

Algunas consideraciones sobre el uso de la prueba de Mc Nemar en las intervenciones educativas

Some considerations on the use of Mc Nemar test in the educational interventions

MsC. Irlán Amaro Guerra y Lic. Yolan Chía Gonzáles

Universidad de Ciencias Médicas, Santiago de Cuba, Cuba.

RESUMEN

En este artículo se exponen algunas consideraciones acerca del empleo de la prueba de Mc Nemar en las intervenciones educativas, así como de los elementos que se deben tener en cuenta para su cálculo y correcta aplicación, dado que el sesgo de confusión constituye uno de los errores que con mayor frecuencia se observan en las investigaciones biomédicas, específicamente en la presentación y el análisis de la información en el mencionado diseño de estudio.

Palabras clave: prueba de Mc Nemar, intervenciones educativas, investigaciones biomédicas.

ABSTRACT

In this work some considerations on the use of Mc Nemar test in the educational interventions, as well as on the elements that should be kept in mind for their calculation and correct application are exposed, since the confusion bias constitutes one of the mistakes that are most frequently observed in the biomedical investigations, specifically in the presentation and analysis of the information of the study design.

Key words: Mc Nemar test, educational intervention, biomedical investigations.

INTRODUCCIÓN

El sesgo de confusión constituye uno de los errores que con mayor frecuencia se observan en las investigaciones, cuya forma más común de controlarlos es en el análisis, específicamente el multivariado, como procedimiento por excelencia, aunque no es el único; sin embargo, algunos optan por su control en el diseño de la investigación.

Uno de los métodos que se emplea para su vigilancia, en esta etapa de la investigación, es el apareamiento. Con este tipo de diseño se intenta dar una mayor validez a las inferencias obtenidas, controlando o eliminando la influencia de variables extrañas, cuyo efecto ya es conocido o sospechado, y tratando de evitar que su ascendencia enmascare el verdadero efecto de la variable de interés.

De hecho, el concepto de diseño pareado se puede aplicar también al análisis de datos cuyo resultado es una categoría. En no pocas investigaciones resulta muy común tal tipo de diseño, con la particularidad de que en un porcentaje elevado de estas, el individuo

constituye su propio control, lo que se podría considerar como el apareamiento perfecto. En el ámbito de los profesionales de las ciencias médicas lo que más comúnmente se realiza son las llamadas intervenciones educativas, con las cuales se le aplica un cuestionario (generalmente acerca del nivel de conocimiento sobre determinada temática) a un grupo de individuos, se desarrolla la intervención y, posteriormente, se vuelve a emplear el mismo cuestionario, para finalmente decidir si los individuos obtuvieron conocimientos, u otro efecto, o no.

DESARROLLO

Generalmente lo que se busca con tal procedimiento es demostrar la eficacia de la intervención en la ganancia de conocimiento sobre la entidad clínica que haya sido tratada. Ante una técnica de este tipo surgen algunas interrogantes:

- ¿Es el nivel de conocimiento un factor teórico, y en la práctica demostrado, que influye o está asociado al efecto que se investiga? De no ser así, qué sentido tendría intervenir sobre este factor.
- ¿Qué tiempo debería mediar entre el fin de la intervención y la reevaluación de los individuos, para tener la certeza de que dicho arbitraje logró realmente su finalidad?
- Partiendo del supuesto de que lo anterior se lograra, ¿acaso no sería necesario evaluar luego el impacto que sobre la incidencia de la enfermedad o el factor de riesgo se alcanzó, con la modificación del supuesto factor (nivel de conocimiento)?

Estas y otras preguntas son inevitables ante tales planteamientos; sin embargo, a todo lo anteriormente expuesto, cabe añadirle las frecuentes incongruencias en la presentación y el análisis de la información en este tipo de diseño.

