# SECCIÓN METODOLÓGICA

*Psicológica (2006), 27*, 269-293.

## Comparación entre índices de tamaño del efecto para variables dicotomizadas en Meta-análisis

Mónica Teresa González Ramírez[*1] y Juan Botella[2]

[1] Universidad Autónoma de Nuevo León (México)
[2] Universidad Autónoma de Madrid (España)

El objetivo principal del presente trabajo es comparar, con datos provenientes de un meta-análisis real, el comportamiento de 7 índices para calcular el tamaño del efecto en variables dicotomizadas: $d_p$, $d_\phi$, $d_{asine}$, $d_{HH}$, $d_{Cox}$, $d_{probit}$, $d_{bis}$. Estos índices fueron estudiados por Sánchez-Meca, Marín-Martínez y Chacón-Moscoso (2003) a través de una simulación Monte Carlo. Considerando los resultados, se concluye que los 7 índices para calcular TE no son equivalentes; asimismo, se pone de manifiesto que el uso de uno u otro índice para calcular el tamaño del efecto puede tener consecuencias importantes en un meta-análisis, conduciendo incluso a conclusiones opuestas. Es por esto que se recomienda que para cálculos del tamaño del efecto para variables dependientes dicotomizadas se utilicen los índices $d_{Cox}$, $d_{probit}$ y $d_{HH}$ para poder comparar sus resultados e incrementar la validez de las conclusiones de un meta-análisis. Estas conclusiones coinciden con las del estudio de simulación de Sánchez-Meca et al. (2003).

El término Meta-análisis fue acuñado por Gene Glass (1976), quien establece la diferencia entre análisis primario, secundario y meta-análisis. El Meta-análisis del griego meta, "después de" y análisis, "descripción o interpretación", consiste en el análisis estadístico de una gran colección de resultados extraídos de trabajos individuales, con el propósito de integrar los hallazgos obtenidos. En otras palabras, el Meta-análisis, corresponde a una metodología de investigación que permite realizar una revisión sistemática de aspectos cualitativos y cuantitativos de estudios individuales (sus características y su metodología) y sus resultados numéricos con el

---

[*] **Agradecimientos**: Los autores agradecen al Dr. Julio Sánchez-Meca por sus comentarios y sugerencias para la realización de este trabajo. **Correspondencia**: Mónica Teresa González Ramírez. Facultad de Psicología de la Universidad Autónoma de Nuevo León, Mutualismo 110, Col. Mitras Centro, 64460 Monterrey, N.L., México. Telf. (81)8333-7859, Fax: (81)8348-3781. Correo electrónico: monygzz77@yahoo.com

afán de resumir la información científica existente en relación a un tema en particular. Los meta-análisis son estudios de tipo correlacional, pues en éstos los investigadores recolectan los estudios pertinentes, los estudian y comparan la evidencia que aportan con la de otros similares, sin cambiar ni alterar el curso de los acontecimientos (Cooper, 1998).

El Meta-análisis, de acuerdo con Botella y Gambara (2002), se refiere "al análisis estadístico de un conjunto de resultados estadísticos obtenido de una variedad de estudios relativamente homogéneos, con el objeto de integrar resultados" (p.23).

Entre los objetivos que se suelen asumir para realizar un Meta-Análisis se encuentran el de sintetizar información para determinar el tamaño del efecto de una intervención; ya que no es inusual detectar algunos estudios que muestran un efecto de cierto factor, mientras que otros no reportan efecto alguno. Por otro lado, a veces todos los estudios muestran cierto efecto, pero en algunos el efecto es marginal mientras que en otros es de gran importancia; surge la pregunta entonces respecto de cuál será el verdadero tamaño del efecto [TE], dado que todos ellos son estimaciones (Manterola, Riedemann y Vial, 2001).

Se puede decir que el tamaño del efecto es el grado en que un fenómeno está presente en la población (Cohen, 1988). En un meta-análisis, estimar el tamaño del efecto es un paso fundamental, ya que es aquí donde se transforman en una escala numérica los datos provenientes de los estudios que se van a integrar (Lipsey y Wilson, 2001).

Los índices de TE más empleados en MA, en psicología (Botella y Gambara, 2002) son: Diferencia de medias estandarizada, Correlación de Pearson y Razón de Ventajas (*odds ratio*).

La elección de uno u otro índice dependerá de la variable de interés; así, si se trata de una variable de tipo continuo lo mas común es utilizar la diferencia de medias estandarizada. Si la variable es dicotómica lo mas común es utilizar la razón de ventaja u *odds ratio* (Manterola et al., 2001). Sin embargo, existen otros índices para calcular el tamaño del efecto en variables dicotomizadas (Sánchez-Meca, Marín-Martínez y Chacón-Moscoso, 2003).

Sánchez-Meca et al. (2003) analizan los siguientes 7 índices de Tamaño del Efecto para Tablas 2x2, cada uno de estos índices será descrito posteriormente: (1) Diferencia estandarizada de proporciones, $d_p$; (2) Transformación del coeficiente phi a diferencia de medias estandarizada, $d_\phi$; (3) Transformación arcoseno, $d_{asine}$; (4) Índice basado en la razón de ventajas, asumiendo distribución logística y varianza homogénea, $d_{HH}$; (5) Índice propuesto por Cox, $d_{Cox}$; (6) Transformación probit, $d_{probit}$; y (7) Coeficiente biserial-phi, $d_{bis}$.

En investigación social y conductual es muy común encontrar estudios que usan variables continuas que han sido dicotomizadas (Sánchez-Meca et al., 2003). Es por esto que se considera importante continuar con la investigación referente a los índices para calcular el tamaño del efecto en este tipo de variables. Así, el presente trabajo pretende establecer guías sobre la elección de alguno de los estimadores de tamaño del efecto para variables dicotomizadas, con el propósito de que los usuarios que inician su aproximación a dichos índices puedan tomar decisiones sobre cual o cuales elegir.

En un mismo meta-análisis puede resultar problemático tratar de integrar resultados de estudios donde se empleó la variable original (continua) y estudios donde ésta ya fue dicotomizada; en estos casos, un índice de tamaño del efecto en la misma métrica *d* para variables dicotómicas permitiría integrar en el meta-análisis los resultados de todos los estudios. Igualmente, hay resultados en los que es más apropiado analizar los datos usando variables dicotomizadas. Sin embargo, hay dos preguntas muy relevantes: ¿todas las opciones para calcular el tamaño del efecto para variables dicotomizadas son equivalentes? y ¿qué consecuencias tiene para el meta-análisis el uso de uno u otro de los índices?

Estas son las preguntas que se intenta responder a través de este trabajo, tomando como base las aportaciones realizadas por Sánchez-Meca et al. (2003) quienes estudian, a través de simulación Monte Carlo, los 7 índices del tamaño del efecto mencionados previamente para variables dicotomizadas.

Por tanto, el objetivo principal del presente estudio es comparar, con datos reales provenientes de un meta-análisis concreto, el comportamiento de los 7 índices explorados por Sánchez-Meca et al. (2003) en su estudio de simulación. Empleando los resultados de un meta-análisis publicado, es posible ilustrar las consecuencias que tendría en un caso concreto la elección entre estos índices alternativos. Los análisis se realizan con datos del meta-análisis de Sánchez-Meca, Marín-Martínez, Olivares y Rosa (1999).

