they are also the primary means of epistemological access to the true effects of treatment. Some may ask at this point, “isn’t it better to have some data than no data.” This could not be further from the truth. Information can cause harm. To ensure we remain faithful to our oath of *primum non nocere* (first, do no harm) we must recognize the iatrogenics of information.

The Iatrogenics of Information

Nassim Taleb, in his book *Antifragile*, gives an apt example of signal and noise.¹⁸ He states that looking at stock prices yearly would give a signal to noise ratio of 50%:50%, daily looks 5%:95%, and hourly 0.5%:99.5% (ie, with more information there is an exponential decrease in the signal to noise ratio, with noise predominating exponentially). This ever-increasing noise is something that is being observed more in the scientific literature. The Open Science Collaboration found that only 36% of $P < .05$ results were found to be reproducible.¹⁹ Vinay Prasad was able to find 396 medical reversals of low value practices after RCTs were performed.²⁰ Given all the low quality data and much of it noise, even with our great training and thoroughly researched bioplausibility of many treatments, we cannot overcome these epistemological hurdles. In fact,

we are more likely to prescribe the wrong treatment based on these studies, thus it is imperative that we wait for high quality data, no matter the temptation to quickly use the latest database, retrospective, phase II, or subgroup analysis from negative RCTs (Fig 3). Unfortunately, this remains a perennial issue. Two articles published 6 years apart regarding the radiation oncology workforce, by MD Anderson, came up with diametrically opposite conclusions. In 2010, they concluded that the “[d]emand for radiation therapy is expected to grow 10 times faster than supply,” but only 6 years later they said “[t]he supply of radiation oncologists is expected to grow more quickly than the demand for radiation therapy.”^{21,22} It comes as no surprise as low-level data was used with wrong generalizations, made from faulty inductive inferences, that their predictions failed, as they did not have epistemological access to accurate outcomes. Now the iatrogenics of this information have affected the job market, which is the principal concern for graduating residents.²³ Furthermore, the head and neck radiation oncology group at Memorial Sloan Kettering published in 2019 their use of 30 Gy

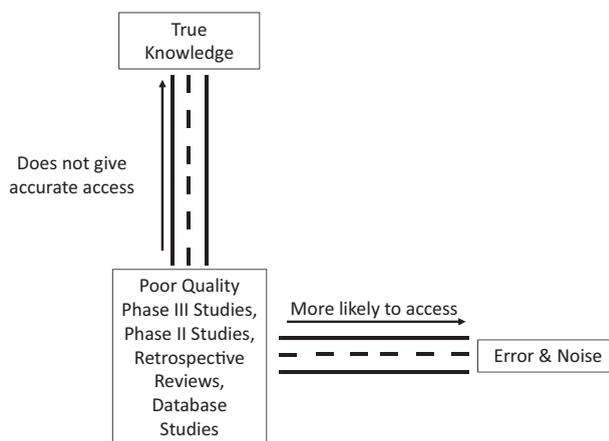


Figure 3 The roads of information.

elective nodal irradiation as their standard of care using only early phase and retrospective data saying “[b]ased on our interpretation of published data, we use the reduced-dose prescriptions described in this article as our current standard of care.”²⁴ In response, Robert Amdur, notes “I need stronger evidence to make the change in my practice. The price of a neck recurrence in this setting is huge because salvage surgery is morbid and rarely successful.”²⁵ The low quality data that is now being used as “standard of care” shows that even the brightest in our field are susceptible to these pitfalls. Both these errors could have been avoided if the limitations of the inductive data were considered, knowing there is a lack of epistemological access to know true future workforce projections or efficacy of reduced elective nodal irradiation, and being wary of the iatrogenics of the information of these predictions.

Why is it that we find it so difficult to go off this path? In Kuhn’s own words, “the man who reads a science text can easily take the applications to be the evidence for the theory, the reasons why it ought to be believed. But science students accept theories on the authority of teacher and text, not because of evidence ... The applications given in texts are not there as evidence but because learning them is part of learning the paradigm at the base of current practice.”³

This article’s goal is to point out some of the fundamental problems coupled with real-life examples of the current crisis. We must not further contribute to the dilemma by continuing to use weak data and perhaps not even publish such findings. We need to realize that to be active participants in fixing this crisis, physicians must take it upon themselves to be well versed in these issues, otherwise the reins of the science of medicine will be driven by others to the detriment of our patients.

