

Importancia del tamaño del efecto para la comparación de dos o más grupos en investigaciones en salud pública

The Importance of Effect Size for Comparing Two or More Groups in Public Health Research

Tomás Caycho-Rodríguez^{1*} <https://orcid.org/0000-0002-5349-7570>

¹Universidad Privada del Norte. Trujillo, Perú.

*Autor para la correspondencia: tomas.caycho@upn.pe

Recibido: 15/11/2016

Aceptado: 06/11/2018

Señor Editor:

Muchos estudios en ciencias de la salud, con el objetivo de comparar las puntuaciones de una variable entre dos o más grupos (por ejemplo, sexo, nivel socioeconómico, nivel de instrucción, tipo de enfermedad, entre otras), recurren a la determinación de la existencia o no de diferencias estadísticamente significativas reportadas, fundamentalmente, a través de los valores del test de significancia de la hipótesis nula (NHST, por sus siglas en inglés), procedimiento que actualmente se encuentra en discusión.^(1,2) En este sentido, el NHST puede considerarse como un procedimiento insuficiente para la valoración de la importancia práctica de los resultados.

Entre las limitaciones del NHST se pueden mencionar las siguientes: considerar solo el error máximo permitido para rechazar una hipótesis nula (H_0) de ausencia de diferencias estadísticamente significativas y la alta sensibilidad respecto al tamaño de la muestra, donde muestras grandes tienden a estar asociadas también con el rechazo de la H_0 .^(3,4,5) Ante esta situación, se hace necesario el reporte de medidas del tamaño del efecto que permiten la cuantificación de la eficacia de diferentes niveles de una variable independiente. De esta manera, de acuerdo a lo sugerido por el Comité Internacional de Editores de Revistas Médicas,⁽⁶⁾ se busca complementar la información acerca de la existencia de diferencias estadísticamente significativas (obtenida a través del NHST) con información que indique la magnitud de tales diferencias entre los grupos.

Para ejemplificar la importancia del reporte del tamaño del efecto en investigaciones de salud pública, nos basaremos en los importantes resultados de *Arias y Muñoz del Carpio*,⁽⁷⁾ quienes reportan la existencia de elevados niveles de agotamiento emocional, baja realización personal y despersonalización (sintomatología característica del síndrome de desgaste profesional; *burnout* en inglés) en una muestra de enfermeras de la ciudad de Arequipa, Perú. Los resultados ponen en relieve un problema de suma importancia para la salud pública peruana y de muchos países latinoamericanos. Entre otros importantes resultados, los autores reportan la existencia de diferencias estadísticamente significativas en el puntaje total del síndrome de desgaste profesional ($p = 0,034$) y la baja realización personal ($p = 0,047$) de acuerdo al tiempo de servicio (menos de 5 años, entre 5 a 10 años, más de 10 años).

En su artículo *Arias y Muñoz del Carpio*⁽⁷⁾ utilizan el análisis de varianza como procedimiento estadístico para valorar las diferencias del puntaje del síndrome de desgaste profesional y de sus dimensiones según el tiempo de servicio. Por lo que, el objetivo de la presente carta es complementar los valiosos resultados de *Arias y Muñoz del Carpio* proponiendo el empleo y cálculo del coeficiente omega al cuadrado (ω^2) como una medida del tamaño del efecto para la comparación de más de dos grupos independientes en investigaciones de salud pública^(4,8) (menos de 5 años, entre 5 a 10 años, más de 10 años).

