

Magnitud del tamaño del efecto y su importancia en la investigación pediátrica

En relación con el artículo: Calidad de vida de los pacientes con inmunodeficiencias primarias de anticuerpos

Señor Editor:

En la actualidad se hace cada vez más común el empleo de medidas del tamaño del efecto al momento de informar los resultados de investigaciones en ciencias de la salud en general¹ (y pediatría en particular)^{2,3} debido a que brindan a la comunidad académica una interpretación más completa en comparación a la brindada por el test de significancia de la hipótesis nula (NHST, por sus siglas en inglés).⁴ A pesar de esto, en investigaciones en ciencias de la salud, que tienen como objetivo la comparación de una variable entre 2 o más grupos (por ejemplo, sexo, tipo de enfermedad, tipo de familia, etc.), es común recurrir a la interpretación del valor del NHST para concluir la existencia o no de diferencias estadísticamente significativas entre las medias de los grupos comparados, expresada en la probabilidad de que esta diferencia ocurra debido a un error en el proceso de muestreo.⁵

En años recientes, el empleo de NHST es discutido por varios investigadores^{6,7} quienes no lo consideran un procedimiento adecuado para la estimación del valor práctico de los resultados,

debido, básicamente, a dos características principales: su dependencia del tamaño o magnitud del efecto y su sensibilidad al tamaño muestral. Así, un estudio podría reportar resultados estadísticamente significativos teniendo un tamaño del efecto grande y una muestra pequeña o teniendo un tamaño del efecto pequeño y una muestra grande.⁵

Ante esta situación, la existencia de diferencias según los grupos de comparación, hace necesario también reportar los estadísticos del tamaño del efecto, como un procedimiento que permite cuantificar la magnitud de la diferencia entre grupos.⁸

Así, el objetivo de la presente carta es dar evidencia acerca de la importancia del reporte de medidas del tamaño y del efecto, para la comparación de dos grupos independientes en investigaciones en pediatría, mediante el cálculo de la *d* de Cohen. Para ello, se utilizarán, a modo de ejemplo, los importantes resultados de Muriel-Vizcaíno y colaboradores,⁹ publicados en esta revista, quienes brindan una contribución acerca de la existencia de diferencias significativas de la calidad de vida entre niños con

inmunodeficiencias primarias por defectos en la producción de anticuerpos y niños sanos en base al NHST, entre otros importantes resultados.

El cálculo mediante la d de Cohen, se realiza a partir de la siguiente fórmula: $d = (M_1 - M_2) / DE_{ponderada}$; donde, $(M_1 - M_2)$ son las diferencias de las medias de los grupos comparados y $DE_{ponderada}$ es la desviación estándar ponderada. Para la valoración de la d , Cohen⁸ señala que valores menores a .20, no presentan efecto; .21 a .49 tienen efecto pequeño; .50 a .70 evidencian efecto moderado; mientras que valores iguales o superiores a .80 presentan un efecto grande.

Luego de realizar un re-análisis del cuadro 1 de Muriel-Vizcaíno y sus colaboradores⁹ podemos señalar que, al comparar niños con inmunodeficiencias primarias y niños sanos, la calidad de vida total ($d_{Cohen} = .80$) tienen un tamaño del efecto grande; mientras que el funcionamiento físico ($d_{Cohen} = .65$) y funcionamiento psicosocial ($d_{Cohen} = .79$) tienen tamaños del efecto moderado. Desde el punto de vista de las medias, estos resultados podrían también interpretarse en el sentido de que las dos medias de los grupos, en relación a la calidad de vida total, funcionamiento físico y funcionamiento psicosocial, se diferencian en 0.80, 0.65 y 0.79 veces la desviación estándar promedio.⁴

Otra manera de interpretar el tamaño del efecto es a través de la proporción de la varianza explicada mediante el coeficiente general de determinación (R^2), que brinda una medida del porcentaje de la varianza total explicada por la variable independiente,⁴ la cual puede obtenerse a través de la transformación de la d de Cohen en base a la siguiente fórmula: $R^2 = d^2 / (d^2 + 4)$; donde d , es el valor obtenido de la d de Cohen. Realizando el cálculo, a partir de los valores de tamaño del efecto, se puede afirmar que las diferencias de acuerdo a padecer o no inmunodeficiencia primaria explican el 14, 9.6

y 13% de las diferencias de las puntuaciones de la calidad de vida total, funcionamiento físico y funcionamiento psicosocial respectivamente.