### *Descripción de los índices para calcular tamaño del efecto en variables dicotomizadas*

Las fórmulas para estimar los diferentes índices presentados, han sido obtenidas de Sánchez-Meca et al. (2003). Todos los índices para calcular el tamaño del efecto que se presentan, son aplicables a diseños de investigación donde se comparan dos grupos (habitualmente grupo control y grupo experimental) y la variable dependiente ha sido dicotomizada, lo que implica que en el estudio original a analizar en el meta-análisis solo se aporta información sobre las tablas de contingencia 2 x 2.

*a) Diferencia estandarizada de proporciones ($d_p$)*

El primero de los 7 índices es la diferencia estandarizada de proporciones ($d_p$), considerando un diseño en el que se cuenta con 2 grupos (uno experimental y uno control), donde se define el éxito de determinada intervención con una puntuación específica en la variable dependiente, la diferencia estandarizada de proporciones se refiere a la diferencia entre la proporción de éxitos o fracasos en los grupos experimental y de control ($p_E$ y $p_C$ respectivamente), dividido por la desviación estándar entre grupos (S'). Las fórmulas necesarias para obtener $d_p$ son:

$$d_p = \frac{p_E - p_C}{S'} \quad (1)$$

$$S' = \sqrt{\frac{(n_E - 1)p_E(1 - p_E) + (n_C - 1)p_C(1 - p_C)}{n_E + n_C - 2}} \quad (2)$$

Para calcular la varianza del índice $d_p$ se emplea la fórmula:

$$S^2_{dp} = \frac{n_E + n_C}{n_E n_C} + \frac{d_P^2}{2(n_E + n_C)} \quad (3)$$

donde: $p_E$ es la proporción de éxitos del grupo experimental; $p_C$ es la proporción de éxitos del grupo control; $n_E$ es el tamaño de muestra del grupo experimental; $n_C$ es el tamaño de muestra del grupo control.

Fleiss (1994) considera que la única ventaja de la diferencia de proporciones es su simplicidad. Sin embargo, en este trabajo se utiliza la diferencia estandarizada de proporciones, donde se incluye en la fórmula la desviación estándar entre grupos (fórmula 1).

Sánchez-Meca et al. (2003) indican que no se han realizado estudios sobre factores que afecten a $d_p$. En la simulación que realizaron encontraron que $d_p$ subestima el efecto en la población, en muestras donde los marginales no son proporcionales; asimismo, $S^2_{dp}$ subestima su variabilidad empírica.

*b) Transformación del coeficiente phi, a diferencia de medias estandarizada ($d_\phi$)*

El segundo de los índices es el coeficiente phi, transformado a la métrica de la diferencia de medias estandarizada ($d_\phi$), las fórmulas necesarias para calcularlo son:

$$\phi = \frac{O_{1E}O_{2C} - O_{2E}O_{1C}}{\sqrt{n_E n_C m_1 m_2}} \tag{4}$$

$$d_\phi = \frac{\phi}{\sqrt{1-\phi^2}}\sqrt{\frac{df(n_E + n_C)}{n_E n_C}} \tag{5}$$

La fórmula para calcular la varianza de $d_\phi$ es:

$$S^2_{d\phi} = \frac{n_E + n_C}{n_E n_C (1-\phi^2)} \tag{6}$$

donde: $O_{1E}$ y $O_{1C}$, son, respectivamente las frecuencias de éxito de los grupos experimental y control, en una tabla 2x2; $O_{2E}$, $O_{2C}$, son, respectivamente las frecuencias de fracaso de los grupos experimental y control, en una tabla 2x2; $df = n_E + n_C - 2$ (grados de libertad); $m_1 = O_{1C} + O_{1E}$; $m_2 = O_{2C} + O_{2E}$; el resto de los términos fueron definidos anteriormente.

Haddock, Rindskopf y Shadish (1998) señalan que $\phi$ subestima el tamaño del efecto y consideran que cuando las distribuciones marginales no son equivalentes en una tabla 2x2, $\phi$ no provee una medida satisfactoria de tamaño del efecto. Al igual que con $d_p$, Sánchez-Meca et al. (2003) indican que no se han realizado estudios sobre factores que afecten a $d_\phi$ y en la simulación encontraron que $d_\phi$ y la fórmula ofrecida para su varianza subestiman, respectivamente, el efecto en la población y su varianza real.

*c) Transformación arcoseno ($d_{asine}$)*

El tercer índice que se describe se basa en la transformación arcoseno, que es una de las transformaciones de datos utilizadas cuando se requiere corregir el incumplimiento de los supuestos estadísticos subyacentes a las técnicas multivariantes (principalmente normalidad), o bien para mejorar la relación entre variables (Hair et al., 1999).

Existen diferentes opciones para la transformación de datos, asociándose algunas de manera especifica al tipo de datos, así, Hair et al.

(1999) señalan que las proporciones se transforman mejor por la transformación del arcoseno ( $X_{nueva} = 2ar\cos eno\sqrt{X_{antigua}}$ ).

Cohen (1988) propuso utilizar esta transformación para obtener un tamaño del efecto en una métrica d ($d_{asine}$), aplicando la fórmula:

$$d_{asine} = 2\text{arcoseno}\sqrt{p_E} - 2\text{arcoseno}\sqrt{p_C} \quad (7)$$

La varianza de $d_{asine}$ se estima con la fórmula:

$$S^2\, d_{asine} = \frac{1}{n_E} + \frac{1}{n_C} \quad (8)$$

De acuerdo a los planteamientos de Lipsey y Wilson (2001) lo esperado es que $d_{asine}$ subestime el tamaño del efecto poblacional, al menos en poblaciones con distribuciones muy asimétricas. Como en los índices anteriores, Sánchez-Meca et al. (2003) encontraron que éste subestima el tamaño del efecto poblacional y la fórmula 8 su variabilidad.

*d) Índice basado en la razón de ventajas, asumiendo distribución logística y varianza homogénea ($d_{HH}$)*

El cuarto índice esta basado en la razón de ventajas, asumiendo distribución logística y varianzas homogénea. Se trata de $d_{HH}$, que se define como:

$$d_{HH} = \log(or)\frac{\sqrt{3}}{\pi} \quad (9)$$

La razón de ventajas tiene inconvenientes en su interpretación. El valor 1 indica que no hay relación; valores entre 0 y 1 indican una relación negativa y valores mayores a 1 indican una relación positiva. Así, una razón de ventajas de .5 seria equivalente en fuerza de relación a una razón de ventajas de 2 (la inversa de 2 es .5), pero en dirección opuesta; para compensar este inconveniente se utiliza el logaritmo natural de la razón de ventajas [Log(or)] (Lipsey y Wilson, 2001).

La distribución del Log(or) se aproxima a una distribución normal, con una media de 0 y una desviación estándar de 1.83; así, un valor negativo representa una relación negativa y un valor positivo refleja una relación positiva (Lipsey y Wilson, 2001).

El índice $d_{HH}$ para calcular el tamaño del efecto toma en cuenta Log(or) (fórmula 9), para lo cual es necesario el cálculo de la razón de ventajas, que se obtiene de la siguiente forma:

$$OR = {p_T(1-p_C)} \Big/ {p_C(1-p_T)} \quad (10)$$

La varianza de $d_{HH}$ se obtiene con la fórmula:

$$S^2 d_{HH} = \frac{3}{\pi^2}\left[\frac{1}{O_{1E}} + \frac{1}{O_{2E}} + \frac{1}{O_{1C}} + \frac{1}{O_{2C}}\right] \qquad (11)$$

Sánchez-Meca et al. (2003) encontraron que $d_{HH}$ subestima el efecto en la población y recomiendan estudiar más en profundidad el comportamiento de este índice cuando no se puede asumir distribución logística.

