

El dilema del valor P en la era de la Medicina basada en la evidencia

Dr. Arturo Martí-Carvajal*

«Y en la receta, que es donde culmina, en efecto,
el interés de la Medicina,
es donde, por lo común, la Medicina malea o pierde
su dignidad de ciencia»
Don Gregorio Marañón.
«[...] en la naturaleza el modo normal de prosperar
los árboles es su asociación en un bosque»
Alfred N. Whitehead

El Papa Benedicto XVI en su obra «Las palabras, la palabra y la verdad» comienza con una pregunta indirecta: «Hasta qué punto no es moderno preguntar por la verdad...» (1). A manera de respuesta analítica — inmediatamente — cita un pasaje de la «Carta del diablo a su sobrino», del filósofo *C.S. Lewis* «que la única cuestión que con seguridad nunca se planteará es la relativa a la verdad de lo leído; en su lugar se pregunta acerca de las repercusiones y dependencias, del desarrollo del respectivo escritor, de la historia de sus influjos, y otras cuestiones análogas» (1). Lo cual pudiera ser interpretado como que la esencia — el *quid* — de la verdad no es importante para el lector, sino quién escribió, el idioma del

autor, el nombre de la revista, etcétera. Ello devela el pensamiento acrítico del lector. Este hilo conductor nos lleva a la obra del científico, filósofo y teólogo *Pável Florenski* (1882-1937), quien en su obra “La columna y fundamento de la verdad” expone: «[...] para el pensamiento teórico, «la Columna de la Verdad» es la certeza, *certitudo*» (2). Certeza y prueba no son sinónimos. Esta puede ser no cierta, mientras que aquella es algo inequívoco, claro y distinto, de lo cual no se puede dudar. La verdad es la esencia de la *episteme*, su contrario es la *doxa*. En la «Alegoría de la Línea» de Platón, la *doxa* es el signo del eslabón más primitivo del mundo sensible, mientras que la *episteme* es el extremo del mundo cognoscente. Este prolegómeno filosófico-epistemológico es necesario dado que la ciencia se hace en una matriz filosófica, Bunge *dixit* (3), y que existen al menos cinco razones para explicar por qué es tan difícil pensar como científico (4).

* MD, MSc, PhD. Invitado de Cortesía de la Academia Nacional de Medicina. Profesor titular (j) de la Universidad de Carabobo. Centro de Filosofía para la Investigación Stanislao Strba, Universidad de Carabobo ORCID iD: 0000-0001-8677-3351.
Email: arturo.marti.carvajal@gmail.com

La medicina es un arte y una ciencia; para quien esto escribe, es un arte científico, rodeado de incertidumbre. Recientemente, la revista médica *New England Journal of Medicine* publicó un editorial relacionado con las nuevas guías para el reporte estadístico en la citada revista (5). El editorial hace hincapié sobre las nuevas pautas que deben ser cumplidas en relación con el imperativo de reemplazar los valores de P con estimaciones de efectos o asociación e intervalos de confianza del 95 % cuando el protocolo ni el plan de análisis estadístico han especificado métodos utilizados para ajustar la multiplicidad. Tal editorial manifiesta la preocupación global de los editores de revistas y consultores estadísticos por el uso excesivo y la interpretación errónea de las pruebas de significación y los valores de P en la literatura médica. Hacer una deconstrucción del valor de P, desde su ontología, no es el quid en esta carta.

