

# Interpretando al *valor p*: ¿Significación estadística o significación clínica?

Martha C. Rodríguez-Vargas<sup>1</sup>

<sup>1</sup> Universidad Nacional Mayor de San Marcos, Facultad de Odontología, Lima, Perú

**Correspondencia:**

Martha C. Rodríguez-Vargas: mrodriguezv@unmsm.edu.pe  
Calle Germán Amézaga 375, Facultad de Odontología, Universidad Nacional Mayor de San Marcos. Lima-Perú.  
ORCID: 0000-0002-7986-8085

Recibido: 24/06/2024

Aceptado: 27/06/2024

Publicado: 30/06/2024

## Interpreting the *p value*: ¿Statistical significance or clinical significance?

---

Las investigaciones en ciencias de la salud, son predominantemente de enfoque cuantitativo, presentando un mayor alcance de los métodos estadísticos aplicados. Las estadísticas han ayudado a desarrollar e incrementar el conocimiento en las diferentes ramas de la salud <sup>1</sup>. Sin embargo, en muchas de estas investigaciones hay cierta preocupación por las conclusiones obtenidas a partir del procesamiento de los datos, las cuales se basan en el uso de las pruebas de hipótesis.

Una de las consecuencias detectadas, ha sido la importancia que se le ha otorgado al resultado en términos de un valor obtenido, olvidando el significado clínico real de la investigación. Desde que se describió por primera vez el *valor p*, ha sido y es usado como una de las medidas más importantes para la lectura de las pruebas estadísticas <sup>2</sup>, donde los datos se examinan en función de rechazar o no la hipótesis estadística “nula”. Este valor dicotomiza los hallazgos en significativos y no significativos <sup>3</sup>, por lo que, su uso es aún motivo de controversia. Esta práctica ha llevado a la creencia errónea de que el objetivo principal de las investigaciones debe ser obtener una significación estadística. Los autores se sienten presionados a mostrar resultados “estadísticamente significativos”, porque asumen que su estudio tendrá mayor oportunidad de ser aceptado, o que sus resultados son más relevantes, por lo que se inclinan a una presentación

selectiva, a la búsqueda de la significación a través del dragado de datos, o a la manipulación de los análisis estadísticos para producir resultados significativos, generando una distorsión de la literatura publicada, afectando la toma de decisiones <sup>4</sup>.

Para aclarar esa controversia sobre el uso y abuso de los *valores p*, la Asociación Estadounidense de Estadística (ASA) publicó una declaración, donde definió al *valor p* como la probabilidad, bajo un modelo estadístico especificado, de que un valor que representa alguna característica de los datos (por ejemplo la diferencia de las medias al comparar dos grupos), sea igual o más extremo que su valor observado <sup>5</sup>.

Para establecer una inferencia estadística, hay que tener en cuenta, que el modelo estadístico también incluye supuestos, como la distribución de los datos, la independencia y número de las observaciones; además de considerar, los tipos de medición, el diseño del estudio, la calidad de las medidas, entre otras <sup>2,5</sup>.

La interpretación de los resultados, es una de las partes más valiosas de un estudio, sacar conclusiones basadas sólo en la lectura de un *valor p* dicotómico, donde en la mayoría de las investigaciones  $p < 0,05$  se interpreta como “estadísticamente significativo” puede llevarnos a concluir erróneamente <sup>1,6</sup>. De igual manera, *valores p*

mayores que el valor preestablecido se informan cómo no significativos; y actualmente, valores cercanos al punto arbitrario, utiliza expresiones como “tendencia hacia la significación estadística”<sup>7</sup>.

El sobreinterpretar el *valor p* sin considerar al tamaño del efecto estimado, puede producir errores en la lectura. Cualquier efecto puede producir significación estadística si el tamaño de la muestra es lo suficientemente grande, incluso si la diferencia en sí misma no es clínicamente relevante, lo que conlleva a concluir que el *valor p* no es necesariamente un indicador de la replicabilidad de los resultados<sup>8-10</sup>. Interpretar la significación estadística como significación clínica, es un error común, el error se produce al interpretar que, cuanto menor sea el *valor p*, mayor es el efecto o más fuerte es la correlación. Sin tener en cuenta el tamaño de la muestra y la variabilidad<sup>2</sup>. De igual manera, un resultado no significativo, no debe interpretarse como que “no hay efecto” o que el efecto no es biológicamente importante<sup>7</sup>.

La diferencia entre relevancia estadística y clínica, radica en que la primera es el resultado de una fórmula estadística aplicada a los datos, que indica hasta que punto estos son incompatibles con el modelo estadístico<sup>5</sup>, mientras que la relevancia clínica va más allá de un cálculo y está determinada por el juicio clínico, es decir, cuando la variable ha tenido efectos reales y cuantificables, manifestándose por ejemplo, con la verificación de un mejor tratamiento o una mejor respuesta al mismo<sup>2</sup>.