Para ejemplificar las consideraciones anteriores: debe suponerse que un grupo de investigadores estima que el desconocimiento acerca de los efectos nocivos para la salud del hábito de fumar pueda estar incidiendo en el incremento de este mal hábito y, por ende, en el alza de la morbilidad y mortalidad por cáncer de pulmón en cierta población. Diseñan una intervención educativa con la finalidad de incrementar el nivel de conocimientos sobre los efectos contra la salud del hábito de fumar, la cual es aplicada a 10 sujetos en 2 momentos, antes y después de la intervención. Los resultados se muestran a continuación (cuadro 1):

Cuadro 1. Resultados por sujetos antes y después de la intervención

Sujetos	Antes	Intervención	Después
1	Adecuado	Intervención	Inadecuado
2	Inadecuado	Intervención	Adecuado
3	Adecuado	Intervención	Inadecuado
4	Inadecuado	Intervención	Adecuado
5	Inadecuado	Intervención	Adecuado
6	Inadecuado	Intervención	Inadecuado
7	Inadecuado	Intervención	Adecuado
8	Inadecuado	Intervención	Adecuado
9	Inadecuado	Intervención	Adecuado
10	Adecuado	Intervención	Adecuado

Como pudo observarse, 3 sujetos presentaban un nivel de conocimiento adecuado antes de la intervención y 7, después de esta; por otra parte, otros 7 evidenciaban conocimientos inadecuados antes de la intervención y 3, después de realizada. La forma incorrecta de presentar estos datos es la siguiente (cuadro 2):

Cuadro 2. Forma incorrecta de presentar los datos en una tabla

Nivel de conocimiento	Antes	Después	Total
Adecuado	3	7	10
Inadecuado	7	3	10
Total	10	10	20

Llamó la atención, en primera instancia, que aparentemente la intervención se realizó en 20 individuos, con la consiguiente falacia que arrojaría una prueba de hipótesis, cuya potencia estará, en gran parte, determinada por el tamaño de la muestra y, en segundo lugar, una presentación de este tipo no expresa la medida de cambio intrasujetos en los 2 momentos, lo cual es, en esencia, lo que se pretende exponer. La presentación correcta sería (cuadro 3):

Cuadro 3. Forma correcta de presentar los datos en una tabla

		Después		
Nivel de conocimiento		Adecuado	Inadecuado	Total
Antes	Adecuado	1	2	3
	Inadecuado	6	1	7
Total		7	3	10

Pudo observarse que en el cuadro 3 solo se muestran los totales marginales que aparecen en el cuadro 4. También es importante apuntar que pareciera que la única forma de probar la eficacia de una intervención educativa fuera a través de la medición, en los mismos individuos, de la variable de respuesta en 2 momentos: antes y después de la intervención. Mucho más ilustrativo sería la selección de 2 grupos independientes y el empleo de intervenciones diferentes en cada uno de ellos, a fin de buscar probar la hipótesis de cuál sería más eficaz; en cuyo caso la dócima de elección no sería la prueba de Mc Nemar. Resulta oportuno agregar que algunos investigadores olvidan que en las investigaciones de este tipo, la población objeto de estudio no es finita, lo que implica que su realización en un lugar y el momento determinado, no significa que sus resultados solo sean extrapolables a dicho grupo poblacional y cometen, entonces, el error, con o sin intención, de repetir la misma intervención en diferentes lugares.

Todo lo anterior no necesariamente significa que se le esté quitando crédito a las investigaciones de esta naturaleza, pero sí se considera necesario advertir sobre el exceso de estudios de este tipo, muchos de ellos con escasa, en ocasiones casi nula, aplicabilidad.

Al respecto, la prueba de Mc Nemar es el test de hipótesis que por excelencia se emplea en los diseños antes señalados, siempre que la variable de respuesta sea dicotómica. No obstante, su uso no es exclusivo en este tipo de diseño (antes y después).

Este planteamiento es habitual en algunos estudios de casos y controles, en los que cada caso se empareja con un control, de acuerdo con un criterio determinado, y en el que se trata de valorar la frecuencia de la presencia de un factor de riesgo. Los resultados se pueden representar en una tabla de la siguiente forma (cuadro 4):

Cuadro 4. Forma correcta de presentar los datos en un estudio de casos y controles emparejados

		Controles		
		Factor presente	Factor ausente	
Casos	Factor presente	a	b	a+b
	Factor ausente	c	d	c+d
		a+c	b+d	n

En cada celda del cuadro anterior se refleja el número de parejas; así "a" es el número de parejas en las que el factor de riesgo está presente, tanto en los casos como en los controles, y "d" es el número de parejas en las que ni en los casos ni los controles se da el factor de riesgo. Es evidente que en esas 2 celdas hay concordancia entre lo observado en casos y lo observado en controles, en la pareja, y que, por tanto, no afectarán el resultado en cuanto a diferencias entre ambos grupos; solo los pares discrepantes "b" y "c" aportan información en ese sentido.