El valor P es una medida de evidencia estadística que aparece prácticamente en todos los trabajos de investigación médica. Su interpretación se hace extraordinariamente difícil porque no es parte de ningún sistema formal de inferencia estadística (6). Esta inquietud científica no es nueva (7,8), y las consecuencias para la epidemiología de un debate histórico ignorado han sido sustentadas con sólidos argumentos (7). En febrero de 2014, *George Cobb*, profesor emérito de Matemática y Estadística en *Mount Holyoke College*, formuló dos preguntas a un foro de discusión de *The American Statistical Association* (ASA): La primera, «¿Por qué tantas universidades y escuelas de posgrado enseñan $p = 0,05$?». La respuesta fue «porque eso sigue siendo lo que usan la comunidad científica y los editores de revistas». La segunda: «¿por qué tanta gente todavía usa $p = 0,05$?», siendo la respuesta: «Porque eso es lo que se les enseñó en la universidad o en la escuela de posgrado». El profesor *Cobb* expresó: «Lo enseñamos porque es lo que hacemos; lo hacemos porque es lo que nosotros enseñamos». Este episodio tautológico generó preocupación en la junta de la ASA (por sus siglas en inglés) (9). Desde entonces, se ha publicado una pléyade de reportes —a favor y en contra— sobre el uso, el reporte e interpretación del valor de la P (10-22).

Lo pernicioso de esta situación es que el valor de P se ha convertido es una especie de

elemento clave ¡incluso en estudios donde no hay hipótesis! Es frecuente ver estudios descriptivos —de prevalencia— reportando el valor de P, lo cual es un ejemplo del llamado «estudio SPIN» (23). Pareciera que no reportar el valor de P menosprecia o desacredita el manuscrito, de lo que se deduce —erróneamente— que el valor de P es la *Alétheia* del estudio. No obstante, la errónea interpretación del valor de P —que con hábito de dogmatismo se transforma en obstáculo epistemológico (24)—, genera conclusiones falsas (25,26) sobre las cuales se toman decisiones médicas. He aquí un componente ético y moral: ¡Primero, no hacer daño! El valor de P no debe ser interpretado como prueba de fortaleza del efecto de datos empíricos (27-29), tampoco es sinónimo de la expresión «clínicamente importante» (30). Un diseño inapropiado de un estudio afecta la credibilidad de sus autores, dado que una investigación clínica tiene un componente ético-moral, que no se limita a la obtención del consentimiento informado (31). Reducir el número de investigaciones dañinas, fútiles, vanas, falsas es un deber global; iniciativas y recomendaciones existen (32). La Academia Nacional de Medicina, a través de la Gaceta Médica de Caracas, tal como aquel jardín donde Platón enseñaba a sus seguidores —los academos— es la corporación idónea para tomar la iniciativa que se plantea en este manuscrito. No olvidemos que los académicos son los creadores de remedios para el pueblo, tal como lo expresa la etimología de Academia.

REFERENCIAS

1. Ratzinger J. Las palabras, la palabra y la verdad. <https://es.scribd.com/document/248244460/Las-Palabras-La-Palabra-y-La-v-Ratzinger-Joseph>. Acceso: 04 de septiembre de 2019.
2. Florenski P. III Carta segunda: La duda. En: La columna y fundamento de la verdad. Ensayo de teodicea ortodoxa en doce cartas. Ediciones Sígueme, Salamanca; 2010.p.47-74.
3. Bunge M. https://www.tendencias21.net/Mario-Bunge-La-ciencia-se-hace-en-una-matriz-filosofica_a33454.html. Acceso: 04 de septiembre de 2019.
4. Jarrett C. 5 Reasons It's So Hard To Think Like A Scientist. <https://digest.bps.org.uk/2017/06/20/5-reasons-its-so-hard-to-think-like-a-scientist/>. Acceso: 04 de septiembre de 2019.