*e) Índice propuesto por Cox ($d_{Cox}$)*

El quinto índice fue propuesto por Cox (1970, en Sánchez-Meca et al., 2003), se trata de $d_{Cox}$ y se obtiene dividiendo el logaritmo de razón de ventajas por la constante 1.65, lo cual provee valores para el TE en una escala normal estandarizada (Haddock et al.,1998):

$$d_{Cox} = \frac{L_{or}}{1.65} \qquad (12)$$

La formula para la varianza de $d_{Cox}$ es:

$$S^2\ d_{Cox} = 0.367\left[\frac{1}{O_{1E}} + \frac{1}{O_{2E}} + \frac{1}{O_{1C}} + \frac{1}{O_{2C}}\right] \qquad (13)$$

El índice $d_{Cox}$ fue el que presentó mejor ejecución en el estudio de simulación de Sánchez-Meca et al. (2003), con una leve sobrestimación del tamaño del efecto poblacional, y $S^2\ d_{Cox}$ una pequeña subestimación de la variabilidad empírica.

El sexto y el séptimo índice están basados en el supuesto de la distribución normal, se trata de la transformación probit ($d_{probit}$) y el coeficiente biserial-phi ($d_{bis}$).

*f) Transformación probit ($d_{probit}$)*

La transformación probit ($d_{probit}$) se define como:

$$d_{probit} = (Z_E - Z_C) \qquad (14)$$

siendo $Z_E$ y $Z_C$ las inversas de los valores de función de distribución $p_E$ y $p_C$, respectivamente, en una distribución normal estandarizada.

La varianza de $d_{probit}$ se estima con:

$$S^2\ d_{probit} = \left[\frac{2\pi p_E(1-p_E)e^{z_E^2}}{n_E} + \frac{2\pi p_C(1-p_C)e^{z_C^2}}{n_C}\right] \qquad (15)$$

Asumiendo la distribución normal, $d_{probit}$ es un estimador insesgado del tamaño del efecto poblacional (Sánchez et al., 2003). Lipsey y Wilson (2001) indican que en teoría, probit es un excelente estimador del tamaño del efecto, sin embargo, en distribuciones desviadas de la normalidad sobreestima el TE, excepto en distribuciones asimétricas donde el punto de corte es el valor mas alto de la distribución.

Sánchez et al. (2003) encontraron que después de $d_{cox}$, $d_{probit}$ presentaba mejor ejecución que el resto de los índices, con una leve sobreestimación del tamaño del efecto poblacional; $S^2 d_{probit}$ presentaba una pequeña subestimación de la varianza.

*g) Coeficiente biserial-phi ($d_{bis}$)*

El séptimo índice consiste en calcular el coeficiente de correlación biserial-phi ($\phi_{bis}$) y transformarlo a una métrica d ($d_{bis}$), las fórmulas necesarias para calcular este índice son:

$$\phi_{bis} = \frac{\sqrt{p'q'}}{y'}\phi \tag{16}$$

$$d_{bis} = \frac{\phi_{bis}}{\sqrt{1-\phi_{bis}^2}}\sqrt{\frac{df(n_E+n_C)}{n_E n_C}} \tag{17}$$

Para estimar la varianza de $d_{bis}$, la fórmula es:

$$S^2 d_{bis} = \frac{p'q'(1-\phi^2)(n_E+n_C)}{y'^2 n_E n_C (1-\phi_{bis}^2)^3} \tag{18}$$

donde p' es la proporción de éxitos total de la tabla 2x2; q' se obtiene de: q'=1-p'; y' es la altura de la ordenada de una curva estándar, correspondiente al valor Z con percentil p' (Longley-Cook, 1981).

Como se mencionó anteriormente Haddock et al. (1998) señalan que cuando las distribuciones marginales no son equivalentes en una tabla 2x2, $\phi$ no provee una medida satisfactoria de tamaño del efecto; esto ha de ser considerado para $d_{bis}$ ya que implica el calculo de $\phi$. Sánchez-Meca et al. (2003) señalan que bajo distribuciones normales $d_{bis}$ tendrá una buena ejecución, en el estudio de simulación encontraron que sobreestima el tamaño del efecto poblacional, mientras que $S^2 d_{bis}$ es una buena estimación de la variabilidad empírica cuando el tamaño del efecto poblacional es pequeño.

### *Ejecución de los diferentes índices*

Sánchez-Meca et al. (2003) indican que los 7 índices pertenecen a la misma métrica, por lo que sus propiedades pueden ser comparadas. Sin embargo, hay que hacer algunas consideraciones previas que se resumen en la tabla 1.

A medida que se incrementa el tamaño del efecto poblacional se incrementan tanto la sobreestimación de $d_{Cox}$, $d_{probit}$ y $d_{bis}$ como la subestimación de $d_{HH}$, $d_p$, $d_\phi$ y $d_{asine}$.

En la misma simulación de Sánchez-Meca et al. (2003), $d_{bis}$ presentó una pobre ejecución en la condición con una gran distancia entre el punto de corte y el parámetro δ, y cuando el tamaño muestral era diferente en los grupos control y experimental. Para el resto de los estimadores del tamaño del efecto, no se detectaron influencias del tamaño de la muestra, el punto de corte o la relación de los tamaños muestrales de los grupos experimental y de control. En síntesis, los índices más recomendados cuando se puede asumir la normalidad son $d_{Cox}$ y $d_{probit}$.

Respecto a la varianza de cada uno de los índices, Sánchez-Meca et al. (2003), encontraron que los 7 índices mostraban mayor variabilidad que la diferencia de medias estandarizada; en orden ascendente tendríamos: $d_p$, $d_\phi$, $d_{asine}$, $d_{HH}$, $d_{Cox}$, $d_{probit}$, $d_{bis}$.

Entre los factores que afectan a la variabilidad de los estimadores se encuentran los siguientes (Sánchez-Meca y cols, 2003): 1) hay una relación directa entre la magnitud de los parámetros δ y los valores empíricos de las varianzas de todos los índices; 2) a mayor tamaño de muestra, menor varianza empírica, y 3) diferencias en el tamaño de los grupos experimental y control (la variabilidad se incrementa cuando aumenta la diferencia). Específicamente en $d_{bis}$ se encontró que cuando el grupo control era más pequeño, la variabilidad se incrementaba (en valores altos de δ).

### *Descripción del meta-análisis base*

Como se mencionó anteriormente, con el objetivo de comparar el comportamiento de los 7 estimadores del tamaño del efecto para variables dicotomizadas, se realizaron análisis basados en datos reales provenientes del meta-análisis realizado por Sánchez-Meca et al. (1999), cuyos resultados proporcionan información sobre la eficacia de las intervenciones conductuales en el tratamiento de la adicción al tabaco, en España.

Este meta-análisis base fue seleccionado después de una amplia búsqueda, por trabajar con una variable dicotómica de origen (abstinencia/no abstinencia) y por contar con datos suficientes en sus apéndices para calcular los 7 estimadores del tamaño del efecto para variables dicotomizadas.