El coeficiente ω^2 indica los cambios de la variable dependiente, que pueden ser explicados por los diferentes niveles de la variable independiente.^(8,9) *Ferguson* y otros⁽²⁾ y *Fritz* y otros⁽⁹⁾ abordan el tema y muestran los elementos de la fórmula de ω^2 . Si bien existen otros coeficientes para la valoración del tamaño del efecto para comparar más de dos grupos independientes, como la eta al cuadrado (η^2) y la épsilon al cuadrado (ϵ^2), el coeficiente ω^2 presenta algunas bondades importantes como el no ser sensible a los cambios respecto al tamaño de la muestra, lo que sí sucede con el NHST. En este sentido, un valor de $\omega^2 = 0,04$, señala un tamaño mínimo del efecto; 0,25 hace referencia a un efecto moderado, finalmente, valores iguales o mayores a 0,64 señalan un efecto fuerte.^(8,9) Luego de realizar un reanálisis de la tabla 4 de *Arias y Muñoz del Carpio*,⁽⁷⁾ se puede plantear que los resultados tienen una interpretación diferente a la reportada por los autores.

Así, al comparar el puntaje total del síndrome de desgaste profesional ($\omega^2 = 0,14$) y la baja realización personal ($\omega^2 = 0,12$), de acuerdo al tiempo de servicio, ninguna de las diferencias son consideradas con un tamaño del efecto grande ($\omega^2 \geq 0,64$), valor que permitiría una adecuada interpretación de la importancia práctica de los resultados.⁽⁴⁾ Más bien, los hallazgos revelan la existencia de tamaños del efecto mínimo. Así, el no reportar medidas del tamaño del efecto podría generar errores en la interpretación, como la sobrevaloración de las diferencias encontradas.

Los resultados aquí presentados exponen la importancia del cálculo del tamaño del efecto en estudios que pueden servir de base para la implementación de políticas en salud pública. Lo cual permite tener una mayor y mejor evidencia y precisión, además de brindar un significado práctico a los resultados obtenidos en busca del establecimiento de acciones de mejora de la salud y bienestar de las enfermeras de la ciudad de Arequipa (Perú).

Referencias bibliográficas

1. Wasserstein RL, Lazar NA. The ASA's statement on p-values: context, process, and purpose. *Am Stat*, 2016;70(2):129-33. DOI: [10.1080/00031305.2016.1154108](https://doi.org/10.1080/00031305.2016.1154108)
2. Ferguson CJ. An effect size primer: a guide for clinicians and researchers. *Prof Psychol Res Pr*. 2009;40(5):532-8. DOI: [10.1037/14805-020](https://doi.org/10.1037/14805-020)
3. Sullivan G, Feinn R. Using effect size - or why the p value is not enough. *J Grad Med Educ*. 2012;4(3):279-82. DOI: [10.4300/JGME-D-12-00156.1](https://doi.org/10.4300/JGME-D-12-00156.1)
4. Dominguez-Lara S. Magnitud del efecto en comparaciones entre 2 o más grupos. *Rev Calid Asist*. 2016;32(2):121-2. DOI: [10.1016/j.cali.2016.04.002](https://doi.org/10.1016/j.cali.2016.04.002)
5. Caycho T, Ventura-León J, Castillo-Blanco R. Magnitud del efecto para la diferencia de dos grupos en ciencias de la salud. *An. Sist. Sanit. Navar*. 2016;39(3):459-61. DOI: [10.23938/ASSN.0242](https://doi.org/10.23938/ASSN.0242)
6. International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. *N Engl J Med*. 1997;(336):309-16. DOI: [10.1056/NEJM199701233360422](https://doi.org/10.1056/NEJM199701233360422)
7. Arias WL, Muñoz del Carpio A. Síndrome de burnout en personal de enfermería de Arequipa. *Rev Cub Sal Públ*. 2016 [acceso 08/11/2016];42(4):1-19. Disponible en: <http://www.revsaludpublica.sld.cu/index.php/spu/article/view/872>
8. Hinkle DE, Wiersma W, Jurs SG. *Applied statistics for the behavioral sciences*. 3rd Ed. Boston, MA: Houghton Mifflin; 1994.
9. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. *J Exp Psychol Gen*. 2012;141:2-18. DOI: [10.1037/a0024338](https://doi.org/10.1037/a0024338)

Conflicto de intereses

El autor declara que no existe conflicto de intereses.