The article equates frequentist methods with simple null hypothesis testing without alternatives, thereby overlooking hypothesis testing methods that control both type I and II errors. The frequentist takes account of type II errors and the corresponding notion of power. If a test has high power to detect a meaningful effect size, then failing to detect a statistically significant difference is evidence against a meaningful effect. Therefore, a *P* value that is not small is informative.

The authors write that frequentist methods do not use background information, but this is to ignore the field of experimental design and all of the work that goes into specifying the test (eg, sample size, statistical power) and critically evaluating the connection between statistical and substantive results. An effect that corresponds to a clinically meaningful effect, or effect sizes well warranted from previous studies, would clearly influence the design.

Although their article engenders important discussion, these differences between frequentist and Bayesian methods may help readers understand why so many researchers around the world still prefer the frequentist approach.

Amit K. Chowdhry, MD, PhD
Department of Radiation Oncology
University of Rochester Medical Center
Rochester, New York

Deborah Mayo, PhD
Department of Philosophy
Virginia Tech
Blacksburg, Virginia

Stephanie L. Pugh, PhD
NRG Oncology Statistical and Data Management Center
American College of Radiology
Philadelphia, Pennsylvania

John Park, MD
Department of Radiation Oncology
Kansas City VA Medical Center
Kansas City, Missouri

Clifton David Fuller, MD, PhD
Department of Radiation Oncology
MD Anderson Cancer Center
Houston, Texas

John Kang, MD, PhD
Department of Radiation Oncology
University of Washington
Seattle, Washington

https://doi.org/10.1016/j.ijrobp.2022.08.034

References

1. Fornacon-Wood I, Mistry H, Johnson-Hart C, et al. Understanding the differences between Bayesian and frequentist statistics. *Int J Radiat Oncol Biol Phys* 2022;112:1076–1082.
2. Mayo DG. *Statistical Inference as Severe Testing: How to Get Beyond the Statistics Wars*. Cambridge, UK: Cambridge University Press; 2018.
3. Ryan EG, Brock K, Gates S, Slade D. Do we need to adjust for interim analyses in a Bayesian adaptive trial design? *BMC Med Res Methodol* 2020;20:150.
4. Jennison C, Turnbull BW. *Group Sequential Methods With Applications to Clinical Trials*. Boca Raton, FL: CRC Press; 1999.
5. Staley K, Park J. Comment on Mayo's "The statistics wars and intellectual conflicts of interest". *Conserv Biol* 2022;36:e13861.

In Reply to Chowdhry et al.

To the Editor:

We thank the authors for their response[1] to our "statistics for the people" article[2] that aimed to introduce perhaps unfamiliar readers to Bayesian statistics and some potential advantages of their use. We agree that frequentist statistics are a useful and widespread statistical analytical approach, and we are not aiming to revisit the frequentist versus Bayesian arguments that have been well articulated in the literature.[3-5] However, there are a couple of points we would like to make.

First, we acknowledge that the majority of phase 3 studies use frequentist designs, and this has the advantage of facilitating meta-analyses using established techniques. However, we would argue that the reason such frequentist designs are so prevalent is likely to have as much to do with convention (from funders/regulators as well as from researchers themselves), the relative exposure of the 2 approaches in educational materials, and the historic difficulties in calculating Bayesian posteriors as it does with the arguments the authors make.[6,7]

Second, although we agree with Chowdhry et al that there are many challenges associated with the estimation of prior probability distributions, we note that similar arguments apply to effect size estimation, which they cite as a strength of the Neyman-Pearson/null hypothesis significance testing approach (ie, the use of power calculations to limit the risk of type II errors).[8,9] We would also re-enforce the point we make in the article about the importance of testing the influence of the prior (represented as the divergent beliefs of the hypothetical radiation oncologist and surgeon in the communication by Chowdhry et al) in the analysis results. If the data are strong enough, the posterior distributions will be in close enough agreement to convince both parties. As we noted, it is also possible to undertake Bayesian analyses without prior information, using an uninformative prior, in which case the analysis is driven directly by the data, as for a frequentist calculation. As an aside, there is continued debate about the relative merits and deficiencies of the different frequentist approaches to significance testing, particularly around the widespread use of the hybrid Neyman-Pearson/null hypothesis significance testing approach.[10]

Disclosures: none.

There have, undoubtedly, been important practice changing studies delivered using frequentist approaches, but equally there is often erroneous interpretation of frequentist results.[11,12] Indeed, the American Statistical Society had to issue guidance on the misuse of frequentist significance testing.[13] We would argue that the often-counterintuitive nature of null hypothesis significance testing likely makes such interpretation errors inevitable. One of the principal strengths of the Bayesian approach we discuss in the article is that the researcher can directly ask the question they are interested in, that is, what is the probable effect size and uncertainty of an intervention compared with an alternative.