**Tabla 1. Información sobre los diferentes estimadores del tamaño del efecto para variables dicotomizadas.**

| Índice | Información previa a la simulación realizada por Sánchez-Meca et al. (2003) | Resultado bajo el supuesto de normalidad en la simulación realizada por Sánchez et al., (2003) | |
|---|---|---|---|
| | Respecto a su ejecución | Estimación del efecto | Estimación de la varianza |
| $d_p$ | No ha sido explorado como lo afectan la magnitud del efecto ni otros factores | (subestimación de δ, y ésta incrementa con el valor de δ) | La subestima (varianza mas pequeña) |
| $d_\phi$ | No ha sido explorado como lo afectan la magnitud del efecto ni otros factores | (subestimación de δ, y ésta incrementa con el valor de δ) | La subestima (varianza mas pequeña) |
| $d_{asine}$ | Lo esperado es que subestime el TE poblacional, al menos en poblaciones con distribuciones muy asimétricas | (subestimación de δ, y ésta incrementa con el valor de δ) | La subestima (varianza mas pequeña) |
| $d_{HH}$ | Está basado en distribución logística. Falta estudiar su ejecución cuando no se permite asumirla | (subestimación de δ, y ésta incrementa con el valor de δ) | La subestima un poco |
| $d_{Cox}$ | - - | **Mejor ejecución de los 7** (muy leve sobreestimación de δ) | La subestima un poco |
| $d_{probit}$ | Bajo distribución normal ofrece los mejores resultados, ya que la asume. En distribuciones desviadas de la normalidad, sobreestima el TE, excepto en distribuciones asimétricas donde el punto de corte es el valor mas alto de la distribución. | Mejor ejecución después de $d_{Cox}$ (leve sobreestimación de δ) | La subestima un poco |
| $d_{bis}$ | Bajo distribución normal ofrece los mejores resultados, ya que la asume | (sobreestimación de δ y esta incrementa con el valor de δ) | Buena estimación para δ pequeños. Se recomienda usarlo cuando se puede asumir que δ es pequeño o medio |

El meta-análisis base incluyó 13 estudios, todos con grupo experimental y grupo control. Los autores obtuvieron las tasas de abstinencia de los grupos experimental y control ($p_E$ y $p_C$, respectivamente) y calcularon tres índices de tamaño del efecto: 1) la diferencia de proporciones; 2) razón de ventajas (*odds ratio* -OR-) y 3) su transformación a diferencia media tipificada (d). La fórmula empleada para realizar esta transformación, es la misma que se utilizó en el presente trabajo, para calcular $d_{HH}$.

## MÉTODO

Para el cumplimiento del objetivo planteado se estimaron los valores de frecuencias observadas de las tablas 2x2 de cada uno de los 13 estudios incluidos en el meta-análisis base. Posteriormente se calcularon los 7 índices de tamaño del efecto. La comparación de los mismos se hizo mediante el cálculo de las correlaciones entre ellos y el estudio de su variabilidad. Asimismo, se realizó la estimación combinada del tamaño del efecto y la prueba de homogeneidad (para cada uno de los 7 estimadores de TE). Por último, se efectuó la búsqueda de algunas variables moderadoras para comparar los resultados del meta-análisis de Sánchez-Meca et al. (1999), con los resultados obtenidos con cada uno de los 7 índices. En el meta-análisis base no se explicita el método de análisis usado (efectos fijos o efectos aleatorios). Partiendo de que el modelo de efectos fijos es el más usado en Ciencias del Comportamiento (Huedo-Medina, Sánchez-Meca y Marín-Martínez, 2004) y de algunos detalles del meta-análisis original, hemos deducido que ese es el modelo de análisis empleado por los autores, por lo cual se trabajó con el modelo de efectos fijos. Ello nos permitía, además, hacer un análisis de sensibilidad dentro del mismo modelo.

Con los datos provenientes sobre los tamaños de los grupos experimental y de control, y de la tasa de abstinencia de cada grupo (porcentaje de sujetos que siguen sin fumar en cada grupo), por cada estudio analizado en el meta-análisis de Sánchez-Meca et al. (1999) se estimaron las frecuencias de éxito ($O_{1E}$, $O_{1C}$) y las frecuencias de fracaso ($O_{2E}$, $O_{2C}$) de los grupos experimental y control. Lipsey y Wilson (2001) indican que un problema al utilizar la razón de ventajas son las celdas con frecuencias observadas igual a 0; en estos casos se puede resolver el problema agregándoles 0.5 a cada una. En el meta-análisis base, los autores proceden de esa manera, agregando 0.5 a las frecuencias de cada celda de todas las tablas 2x2; es decir:

$$O'_{1E} = O_{1E} + 0.5 \quad (19)$$

$$O'_{1C} = O_{1C} + 0.5 \quad (20)$$

donde $O_{1E}$ y $O_{1C}$ son las frecuencias observadas (de éxito) para los grupos experimental y de control, respectivamente, y $O'_{1E}$ y $O'_{1C}$ son las frecuencias corregidas para los grupos experimental y control. Estas correcciones a las frecuencias observadas dan lugar a que se agregara 1 sujeto a $n_E$ y a $n_C$, es decir:

$$n'_E = n_E + 1 \quad (21)$$

$$n'_C = n_C + 1 \quad (22)$$

donde $n_E$ y $n_C$ son los tamaños de los grupos experimental y control, respectivamente, y $n'_E$ y $n'_C$ son las *n* corregidas para ambos grupos.

Con base en las proporciones de abstinencia ($p_E$ y $p_C$) y los tamaños de muestra corregidos ($n'_E$ y $n'_C$) se obtuvieron las frecuencias corregidas para recomponer las tablas 2x2 de cada estudio (tabla 2). Para lo anterior se usaron las siguientes fórmulas:

$$O'_{1E} = (p_E)(n'_E) \quad (23)$$

$$O'_{1C} = (p_C)(n'_C) \quad (24)$$

Posteriormente, se calcularon $O'_{2E}$ y $O'_{2C}$ (frecuencias de fracaso corregidas), con base en $O'_{1E}$, $O'_{1C}$, $n'_E$ y $n'_C$:

$$O'_{2E} = n'_E - O'_{1E} \quad (25)$$

$$O'_{2C} = n'_C - O'_{1C} \quad (26)$$

Asimismo, se calcularon los valores de $m_1$ y $m_2$, siguiendo las fórmulas:

$$m'_1 = O'_{1C} + O'_{1E} \quad (27)$$

$$m'_2 = O'_{2C} + O'_{2E} \quad (28)$$

**Tabla 2. Datos de frecuencias corregidas de cada estudio.**

| | $O'_{1E}$ | $O'_{1C}$ | $O'_{2E}$ | $O'_{2C}$ | $m'_1$ | $m'_2$ |
|---|---|---|---|---|---|---|
| e1 | 16.5 | 0.5 | 8.5 | 16.5 | 17 | 25 |
| e2 | 11.5 | 0.5 | 8.5 | 16.5 | 12 | 25 |
| e3 | 7.5 | 0.5 | 14.5 | 16.5 | 8 | 31 |
| e4 | 7.5 | 0.5 | 15.5 | 16.5 | 8 | 32 |
| e5 | 0.5 | 0.5 | 7.5 | 14.5 | 1 | 22 |
| e6 | 7.5 | 0.5 | 1.5 | 14.5 | 8 | 16 |
| e7 | 36.5 | 3.5 | 2.5 | 25.5 | 40 | 28 |
| e8 | 8.5 | 3.5 | 5.5 | 10.5 | 12 | 16 |
| e9 | 7.5 | 3.5 | 6.5 | 10.5 | 11 | 17 |
| e10 | 12.5 | 3.5 | 4.5 | 10.5 | 16 | 15 |
| e11 | 34.5 | 0.5 | 6.5 | 40.5 | 35 | 47 |
| e12 | 22.5 | 0.5 | 14.5 | 40.5 | 23 | 55 |
| e13 | 51.5 | 3.5 | 30.5 | 33.5 | 55 | 64 |