Los resultados publicados, muestran a menudo que han encontrado una diferencia estadísticamente significativa, sin embargo, pocas veces se presenta el tamaño de dicha diferencia, por lo que, éstos resultados, podrían no representar el mismo significado en la práctica, es decir, ¿tendrán alguna relevancia clínica?

Por ello, se recomienda utilizar, además del *valor p*, el cálculo del tamaño del efecto, el cual mide la magnitud de la diferencia entre los grupos estudiados<sup>9</sup>. Estos efectos deben presentarse con sus respectivos intervalos de confianza y de esta manera proporcionar una mejor idea de la posible importancia clínica de los hallazgos estadísticamente significativos. Un *valor p* por sí solo, por pequeño que sea, puede ser insuficiente para explicar estos hallazgos, por lo que para permitir una interpretación más completa, honesta y útil se requieren de estos parámetros adicionales para su publicación<sup>11</sup>.

Tener en cuenta, que el *valor p*, seguirá desempeñando un papel fundamental en las investigaciones científicas basadas en pruebas de hipótesis, sin embargo, se debe mejorar su comprensión e interpretación, considerando sus debilidades, ventajas, y desventajas.

## Referencias bibliográficas

1. Sabetian PW, Domb BG. Editorial Commentary: The Power of Interpretation: Utilizing the P Value as a Spectrum, in Addition to Effect Size, Will Lead to Accurate Presentation of Results. *Arthroscopy*. 2022;38(4):1324-5. DOI: <https://doi.org/10.1016/j.arthro.2021.10.009>
2. Verploegh ISC, Lazar NA, Bartels RHMA, Volovici V. Evaluation of the Use of P Values in Neurosurgical Literature: from Statistical Significance to Clinical Irrelevance. *World Neurosurg*. 2022;161:280-283.e3. DOI: <https://doi.org/10.1016/j.wneu.2022.02.018>
3. Cote MP, Lubowitz JH, Brand JC, Rossi MJ. Misinterpretation of P Values and Statistical Power Creates a False Sense of Certainty: Statistical Significance, Lack of Significance, and the Uncertainty Challenge. *Arthroscopy*. 2021;37(4):1057-63. DOI: <https://doi.org/10.1016/j.arthro.2021.02.010>
4. To MS, Lu L, Tran M, Chong C. Preferential reporting of significant p-values in radiology journal abstracts. *Clin Radiol*. 2022;77(10):743-8. DOI: <https://doi.org/10.1016/j.crad.2022.05.025>
5. Wasserstein RL. The ASA's statement on p-values: context, process, and purpose. *Am Stat*. 2016;70(2):129-33. DOI: <https://doi.org/10.1080/00031305.2016.1154108>
6. Mansournia MA, Nazemipour M, Etminan M. P-value, compatibility, and S-value. *Glob Epidemiol*. 2022;4:100085. DOI: <https://doi.org/10.1016/j.gloepi.2022.100085>
7. El Teclé NE, Urquiaga JF, Griffin ST, Alexopoulos G, El Ahmadieh TY, Aoun SG, et al. Misinterpretations of Null Hypothesis Significance Testing Results Near the P-Value Threshold in the Neurosurgical Literature. *World Neurosurg*. 2022;159:e192-8. DOI: <https://doi.org/10.1016/j.wneu.2021.12.030>
8. Harris JD, Brand JC, Cote MP, Faucett SC, Dhawan A. Research Pearls: The Significance of Statistics and Perils of Pooling. Part 1: Clinical Versus Statistical Significance. *Arthroscopy*. 2017;33(6):1102-12. DOI: <https://doi.org/10.1016/j.arthro.2017.01.053>
9. Cortés J, Casals M, Langohr K, González JA. Importancia de la potencia y la hipótesis en el valor p. *Med Clínica*. 2016;146(4):178-81. DOI: <https://doi.org/10.1016/j.medcli.2015.10.011>
10. Chén OY, Bodelet JS, Saraiva RG, Phan H, Di J, Nagels G, et al. The roles, challenges, and merits of the p value. *Patterns*. 2023;4(12):100878. DOI: <https://doi.org/10.1016/j.patter.2023.100878>
11. Solla F, Tran A, Bertocelli D, Musoff C, Bertocelli CM. Why a P-Value is Not Enough. *Clin Spine Surg Spine Publ*. 2018;31(9):385-8. DOI: <https://doi.org/10.1097/BSD.0000000000000695>