References

- Ioannidis J. Why most published research findings are false. *PLoS Med.* 2005;2:e124.
- Topol E. Ioannidis: Most research is flawed; let’s fix it. Medscape. Available at: <https://www.medscape.com/viewarticle/898405>. Accessed December 20, 2019.
- Kuhn T. *The Structure of Scientific Revolutions*. 2nd ed. Chicago, IL: The University of Chicago Press; 1970.
- Fisher RA. *The Design of Experiments*. 8th ed. Edinburgh, UK: Oliver & Boyd; 1966:316 (as cited in Gigerenzer G. *The Superego, the Ego, and the ID in Statistical Reasoning in A Handbook for Data Analysis in the Behavioral Sciences: Methodological Issues*. Gideon K, Lewis C, eds. Hillsdale, NJ: Lawrence Erlbaum Associates; 1993).
- Gigerenzer G. In: Gideon K, Lewis C, eds. *The Superego, the Ego, and the ID in Statistical Reasoning in A Handbook for Data Analysis in the Behavioral Sciences: Methodological Issues*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1993:316.
- Wasserstein R, Lazar N. The ASA statement on *P* values: Context, process, and purpose. *Am Stat.* 2016;70:129-133.
- Wasserstein R, Schirm A, Lazar N. Moving to a world beyond “*P* < .05.”. *Am Stat.* 2019;73(Sup1):1-19.
- Grimes D, Bauch C, Ioannidis J. Modelling science trustworthiness under publish or perish pressure. *R Soc Open Sci.* 2018;5:171511.
- Einstein A. Physics and reality. *J Franklin I.* 1936;221:349-382.
- Pakaluk M. Quine’s 1946 lectures on Hume. *Journal of the History of Philosophy.* 1989;7:445-459.
- En.wikipedia.org. Willem de Vlamingh. Available at: https://en.wikipedia.org/wiki/Willem_de_Vlamingh. Accessed December 20, 2019.
- Stephen L. Problem of induction explained simply... (from my book The Philosophy Gym). Stephenlaw.blogspot.com. Available at: <https://stephenlaw.blogspot.com/2012/10/problem-of-induction-explained-simply.html>. Accessed December 20, 2019.
- Berger J. The case for objective Bayesian analysis. *Bayesian analysis.* 2006;3:385-402 (as cited in Mayo D. *Statistical Inference as Severe Testing: How to Get Beyond the Statistics Wars*. Cambridge, UK: Cambridge University Press; 2018).
- Bernier J, Cooper J, Pajak T, et al. Defining risk levels in locally advanced head and neck cancers: A comparative analysis of concurrent postoperative radiation plus chemotherapy trials of the EORTC (#22931) and RTOG (# 9501). *Head Neck.* 2005;27:843-850.
- Antonia S, Villegas A, Daniel D, et al. Overall survival with durvalumab after chemoradiotherapy in stage III NSCLC. *N Engl J Med.* 2018;379:2342-2350.
- Soni P, Hartman H, Dess R, et al. Comparison of population-based observational studies with randomized trials in oncology. *J Clin Oncol.* 2019;37:1209-1216.
- Zia M, Siu L, Pond G, et al. Comparison of outcomes of phase ii studies and subsequent randomized control studies using identical chemotherapeutic regimens. *J Clin Oncol.* 2005;23:6982-6991.
- Taleb N. *Antifragile: Things that Gain from Disorder*. New York, NY: Random House Publishing Group; 2014:126.
- Aarts A, Anderson J, Anderson C, et al. Estimating the reproducibility of psychological science. *Science.* 2015;349.
- Herrera-Perez D, Haslam A, Crain T, et al. Meta-research: A comprehensive review of randomized clinical trials in three medical journals reveals 396 medical reversals. *eLife.* 2019;8:e45183.
- Smith B, Haffty B, Wilson L, et al. The future of radiation oncology in the United States from 2010 to 2020: Will supply keep pace with demand? *J Clin Oncol.* 2010;28:5160-5165.
- Pan H, Haffty B, Falit B, et al. Supply and demand for radiation oncology in the United States: Updated projections for 2015 to 2025. *Int J Radiat Oncol Biol Phys.* 2016;96:493-500.
- Kahn J, Goodman C, Albert A. Top concerns of radiation oncology trainees in 2019: Job market, board examinations, and residency expansion. *Int J Radiat Oncol Biol Phys.* 2020;106:19-25.
- Tsai C, McBride S, Riaz N, et al. Reducing the radiation therapy dose prescription for elective treatment areas in human papillomavirus-associated oropharyngeal carcinoma being treated with primary chemoradiotherapy at Memorial Sloan Kettering Cancer Center. *Pract Radiat Oncol.* 2019;9:98-101.
- Amdur R. Throwing down the gauntlet regarding lowering the elective nodal irradiation dose in HPV-associated oropharyngeal cancer. *Pract Radiat Oncol.* 2019;9:63-64.