# RESULTADOS

En la tabla 3 se muestran los valores obtenidos en los diferentes índices de TE, para cada estudio. Se presentan 3 valores para el índice $d_{bis}$, debido a que en 3 de los estudios se requirieron ajustes para poder estimarlo. Al calcular la fórmula 16, para el coeficiente phi biserial ($\phi_{bis}$), se obtenían valores mayores a 1 en los estudios 6, 7 y 11:

$$\phi_{bis} = \frac{\sqrt{p'q'}}{y'}\phi, \qquad e6= 1.07; \qquad e7= 1.04; \qquad e11= 1.06$$

Estos valores ocasionaban que al calcular la fórmula 17 para obtener $d_{bis}$, se requiriera calcular la raíz cuadrada de un número negativo. Para resolverlo se intentaron 3 procedimientos[1]:

1) Calcular la raíz cuadrada del valor absoluto en la fórmula para $d_{bis}$:

$$d_{bis} = \frac{\phi_{bis}}{\sqrt{\left|1-\phi_{bis}^2\right|}}\sqrt{\frac{df(n_E + n_C)}{n_E n_C}}$$

2) Cambiar los valores de $\phi_{bis}$ por 0.99, en los 3 estudios

3) Cambiar los valores de $\phi_{bis}$ por 0.90, en los 3 estudios

---

[1] Agradecemos a Julio Sánchez-Meca, quien sugirió las opciones 2 y 3 en comunicación personal, con el propósito de resolver los problemas con el denominador de la fórmula 17.

**Tabla 3. Índices de tamaño del efecto.**

| | $d_p$ | $d_\phi$ | $d_{asin}$ | $d_{HH}$ | $d_{cox}$ | $d_{probit}$ | $d_{bis}$ (aj. Abs)* | $d_{bis}$ (aj.0.99)* | $d_{bis}$ (aj.0.90)* |
|---|---|---|---|---|---|---|---|---|---|
| e1 | 1.65 | 1.61 | 1.55 | 2.29 | 2.51 | 2.29 | 2.63 | -- | -- |
| e2 | 1.43 | 1.39 | 1.37 | 2.09 | 2.29 | 2.06 | 2.25 | -- | -- |
| e3 | 0.83 | 0.81 | 0.90 | 1.55 | 1.70 | 1.47 | 1.26 | -- | -- |
| e4 | 0.80 | 0.78 | 0.87 | 1.52 | 1.67 | 1.44 | 1.21 | -- | -- |
| e5 | 0.15 | 0.15 | 0.15 | 0.40 | 0.44 | 0.33 | 0.33 | -- | -- |
| e6 | 3.03 | 2.87 | 1.94 | 2.80 | 3.07 | 2.84 | 5.57 | 13.87 | 4.08 |
| e7 | 2.91 | 2.87 | 1.93 | 2.59 | 2.84 | 2.73 | 7.39 | 13.97 | 4.11 |
| e8 | 0.78 | 0.75 | 0.75 | 0.85 | 0.94 | 0.95 | 0.99 | -- | -- |
| e9 | 0.62 | 0.60 | 0.60 | 0.69 | 0.76 | 0.77 | 0.78 | -- | -- |
| e10 | 1.11 | 1.07 | 1.01 | 1.17 | 1.28 | 1.29 | 1.47 | -- | -- |
| e11 | 3.10 | 3.06 | 2.11 | 3.41 | 3.74 | 3.32 | 5.98 | 13.86 | 4.08 |
| e12 | 1.74 | 1.71 | 1.58 | 2.73 | 3.00 | 2.61 | 3.42 | -- | -- |
| e13 | 1.23 | 1.22 | 1.21 | 1.54 | 1.69 | 1.67 | 1.70 | -- | -- |

* Los ajustes sólo se realizaron para los estudios 6, 7 y 11.

La característica que tienen en común estos 3 estudios (6, 7 y 11) es que las frecuencias observadas son desiguales entre los grupos experimental y control; asimismo, se puede asumir fácilmente que el tamaño del efecto será grande, tomando en cuenta los valores de la diferencia de proporciones de éxito (abstinencia) entre el grupo experimental y el grupo de control (.80, .81 y .83, respectivamente, ver tabla 4). El valor que les seguiría en la diferencia de proporciones es la que corresponde al estudio 12 (dp= .60); el valor de $d_{bis}$ para este estudio también es mayor que el valor de $d_{Cox}$. Conviene tener presentes los resultados y recomendaciones de Sánchez-Meca et al. (2003) respecto a que $d_{bis}$ sobreestima el tamaño del efecto y que solo se recomienda cuando se puede asumir que el tamaño del efecto poblacional es pequeño o medio.

Con los valores presentados en la tabla 3, puede observarse que los estudios 1, 2, 3 y 4 tienen un comportamiento muy similar en lo que respecta al ordenamiento de los índices de TE. En los estudios 5, 8, 9 y 10 todos los índices abarcan un rango menor; en otras palabras, están concentrados prácticamente en el mismo valor. El comportamiento de los índices en los estudios 6 y 7 es muy similar; así como entre los estudios 11, 12 y 13. Asimismo, los valores de $d_{bis}$ en los estudios 6 y 7 se alejan de los valores obtenidos mediante otros índices.

**Tabla 4. Tamaño de cada grupo por estudio y diferencia de proporciones.**

| | n grupo experimental | n grupo control | muestra total | dp* |
|---|---|---|---|---|
| e1 | 24 | 16 | 40 | 0.63 |
| e2 | 19 | 16 | 35 | 0.55 |
| e3 | 21 | 16 | 37 | 0.31 |
| e4 | 22 | 16 | 38 | 0.3 |
| e5 | 7 | 14 | 21 | 0.03 |
| e6 | 8 | 14 | 22 | 0.8 |
| e7 | 38 | 28 | 66 | 0.81 |
| e8 | 13 | 13 | 26 | 0.36 |
| e9 | 13 | 13 | 26 | 0.29 |
| e10 | 16 | 13 | 29 | 0.48 |
| e11 | 40 | 40 | 80 | 0.83 |
| e12 | 36 | 40 | 76 | 0.6 |
| e13 | 81 | 36 | 117 | 0.53 |

*dp: diferencia de proporciones de éxito (dp = pt-pc)

En los estudios 5, 8 y 9, los valores de tamaño del efecto calculados con los 7 índices son muy cercanos; estos 3 estudios presentan valores pequeños en la diferencia de proporciones de éxito (ver tabla 4); sin embargo los estudios 3 y 4 también presentan esta característica, sólo que en ambos el tamaño del grupo experimental es mayor al tamaño del grupo control. En el estudio 10 también es mayor el tamaño del grupo experimental que del grupo control, la diferencia de proporciones es levemente mayor a la observada en el estudio 8 (ver tabla 4) y los valores de los 7 índices de tamaño del efecto son cercanos.

Tomando como referencia $d_{Cox}$, debido a que fue el índice con mejor ejecución en el estudio de simulación de Sánchez-Meca et al. (2003), podemos observar en la tabla 3 que $d_p$, $d_\phi$, $d_{HH}$ y $d_{asine}$ muestran tamaños del efecto menores a $d_{Cox}$ en todos los estudios, excepto el estudio 7, donde $d_p$ y $d_\phi$ muestran valores levemente mayores que $d_{Cox}$.