Finally, as we note in our original concluding paragraph, "both frequentist and Bayesian approaches are useful for data analysis as long as they are interpreted correctly" and that "however data are analyzed, it is of utmost importance to be transparent and to correctly interpret the results in a manner consistent with... limitations in how data were collected." That is, that the quality of the whole study design and its execution, including its assumptions and data collection approaches, is likely more important to the inferences one can make than the analytical approach itself.

Isabella Fornacon-Wood, MRes
Hitesh Mistry, PhD
Division of Cancer Sciences
University of Manchester
Manchester, United Kingdom

Gareth J. Price, PhD
Department of Medical Physics
The Christie Hospital NHS Foundation Trust
Manchester, UK

Corinne Faivre-Finn, PhD
Division of Cancer Sciences
University of Manchester
Manchester, United Kingdom
Department of Clinical Oncology
The Christie Hospital NHS Foundation Trust
Manchester, United Kingdom

James P.B. O'Connor, PhD
Division of Cancer Sciences
University of Manchester
Manchester, United Kingdom
Department of Diagnostic Radiology
The Christie Hospital NHS Foundation Trust
Manchester, United Kingdom

https://doi.org/10.1016/j.ijrobp.2022.08.034

References

1. Chowdhry AK, Mayo D, Pugh SL, Park J, Fuller CD, Kang J. In regard to Fornacon-Wood et al. [Letter to the Editor]. *Int J Radiat Oncol Biol Phys* 2022.
2. Fornacon-Wood I, Mistry H, Johnson-Hart C, Faivre-Finn C, O'Connor JPB, Price GJ. Understanding the differences between Bayesian and frequentist statistics. *Int J Radiat Oncol Biol Phys* 2022;112:1076–1082.
3. Dienes Z. Bayesian versus orthodox statistics: Which side are you on? *Perspect Psychol Sci* 2011;6:274–290.
4. Harrell FE. My journey from frequentist to Bayesian statistics. Available at: https://www.fharrell.com/post/journey/. Accessed July 25, 2022.
5. Zampieri FG, Casey JD, Shankar-Hari M, Harrell FE, Harhay MO. Using Bayesian methods to augment the interpretation of critical care trials. An overview of theory and example reanalysis of the alveolar recruitment for acute respiratory distress syndrome trial. *Am J Respir Crit Care Med* 2021;203:543–552.
6. Natanegara F, Neuenschwander B, Seaman Jr JW, et al. The current state of Bayesian methods in medical product development: Survey results and recommendations from the DIA Bayesian Scientific Working Group. *Pharmaceut Stat* 2013;13:3–12.
7. Clark J, Muhlemann N, Natanegara F, et al. Why are not there more Bayesian clinical trials? Perceived barriers and educational preferences among medical researchers involved in drug development. *Ther Innov Regul Sci* 2022;1:1–9.
8. Fern EF, Monroe KB. Effect-size estimates: Issues and problems in interpretation. *J Consum Res* 1996;23:89–105.
9. Marino MJ. The use and misuse of statistical methodologies in pharmacology research. *Biochem Pharmacol* 2014;87:78–92.
10. Perezgonzalez JD. Fisher, Neyman-Pearson or NHST? A tutorial for teaching data testing. *Front Psychol* 2015;6:1–11.
11. Hoekstra R, Morey RD, Rouder JN, Wagenmakers EJ. Robust misinterpretation of confidence intervals. *Psychon Bull Rev* 2014;21:1157–1164.
12. Greenland S, Senn SJ, Rothman KJ, et al. Statistical tests, *P* values, confidence intervals, and power: A guide to misinterpretations. *Eur J Epidemiol* 2016;31:337–350.
13. Wasserstein RL, Lazar NA. The ASA's statement on p-values: Context, process, and purpose. *Am Stat* 2016;70:129–133.

Meta-analysis of Low-Dose Irradiation for COVID-19

In Regard to Kolahdouzan et al.

To the Editor:

With great enthusiasm, we read the article by Kolahdouzan et al[1] recently published in the Red Journal. The authors performed a systematic review and meta-analysis to synthesize the evidence of low-dose whole lung irradiation for treatment of COVID-19 pneumonia. It is a significant breakthrough for this controversial subject,[2,3] because meta-analysis is an effective method to resolve clinical disputes and reach a final conclusion. However, several issues should be noted as potential pitfalls leading to biases, which are opposed by the Cochrane collaboration.

Disclosures: Zheng Li, Yue Hu, and Qiang Li developed the conception and design. All authors investigated, analyzed and synthesized the supporting data/references of the key viewpoints. All authors participated in discussion of the key viewpoints in order to form the ultimate consensus. Zheng Li and Yue Hu wrote the initial draft. Qiang Li revised the manuscript. All authors reviewed and approved the final manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (Grant No. 11875299).

Zheng Li and Yue Hu contributed equally to this work.