Un punto importante a tomar en cuenta al momento del cálculo del tamaño del efecto es que sus valores deben considerarse como una aproximación al valor poblacional verdadero y, por lo tanto, no se encuentra libre de error, el cual puede estimarse mediante el cálculo de los intervalos de confianza de la d .⁵ Los intervalos de confianza permiten conocer el rango de todos los valores posibles de la d en función a un determinado nivel de confianza (95, 99%, etc.), lo que permite tener información de si, a pesar de las condiciones que puedan afectar el valor de la d ; éste, aún tendría utilidad práctica. Si dentro del intervalo se encuentra incluido el cero, el resultado encontrado no sería estadísticamente significativo; caso contrario, la no presencia del cero, permite tener certeza de que nuestros hallazgos son estadísticamente significativos. Para el cálculo de los intervalos de confianza de la d nos basaremos en lo propuesto por Hedges y Olkin¹⁰ y Bologna.⁴ Primero, se obtiene el valor del error típico de la d ($\sigma[d]$):

$$\sigma[d] = \sqrt{\frac{(N_1 + N_2)}{N_1 * N_2} + 2 \left(\frac{d^2}{N_1 * N_2} \right)}$$

Donde N_1 y N_2 corresponden al tamaño de los grupos comparados. De esta manera el intervalo de confianza al 95% para la d estaría dado desde $d - 1.96 (\sigma[d])$ a $d + 1.96 (\sigma[d])$. En el caso de Muriel-Vizcaíno y sus colaboradores⁹ los intervalos de confianza al 95% de la d para calidad de vida total [0.317, 1.276], funcionamiento físico [0.171, 1.118] y funcionamiento psicosocial [0.312, 1.27] son estadísticamente significativos, teniendo los resultados utilidad práctica.

Como se mencionó líneas arriba, el objetivo de la presente carta, es dar a conocer lo importante

que es reportar medidas del tamaño del efecto. Los hallazgos revelan que las comparaciones se ubican en la categoría de magnitud moderada y grande con intervalos de confianza significativos, complementando y brindando mayor evidencia empírica acerca de la utilidad práctica de los valiosos resultados de Muriel-Vizcaíno y su equipo.⁹ Los reportes acerca de la magnitud del tamaño de efecto son importantes, pues permiten una mayor precisión a las conclusiones, siendo útiles también dentro de los estudios de metaanálisis, ya que al ser valores estandarizados ayudan para la sistematización de resultados de diversas investigaciones en pediatría y ciencias de la salud en general, importantes como base para la formulación de políticas públicas.

REFERENCIAS

1. Casado A, Prieto L, Alonso J. El tamaño del efecto de la diferencia entre dos medias: ¿estadísticamente significativo o clínicamente relevante? *Med Clin*. 1999; 112(15):584-588.
2. Domínguez-Lara S. Diferencias entre grupos y magnitud del efecto: un análisis complementario. *Arch Argent Pediatr*. 2016;114(4):e300-e309. Disponible en <http://www.sap.org.ar/docs/publicaciones/archivosarg/2016/v114n4a34.pdf>
3. Domínguez-Lara S. Diferencias en el estilo de vida según sexo y edad, análisis de la significancia práctica. *Rev Cubana Pediatr*. 2016; 88(3). Disponible en: <http://www.revpediatria.sld.cu/index.php/ped/article/view/185>
4. Bologna E. Estimación por intervalo del tamaño del efecto expresado como proporción de varianza explicada, *Evaluar*; 14:43-66. Disponible en <https://revistas.unc.edu.ar/index.php/revaluar/article/view/11521/11960>
5. Coe R, Merino C. Magnitud del efecto: una guía para investigadores y usuarios. *Revista de Psicología*. 2003; 21(1): 146-177. Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=993949>
6. Ferguson CJ. An effect size primer: a guide for clinicians and researchers. *Prof Psychol Res Pr*. 2009;40(5):532-538. Disponible en <http://psycnet.apa.org/journals/pro/40/5/532/>
7. Sullivan G, Feinn R. Using effect size - or why the p value is not enough. *J Grad Med Educ*. 2012;4(3):279-282. Disponible en <http://www.jgme.org/doi/full/10.4300/JGME-D-12-00156.1>
8. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. (2. da ed.). Hillsdale: Erlbaum, 1988.
9. Muriel-Vizcaíno R, Treviño-Garza G, Murata C, Staines-Boone A T, Yamazaki-Nakashimada M A, Espinosa-Padilla

S E, et al. Calidad de vida de los pacientes con inmunodeficiencias primarias de anticuerpos. *Acta Pediatr Mex*. 2016;37(1):17-25. Disponible en <http://ojs.actapediatrica.org.mx/index.php/APM/article/view/1126/885>

10. Hedges L, Olkin I. *Statistical methods for meta-analysis*. Nueva York: Academic Press, 1985.

Tomás Caycho

tomas.caycho@upn.pe

Universidad Privada del Norte Perú

Recibida: 17 de noviembre del 2016

Este artículo debe citarse como

Caycho T. Magnitud del tamaño del efecto y su importancia en la investigación pediátrica. En relación con el artículo: Calidad de vida de los pacientes con inmunodeficiencias primarias de anticuerpos. *Acta Pediatr Mex*. 2017;38(2):134-136.