Para cuantificar la similitud entre los valores de los 7 índices de tamaño del efecto se calculó la correlación de Pearson entre ellos, incluyendo las 3 opciones de ajuste para $d_{bis}$, lo que da lugar a la matriz que se presenta en la tabla 5. En dicha tabla puede observarse que las correlaciones menores se obtienen cuando en el par de índices está involucrada $d_{bis}$; en el resto, la correlación es siempre mayor de .90.

Asimismo, se observa que al utilizar la tercera opción de ajuste para $d_{bis}$ (ajustando el valor de $\phi_{bis}$ a 0.90 al aplicar la fórmula para obtener $d_{bis}$) las correlaciones son más fuertes que con las opciones 1 y 2; es decir, los valores para los tamaños del efecto se asemejan más a los obtenidos por otros estimadores.

**Tabla 5. Correlaciones entre los diferentes estimadores del tamaño del efecto*.**

| | $d_p$ | $d_\phi$ | $d_{asin}$ | $d_{HH}$ | $d_{cox}$ | $d_{probit}$ | $d_{bis}$ (aj. Abs) | $d_{bis}$ (aj.0.99) | $d_{bis}$ (aj.0.90) |
|---|---|---|---|---|---|---|---|---|---|
| $d_p$ | 1.000 | | | | | | | | |
| $D_\phi$ | 1.000 | 1.000 | | | | | | | |
| $d_{asin}$ | .961 | .963 | 1.000 | | | | | | |
| $d_{HH}$ | .913 | .916 | .964 | 1.000 | | | | | |
| $d_{cox}$ | .912 | .915 | .963 | 1.000 | 1.000 | | | | |
| $d_{probit}$ | .939 | .942 | .986 | .995 | .995 | 1.000 | | | |
| $d_{bis}$ (aj. Abs) | .969 | .972 | .905 | .858 | .858 | .885 | 1.000 | | |
| $d_{bis}$ (aj.0.99) | .949 | .948 | .830 | .784 | .783 | .809 | .961 | 1.000 | |
| $d_{bis}$ (aj.0.90) | .973 | .974 | .972 | .954 | .954 | .970 | .952 | .884 | 1.000 |

* En el recuadro aparecen los coeficientes que no implican $d_{bis}$.

Las correlaciones más estrechas se obtienen entre $d_p$ y $d_\phi$ y entre $d_{cox}$ y $d_{HH}$, encontrándose una correlación perfecta (r=1.00) entre los pares de índices mencionados. Seguida por las correlaciones entre $d_{HH}$ y $d_{probit}$ y entre $d_{cox}$ y $d_{probit}$ (r=.99).

Las correlación más débil es la obtenida entre $d_{Cox}$ y la segunda opción de ajuste para $d_{bis}$ (ajustando el valor de $\phi_{bis}$ a 0.99). En síntesis, los estimadores que muestran consistentemente valores más cercanos son $d_{Cox}$, $d_{probit}$ y $d_{HH}$.

En la tabla 6 se pueden observar las varianzas estimadas para los 7 índices en cada uno de los 13 estudios. $S^2d_{bis}$ presenta en 3 de los estudios presenta valores anómalos. Al ajustarlos obteniendo la raíz cuadrada de un valor absoluto en la fórmula de $d_{bis}$ el valor de $S^2d_{bis}$ es negativo y al realizar el ajuste de $\phi_{bis}$=0.99, el valor es excesivamente grande. Realizando el ajuste de de $\phi_{bis}$=0.90, $S^2d_{bis}$ continua presentando valores grandes.

**Tabla 6. Varianzas.**

| | $S^2\ d_p$ | $S^2\ d_\phi$ | $S^2\ d_{asin}$ | $S^2\ d_{HH}$ | $S^2\ d_{cox}$ | $S^2\ d_{probit}$ | $S^2\ d_{bis}$ (aj. Abs)* | $S^2\ d_{bis}$ (aj.0.99)* | $S^2\ d_{bis}$ (aj.0.90)* |
|---|---|---|---|---|---|---|---|---|---|
| e1 | 0.13 | 0.27 | 0.10 | 0.67 | 0.81 | 0.44 | 1.99 | -- | -- |
| e2 | 0.14 | 0.25 | 0.11 | 0.68 | 0.82 | 0.45 | 1.54 | -- | -- |
| e3 | 0.11 | 0.14 | 0.10 | 0.68 | 0.82 | 0.45 | 0.51 | -- | -- |
| e4 | 0.11 | 0.14 | 0.10 | 0.67 | 0.81 | 0.44 | 0.47 | -- | -- |
| e5 | 0.19 | 0.19 | 0.19 | 1.37 | 1.65 | 0.92 | 1.06 | -- | -- |
| e6 | 0.37 | 1.72 | 0.18 | 0.94 | 1.13 | 0.66 | -32.27 | 12232.94 | 14.05 |
| e7 | 0.12 | 0.57 | 0.06 | 0.23 | 0.28 | 0.20 | -65.10 | 3975.52 | 4.57 |
| e8 | 0.15 | 0.19 | 0.14 | 0.21 | 0.25 | 0.25 | 0.40 | -- | -- |
| e9 | 0.15 | 0.17 | 0.14 | 0.20 | 0.25 | 0.25 | 0.33 | -- | -- |
| e10 | 0.15 | 0.22 | 0.13 | 0.21 | 0.25 | 0.24 | 0.61 | -- | -- |
| e11 | 0.11 | 0.56 | 0.05 | 0.74 | 0.89 | 0.43 | -12.46 | 2895.67 | 3.32 |
| e12 | 0.07 | 0.16 | 0.05 | 0.72 | 0.86 | 0.42 | 3.27 | -- | -- |
| e13 | 0.05 | 0.07 | 0.04 | 0.11 | 0.14 | 0.11 | 0.20 | -- | -- |

* Los ajustes solo se realizaron para los estudios 6, 7 y 11.

Con respecto a las varianzas estimadas, Sánchez-Meca et al. (2003), encontraron el siguiente orden ascendente: $S^2\ d_p$, $S^2\ d_\phi$, $S^2\ d_{asin}$, $S^2\ d_{HH}$, $S^2\ d_{cox}$, $S^2\ d_{probit}$, $S^2\ d_{bis}$. Con los datos del meta-análisis base, el orden ascendente (en la mayoría de los estudios) es: $S^2\ d_{asin}$, $S^2\ d_p$, $S^2\ d_\phi$, $S^2\ d_{probit}$, $S^2\ d_{HH}$, $S^2\ d_{cox}$ (ver tabla 6). $S^2 d_{bis}$ no muestra un patrón característico. Incluso en los estudios 6, 7 y 11, con las 3 opciones de ajuste realizado los valores son grandes y en una de las opciones los valores son anómalos (negativos). Se presume que estos resultados pueden tener relación con lo mencionado previamente respecto a la mala ejecución de $d_{bis}$ cuando el tamaño del efecto poblacional es grande. Además, este índice esta basado en el supuesto de una distribución normal, pero es lógico suponer que los datos correspondientes a estos estudios no tienen distribución normal.

Por otro lado, de las 3 opciones de ajuste para $d_{bis}$ la única con un valor aceptable para $S^2 d_{bis}$ es con la tercera opción (ajustando el valor de $\phi_{bis}$ a 0.90 al aplicar la fórmula para obtener $d_{bis}$). Aun así, los valores de $S^2\ d_{bis}$ son muy distantes del resto de las varianzas estimadas.