La proporción de controles que presentan el factor de riesgo es:

$$p_1 = \frac{a+c}{n}$$

Y la proporción de casos con el factor de riesgo:

$$p_2 = \frac{a+b}{n}$$

La diferencia de proporciones en cuanto a la presencia del factor de riesgo entre casos y controles es:

$$p_2 - p_1 = \frac{b-c}{n}$$

Donde, como fue anticipado, las cantidades a y d no intervienen. El error estándar de esa diferencia viene dado por:

$$e.s.(p_2 - p_1) = \frac{\sqrt{b+c}}{n}$$

El cuadrado del cociente entre la diferencia y su error estándar, se distribuye bajo la hipótesis de igualdad como una X² con 1 grado de libertad, y el contraste se conoce como prueba de Mc Nemar:

$$\chi^2 = \frac{(b - c)^2}{b + c}$$

Si se aplica la corrección de continuidad (recomendable sobre todo si el tamaño de muestra es pequeño o si hay celdas con frecuencias pequeñas), la fórmula anterior se modifica ligeramente:

$$\chi_c^2 = \frac{(|b - c| - 1)^2}{b + c}$$

Para estimar la oportunidad relativa (*odds ratio*) en este tipo de diseño se utiliza la fórmula:

$$OR = \frac{b}{c}$$

Donde, de nuevo, solo intervienen los pares con desacuerdo. El error estándar de esta oportunidad relativa se calcula como:

$$e. s. (OR) = OR \sqrt{\frac{1}{b} + \frac{1}{c}}$$

En una primera impresión puede sorprender la fórmula para el cálculo de la oportunidad relativa, pero su obtención es sencilla si se considera que en realidad cada pareja es un estrato con 2 elementos, y que no se deben combinar las tablas obtenidas en cada estrato al unirlos sin más. Si se aplica en el cálculo de la oportunidad relativa combinada, el método habitual conocido como de Mantel-Haenszel, se obtendrá la fórmula anterior.

Tal planteamiento se puede extender también al caso de una variable con más de 2 respuestas (prueba de Stuart-Maxwell), o también al caso de agrupaciones de más de 2 elementos por bloque.

REFERENCIAS BIBLIOGRÁFICAS

1. Siegel S, Castellana NJ. Estadística no paramétrica: aplicada a las ciencias de la conducta. 2 ed. Méjico: Editorial Trillas; 1995.
2. Martínez González MA, Sánchez Villegas A, Faulin FJ. Bioestadística amigable. 2 ed. Madrid: Editorial Díaz de Santos; 2006.
3. Sprent P, Smeeton NC. Applied nonparametric statistical methods. 3 ed. Boca Ratón, FL: Taylor & Francis; 2000. (Serie Chapman & Hall/CRC Texts in Statistical Science)
4. Gómez Gómez M, Danglot Banck C, Vega Franco L. Sinopsis de pruebas estadísticas no paramétricas. Cuándo usarlas. Rev Mex Pediatr. 2003; 70(2); 91-9.
5. Artilles Visbal L, Otero Iglesias J, Barrios Osuna I. Metodología de la investigación para las ciencias de la salud. La Habana: Editorial Ciencias Médicas; 2008 [citado 20 Ago 2009].

6. Arroniz Padilla S. Guía para la confección de un proyecto de investigación [citado 20 Ago 2009].
7. Molinero LM. Algunas consideraciones sobre el diseño del protocolo de investigación. Asociación de la Sociedad Española de Hipertensión; 2002 [citado 10 Sep 2009].
8. Ortiz Guerrero NA. La elaboración de los proyectos de investigación [citado 7 Sep 2009].

Recibido: 1 de julio de 2013.
Aprobado: 1 de julio de 2013.

Irlán Amaro Guerra. Universidad de Ciencias Médicas, avenida de las Américas, entre calles I y E, reparto Sueño, Santiago de Cuba, Cuba. Correo electrónico: irlan@medired.scu.sld.cu