DOI: <http://dx.doi.org/10.18233/APM38No2pp134-1361366>

Respuesta a la carta al Editor

Estimado Señor Editor:

Le agradecemos por darnos la oportunidad de contestar a la carta “Magnitud del tamaño del efecto y su importancia en la investigación pediátrica”¹ del Dr. Tomás Caycho. También quisiéramos agradecer por las observaciones y comentarios del Dr. Caycho, a través de los cuales se señala la importancia y necesidad de reportar tamaño de efecto en los resultados de análisis estadístico en las investigaciones y de nuestro trabajo: “Calidad de vida de los pacientes con inmunodeficiencias primarias de anticuerpos”² publicado en esta revista.

Estamos de acuerdo en que es de suma importancia estar conscientes del problema de depender del valor de *P* para interpretar y concluir los resultados de análisis estadísticos, y por ende, con respecto a los objetivos de una investigación. Las críticas sobre el uso del valor de *P* como un dispositivo para la toma de decisión en la interpretación de los resultados en forma dicotómica de “Significativo vs. No significativo” estadísticamente, desarrollado por Neyman y Pearson, surgieron inmediatamente después

de la aparición de esta propuesta, iniciado por el propio inventor del concepto original del valor de P : R. A. Fisher,³ seguido por R. W. Tyler⁴ y otros investigadores. Como señaló el Dr. Caycho, el problema de absolutizar el valor de P ha sido discutido actualmente, siendo uno de los autores que contribuyeron de manera más importante J. Cohen, para dar direccionalidad en la solución del problema y establecer las recomendaciones.^{5,6,7}

Aquí mencionamos que siguiendo a la sugerencia hemos calculado el tamaño de efecto de nuestros datos que presentamos en el artículo, y verificamos que todas las cifras se concordaron con los valores que obtuvo el Dr. Caycho. Al retomar los valores:

El tamaño de efecto del puntaje total de la escala del PedsQL observado entre los niños con IDP y niños sanos fueron 0.80 [IC95%: 0.32-1.28] en el puntaje total; 0.65 [IC95%: 0.17, 1.12] en la subescala del funcionamiento físico y 0.79 [IC95%: 0.31, 1.27] en la de funcionamiento psico-social. Aprovechando la oportunidad para complementar la información de nuestros resultados sobre el efecto del retraso de diagnóstico, en el grupo de niños con complicación pulmonar presentó tamaño de efecto muy grande, siendo el valor de $R^2=0.59$, el cual corresponde a d de Cohen de 2.5, mientras que en los niños con IDP sin esta complicación, la asociación no fue significativa ($p=0.296$) con $R^2=0.09$ que corresponde al tamaño de efecto pequeño d de Cohen ($d=0.2$).

Como se observa aquí, reportar los tamaños de efecto da mayor información y enriquece la interpretación de resultados. Los valores que hemos obtenido nos permiten argumentar que la diferencia de la calidad de vida de los niños con IDP es grande frente a los niños sanos en forma global, y al examinar por diferentes aspectos, el aspecto del funcionamiento psicológico puede

ser más afectado que el aspecto físico. Sin embargo, esta última interpretación no necesariamente está apoyada por los datos, ya que los intervalos de confianza de los tamaños de efecto están traslapados y puede no haber diferencia entre los dominios. Si pretendemos saber si el nivel afección en la calidad de vida se difiere entre los diferentes aspectos, es necesario aumentar la precisión de los intervalos de confianza, y para este fin, es necesario realizar un estudio con mayor número de sujetos. Por otro lado, este análisis puso en relieve la magnitud del efecto del retraso de diagnóstico en niños con complicación pulmonar.

A pesar de que el problema de reducir los resultados en valor de P ha sido señalado y criticado de manera convincente casi desde su aparición en la literatura médica (y de otros campos de ciencia),³⁻⁸ persiste el uso del valor de P como herramienta principal para sustentar “evidencias”, y nuestro estudio tampoco fue excepción.