Se obtuvieron también las correlaciones de Pearson entre las desviaciones típicas (DT) de los índices de TE, con la finalidad de cuantificar la cercanía de los valores. Se incluye sólo la opción 3 de ajuste para $d_{bis}$, ya que $S^2 d_{bis}$ en las otras 2 opciones resulta anómalo en los estudios 6, 7 y 11. Estas correlaciones se presentan en la tabla 7.

Las correlaciones más fuertes son las que se dan entre las desviaciones típicas de $d_{cox}$ y $d_{HH}$ (r=.996), seguida por las correlaciones entre las desviaciones típicas de $d_p$ y $d_{asin}$ (r=.947), de $d_{cox}$ y $d_{probit}$ (r=.929) y de $d_{HH}$ y $d_{probit}$ (r=.821). Por otro lado se encuentra un valor alto (r=.802) al correlacionar la desviación típica de $d_\phi$ con la de $d_{bis(aj.\ 0.90)}$.

**Tabla 7. Correlación entre las desviaciones típicas de los estimadores de tamaño del efecto.**

| | $d_p$_dt | $d_\phi$_dt | $d_{asin}$_dt | $d_{HH}$_dt | $d_{cox}$_dt | $d_{probit}$_dt | $d_{bis\ (aj.0.90)}$_dt |
|---|---|---|---|---|---|---|---|
| $d_p$_dt | 1.000 | | | | | | |
| $d_\phi$_dt | .513 | 1.000 | | | | | |
| $d_{asin}$_dt | .947** | .249 | 1.000 | | | | |
| $d_{HH}$_dt | .187 | .361 | .199 | 1.000 | | | |
| $d_{cox}$_dt | .188 | .350 | .204 | .996** | 1.000 | | |
| $d_{probit}$_dt | .377 | .178 | .490 | .821** | .829** | 1.000 | |
| $d_{bis\ (aj.0.90)}$_dt | .144 | .802** | -.055 | .676* | .665* | .328 | 1.000 |

** p< .01; * p< .05

A partir de los análisis precedentes decidimos trabajar a partir de aquí solamente con la opción 3 para $d_{bis}$ (ajustando el valor de $\phi_{bis}$ a 0.90), por lo que en los análisis siguientes se excluyen las otras 2 alternativas exploradas anteriormente.

***Estimación combinada del tamaño del efecto y prueba de homogeneidad***

Para la estimación combinada del tamaño del efecto, su intervalo de confianza y la prueba de homogeneidad, se utilizaron las fórmulas descritas en el texto de Botella y Gambara (2001). La siguiente fórmula se refiere al cálculo del tamaño del efecto global:

$$d_\bullet = \frac{\sum w_i * d_i}{\sum w_i} \qquad (29)$$

donde $d_i$ y $w_i$ son las estimaciones y pesos de cada estudio (Botella y Gambara, 2002); el peso es el resultado del inverso de la varianza.

La fórmula 29 fue aplicada a los 7 índices, para estimar un tamaño del efecto global de cada uno de ellos. En los cálculos realizados para obtener

$w_i$ se utilizó el inverso de la varianza obtenida de la fórmula correspondiente a cada uno de los estimadores del tamaño del efecto.

Para estimar los intervalos de confianza, considerando un nivel de confianza de $\alpha$=.95; se aplicó la fórmula:

$$d_{\bullet} \pm z_{1-\alpha/2} * \sqrt{\frac{1}{\sum w_i}} \qquad (30)$$

Para el contraste sobre la homogeneidad, se utilizó la fórmula:

$$Q_t = \sum w_i * d^2 - \frac{(\sum w_i * d_i)^2}{\sum w_i} \qquad (31)$$

En la tabla 8 se presentan los resultados de la estimación combinada del tamaño del efecto, el intervalo de confianza y la prueba de homogeneidad con los 13 estudios, para cada uno de los 7 índices.

En la estimación combinada del tamaño del efecto, el índice con mayor valor fue $d_{probit}$ (1.785) y el menor fue $d_{\phi}$ (1.131); se confirma que los valores mas cercanos son los de $d_{HH}$, $d_{Cox}$ y $d_{probit}$.

Sánchez-Meca et al. (2003) reportaron en su estudio de simulación que $d_{Cox}$ fue el índice con mejor ejecución, seguido por $d_{probit}$. Asimismo, encontraron que conforme se incrementa el tamaño del efecto poblacional, se incrementa la sobreestimación de $d_{Cox}$, $d_{probit}$ y $d_{bis}$, y la subestimación de $d_{HH}$, $d_p$, $d_{\phi}$ y $d_{asine}$. Tomando esto en cuenta, vemos que efectivamente se obtienen los valores menores para $d_{\phi}$, $d_{asine}$, $d_p$ y $d_{HH}$ (en ese orden) y los valores mayores para $d_{Cox}$ y $d_{probit}$. Consideramos que $d_{bis}$ es un índice con mala ejecución; posteriormente retomaremos este punto.

Basándose en los intervalos de confianza presentados en la tabla 8, con los 7 índices se rechaza la hipótesis nula de que el tamaño del efecto es nulo ($\delta$=0), dado que el intervalo no incluye el valor 0 en ninguno de los casos.

Respecto a la prueba de homogeneidad, debido a que $Q_t$ se distribuye según $\chi^2$ con k-1 grados de libertad y considerando que todos los valores de $Q_t$ son mayores que ${}_{.95}\chi^2(12)=21.03$, con excepción del índice $d_{bis}$, se rechaza la hipótesis nula y se concluye que los estudios son heterogéneos. En el meta-análisis base se llegó a esta misma conclusión. Es importante señalar que $d_{bis}$ no nos lleva a esta conclusión, lo que pone en duda la ejecución de este índice y la validez de las conclusiones alcanzadas mediante su uso.

**Tabla 8. Distribución de los índices del tamaño del efecto.**

| | Estimación combinada del tamaño del efecto | Intervalo de confianza al 95% | Prueba de homogeneidad ($Q_t$) |
|---|---|---|---|
| $d_p$ | 1.466 | 1.282; 1.649 | 76.317* |
| $d_\phi$ | 1.131 | .893; 1.379 | 26.159* |
| $d_{asine}$ | 1.378 | 1.219; 1.536 | 40.441* |
| $d_{HH}$ | 1.590 | 1.273; 1.907 | 21.795* |
| $d_{Cox}$ | 1.744 | 1.396; 2.092 | 21.722* |
| $d_{probit}$ | 1.785 | 1.486; 2.084 | 24.989* |
| $d_{bis\ (aj.\ 0.90)}$ | 1.423 | .975; 1.871 | 9.959 |
| Meta-análisis base | 1.800 | 1.608; 1.992 | 67.325* |

*valor mayor que el de ${}_{.95}\chi^2(12)=21.03$

***Búsqueda de variables moderadoras***

Para el análisis de las variables que expliquen la variabilidad en los resultados de los estudios, los autores del meta-análisis base realizaron un análisis de regresión por mínimos cuadrados ponderados (MCP), desde el enfoque de Hedges y Olkin (1985 en Sánchez-Meca et al., 1999), donde se obtiene $Q_R$ de la suma de cuadrados del modelo de regresión y $Q_E$ de la suma de cuadrados residual. Si $Q_R$ es significativa el modelo explica una parte significativa de la varianza. Si $Q_E$ es significativa, aunque $Q_R$ explique parte de la varianza, queda una parte significativa sin explicar y deben buscarse otras variables moderadoras de los resultados.