5. Harrington D, D'Agostino RB Sr, Gatsonis C, Hogan JW, Hunter DJ, Normand ST, et al. New Guidelines for Statistical Reporting in the Journal. *N Engl J Med*. 2019;381(3):285-286.
6. Goodman S. A dirty dozen: Twelve P-value misconceptions. *Semin Hematol*. 2008;45(3):135-140.
7. Goodman SN. P values, hypothesis tests, and likelihood: Implications for epidemiology of a neglected historical debate. *Am J Epidemiol*. 1993;137(5):485-496.
8. Lang J, Rothman KJ, Cann CI. That confounded P-value. (Editorial). *Epidemiology*. 1998;9(1):7-8.
9. Wasserstein RL, Lazar NA. The ASA's statement on p-values: Context, process, and purpose. *Am Stat*. 2016;70:129-133.
10. Griffiths P, Needleman J. Statistical significance testing and P-values: ¿Defending the indefensible? A discussion paper and position statement. *Int J Nurs Stud*. 2019;99:103384.
11. Lytsy P. P in the right place: Revisiting the evidential value of P-values. *J Evid Based Med*. 2018;11(4):288-291.
12. Wasserstein RL, Schirm AL, Lazar NA. Moving to a world beyond "P<0.05." *Am Stat* 2019;73: Suppl 1:1-19.
13. Nuzzo R. Scientific method: statistical errors. *Nature*. 2014;506(7487):150-152.
14. O'Brien SF, Osmond L, Yi QL. How do I interpret a P value? *Transfusion*. 2015;55(12):2778-2782.
15. Kennedy-Shaffer L. Before P < 0.05 to Beyond P < 0.05: Using History to Contextualize P-Values and Significance Testing. *Am Stat*. 2019;73(Suppl 1):82-90.
16. Rose S, McGuire TG. Limitations of P-Values and R-squared for Stepwise Regression Building: A Fairness Demonstration in Health Policy Risk Adjustment. *Am Stat*. 2019;73(Suppl 1):152-156.
17. Schreiber JB. New paradigms for considering statistical significance: A way forward for health services research journals, their authors, and their readership. *Res Social Adm Pharm*. 2019. pii: S1551-7411(19)30508-X.
18. Johnson VE. Evidence from marginally significant t statistics. *Am Stat*. 2019;73(Suppl 1):129-134.
19. Brownstein NC, Louis TA, O'Hagan A, Pendergast J. The Role of Expert Judgment in Statistical Inference and Evidence-Based Decision-Making. *Am Stat*. 2019;73(1):56-68.
20. Demidenko E. The p-Value You Can't Buy. *Am Stat*. 2016;70(1):33-38.
21. van der Bles AM, van der Linden S, Freeman ALJ, Mitchell J, Galvao AB, Zaval L, et al. Communicating uncertainty about facts, numbers and science. *R Soc Open Sci*. 2019;6(5):181870.
22. Johnson VE. Revised standards for statistical evidence. *Proc Natl Acad Sci USA*. 2013;110(48):19313-7.
23. Boutron I, Dutton S, Ravaut P, Altman DG. Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. *JAMA*. 2010;303(20):2058-2064.
24. Bachelard G. La formación del espíritu científico. Buenos Aires: Ed. Argos; 1948.p.305.
25. Ioannidis JP. Why most published research findings are false. *PLoS Med*. 2005;2(8):e124.
26. Ioannidis JP. Why most discovered true associations are inflated. *Epidemiology*. 2008;19(5):640-648.
27. Martí Carvajal A. El número de pacientes a tratar: la cuantificación del beneficio clínico de una intervención en la era de la medicina basada en la evidencia. *Salus*. 2002;6(1). <http://servicio.bc.uc.edu.ve/fcs/vol6n1/3nume.pdf>. Acceso: 04 de septiembre de 2019.
28. Martí Carvajal A. ¿Existe asociación entre el virus de la hepatitis B y el carcinoma hepatocelular? Cómo demostrarlo mediante la gastroenterología basada en la evidencia. *G.E.N*. 2006;60:144-146.
29. Martí Carvajal A, Peña de Martí G, Muñoz Navarro S, Comunián Carrasco G, Martí Peña A. Significado de la razón de posibilidades (Odds ratio). *Gac Méd Caracas*. 2006;114(1):13-16.
30. Frank L, Basch E, Selby JV; Patient-Centered Outcomes Research Institute. The PCORI perspective on patient-centered outcomes research. *JAMA*. 2014;312(15):1513-1514.
31. Ellis LE, Kass NE. How are PCORI-funded researchers engaging patients in research and what are the ethical implications? *AJOB Empir Bioeth*. 2017;8(1):1-10.
32. Moher D, Glasziou P, Chalmers I, Nasser M, Bossuyt PMM, Korevaar DA, et al. Increasing value and reducing waste in biomedical research: who's listening? *Lancet*. 2016;387(10027):1573-1586.