Diferentes autores consideran que esta situación está causada por la deficiencia de la enseñanza de estadística en la carrera universitaria, tanto de pregrado como de posgrado.^{9,10} Hace dos años se publicó un artículo que señalaba el problema del valor de P en *Nature*¹¹ que de nuevo llamó mucho la atención de los investigadores y en este año la *American Association of Statistics* por primera vez hizo pública la recomendación sobre el uso de los diferentes indicadores del tamaño de efecto.¹² En este contexto, consideramos que son muy pertinentes las observaciones del Dr. Caycho sobre nuestro trabajo y actual falta de práctica del uso de tamaño de efecto.

Agradeciendo a la discusión que planteó el Dr. Caycho y junto con él quisiéramos proponer a la revista *Acta Pediátrica de México* y sus lectores que hagamos un esfuerzo para fomentar mejoramiento de la manera de reportar resultados de nuestras investigaciones.

REFERENCIAS

1. Caycho T. Magnitud del Tamaño del Efecto y su importancia en la Investigación Pediátrica. Acta Pediatr Mex. En prensa.
2. Muriel-Vizcaíno R, Treviño-Garza G, Murata C, Staines-Boone AT, Yamazaki-Nakashimada MA, Espinosa-Padilla SE, et al. Calidad de vida de los pacientes con inmunodeficiencias primarias de anticuerpos. Acta Pediatr Mex. 2016;37:17-25.
3. Fisher RA. The statistical methods in psychical research. Proc Soc Psych Res. 1929;39:189-192
4. Tyler RW. What is statistical significance? Educational Research Bulletin. 1931; 10; 115-142
5. Daniel L. Statistical Significance Testing. A Historical Overview. Res Sch. 1998; 5; 23-32
6. Kline RB. Beyond significance testing. Reforming data analysis methods in behavioral research. 1st ed. Washington, DC: APA; 2004:61-92
7. Perez-Gonzales JD. Fisher, Neyman-Pearson or NHST? A tutorial for teaching data testing. Front Psychol. 2015; doi:10.3389/fpsyg.2015.00223; <http://journal.frontiersin.org/article/10.3389/fpsyg.2015.00223/full> (Acceso: 19 de diciembre de 2016)
8. Cohen J. The Earth Is Round ($p < 0.05$). Am Psychol. 1994;49:997-1003
9. Schmidt F. Statistical significance testing and cumulative knowledge in psychology: Implications for training of researchers. Psychol Methods. 1996; 1(2):115-129
10. Sullivan G, Feinn R. Using effect size - or why the p value is not enough. J Grad Med Educ. 2012;4(3):279-282
11. Nuzzo R. Statistical errors. Nature. 2016; 506; <http://www.nature.com/news/scientific-method-statistical-errors-1.14700>. (Acceso: 19 de diciembre de 2016)
12. 1Ronald L. Wasserstein & Nicole A. Lazar (2016) The ASA's Statement on p-Values: Context, Process, and Purpose, The American Statistician, 70:2, 129-133, DOI:10.1080/00031305.2016.1154108

Rodolfo Muriel-Vizcaíno¹
 Gabriela Treviño-Garza²
 Chiharu Murata³

Aidé Tamara Staines-Boone⁴

Marco Antonio Yamazaki-Nakashimada⁵

Sara Elba Espinosa-Padilla⁶

Francisco Javier Espinosa-Rosales⁶

¹ Residente de tercer año, especialidad de pediatría médica, Instituto Nacional de Pediatría, México, D.F.

² Servicio de Inmunología, Hospital Regional Materno Infantil de Alta Especialidad de Monterrey.

³ Departamento de Metodología de la Investigación, Instituto Nacional de Pediatría, México, D.F.

⁴ Servicio de Inmunología, Unidad Médica de Alta Especialidad 25, Instituto Mexicano del Seguro Social, Monterrey Nuevo León.

⁵ Servicio de Inmunología, Instituto Nacional de Pediatría, SS, México D.F.

⁶ Unidad de Inmunología Clínica y Alergia, Instituto Nacional de Pediatría SS, México, D.F.

Correspondencia

Dr. Francisco Javier Espinosa Rosales
espinosa_francisco@yahoo.com.mx

Este artículo debe citarse como

Muriel-Vizcaíno R, Treviño-Garza G, Murata C, Staines-Boone AT, Yamazaki-Nakashimada MA, Espinosa-Padilla SE, Espinosa-Rosales FJ. Magnitud del tamaño del efecto y su importancia en la investigación pediátrica. En relación con el artículo: Calidad de vida de los pacientes con inmunodeficiencias primarias de anticuerpos. Acta Pediatr Mex. 2017;38(2):136-138.

DOI: <http://dx.doi.org/10.18233/APM38No2pp136-1381367>