En el meta-análisis base, los autores encontraron como variables influyentes en los resultados la duración de la intervención, la intensidad de la misma y el modo de entrenamiento. Asimismo, encontraron que la edad y el género son factores que influyen en la eficacia de los tratamientos.

Para comparar los resultados obtenidos por los diferentes estimadores del tamaño del efecto, se tomaron solamente 3 de estas variables para realizar los análisis de regresión. En concreto, se tomaron la duración de la intervención, su intensidad y la edad de los sujetos. Debemos aclarar que en las tablas 9, 10 y 11 el valor del coeficiente de regresión de $d_{HH}$ no coincide con el valor del meta-análisis base debido a las fórmulas empleadas para estimar la variabilidad. En el presente estudio se siguió la fórmula para $S^2 d_{HH}$ presentada por Sánchez-Meca et al. (2003), que es diferente a la empleada en el meta-análisis base.

**Tabla 9. Análisis de regresión tomando como variable moderadora la duración del tratamiento (k=11).**

| Índice de TE | Coeficient e regresión | $R_{aj}^2$ | $Q_R$ | $Q_E$ | gl |
|---|---|---|---|---|---|
| $d_p$ | 0.040 | -0.110 | 0.074 | 60.612** | 9 |
| $d_\phi$ | 0.104 | -0.092 | 0.328 | 18.752** | 9 |
| $d_{asine}$ | 0.100 | -0.090 | 0.573 | 29.968** | 9 |
| $d_{HH}$ | 0.596 | 0.197 | 5.332* | 13.910 | 9 |
| $d_{Cox}$ | 0.653 | 0.197 | 5.295* | 13.818 | 9 |
| $d_{probit}$ | 0.428 | 0.067 | 3.369 | 17.701** | 9 |
| $d_{bis\ (aj.\ 0.90)}$ | 0.316 | 0.015 | 0.945 | 7.367 | 9 |
| Meta-análisis base | 1,105 | 0.209 | 5.561* | 13.719 | 9 |

k: Número de estudios. C. regres: coeficiente de regresión. $Q_R$: Suma de cuadrados debida a la regresión. $Q_E$: suma de cuadrados de error. gl: Grados de libertad asociados a $Q_E$
*valor mayor que el de ${}_{.95}\chi^2(1)=3.841$ (para $Q_R$)
**valor mayor que el de ${}_{.95}\chi^2(9)=16.92$ (para $Q_E$)

En el meta-análisis base se concluye que las intervenciones más prolongadas en el tiempo logran mayores tasas de abstinencia respecto de los grupos de control. Esta conclusión se apoya en el resultado obtenido por el modelo de regresión por MCP [$Q_R(1)=5.561$; $p<.05$], con un 21% de varianza explicada. Ninguno de los índices de TE aplicados en este estudio llevan a un porcentaje de varianza explicada tan alto. La dirección de la relación se refleja en el signo del coeficiente de regresión para todos los índices; sin embargo, $Q_R$ solo es significativa para $d_{HH}$ y $d_{Cox}$ (ver tabla 9).

Respecto a la intensidad del tratamiento, definida como el número medio de horas semanales de tratamiento por sujeto, en el meta-análisis base se concluye que está negativamente relacionada con la eficacia; es decir, las intervenciones más intensas son menos eficaces para lograr la abstinencia al tabaco. Esta conclusión se apoya en el resultado obtenido por el modelo de regresión [$Q_R(1)=4.559$; $p<.05$]. Nuevamente, la dirección de esta relación se manifiesta en los signos de los coeficientes de regresión para todos los índices. $Q_R$ es significativa para $d_{HH}$ y $d_{Cox}$, y además, para $d_{probit}$ y $d_{asin}$; solo que para $d_{asin}$ también $Q_E$ es significativa, lo que indicaría que queda un porcentaje significativo de varianza por explicar (tabla 10).

Por último, en la tabla 11 se presenta el análisis de regresión tomando la edad media de los sujetos como variable moderadora. En el meta-análisis base la edad media de las muestras de sujetos alcanzó el porcentaje más alto de varianza explicada (55.7%); solamente al hacer los análisis con $d_{HH}$, $d_{Cox}$ y $d_{probit}$ se obtienen varianzas cercanas a este valor.

**Tabla 10. Análisis de regresión tomando como variable moderadora la intensidad del tratamiento (k=11).**

| Índice de TE | Coeficiente regresión | $R_{aj}^2$ | $Q_R$ | $Q_E$ | gl |
|---|---|---|---|---|---|
| $d_p$ | -0.245 | -0.043 | 3.717 | 56.969** | 9 |
| $d_\phi$ | -0.163 | -0.046 | 1.111 | 17.969** | 9 |
| $d_{asine}$ | -0.235 | 0.034 | 3.988* | 26.553** | 9 |
| $d_{HH}$ | -0.353 | 0.147 | 4.478* | 14.765 | 9 |
| $d_{Cox}$ | -0.388 | 0.149 | 4.480* | 14.633 | 9 |
| $d_{probit}$ | -0.365 | 0.131 | 4.584* | 16.487 | 9 |
| $d_{bis\ (aj.\ 0.90)}$ | -0.219 | -0.020 | 0.680 | 7.632 | 9 |
| Meta-análisis base | -0.646 | 0.152 | 4.559* | 14.721 | 9 |

k: Número de estudios. C. regres: coeficiente de regresión. $Q_R$: Suma de cuadrados debida a la regresión. $Q_E$: suma de cuadrados de error. gl: Grados de libertad asociados a $Q_E$
*valor mayor que el de $_{.95}\chi^2(1)=3.841$ (para $Q_R$)
**valor mayor que el de $_{.95}\chi^2(9)=16.92$ (para $Q_E$)

**Tabla 11. Análisis de regresión tomando como variable moderadora la edad media de los sujetos (k=13).**

| Índice de TE | Coeficiente regresión | $R_{aj}^2$ | $Q_R$ | $Q_E$ | gl |
|---|---|---|---|---|---|
| $d_p$ | 0.086 | 0.105 | 13.749* | 62.651** | 11 |
| $d_\phi$ | 0.041 | -0.006 | 2.018 | 24.056** | 11 |
| $d_{asine}$ | 0.062 | 0.157 | 9.157* | 31.104** | 11 |
| $d_{HH}$ | 0.113 | 0.562 | 13.067* | 8.779 | 11 |
| $d_{Cox}$ | 0.124 | 0.561 | 12.972* | 8.730 | 11 |
| $d_{probit}$ | 0.109 | 0.440 | 12.205* | 12.851 | 11 |
| $d_{bis\ (aj.\ 0.90)}$ | 0.064 | 0.091 | 1.665 | 8.306 | 11 |
| Meta-análisis base | 0,206 | 0.557 | 13.056* | 8.932 | 11 |

k: Número de estudios. C. regres: coeficiente de regresión. $Q_R$: Suma de cuadrados debida a la regresión. $Q_E$: suma de cuadrados de error. gl: Grados de libertad asociados a $Q_E$
*valor mayor que el de $_{.95}\chi^2(1)=3.841$ (para $Q_R$)
**valor mayor que el de $_{.95}\chi^2(11)=19.68$ (para $Q_E$)

Asimismo, en el meta-análisis base se encontró una asociación significativa entre la edad y los tamaños del efecto [$Q_R(1)=13.056$; $p<.001$]. Según estos datos, las muestras de sujetos de mayor edad tendieron a obtener mejores resultados. Esta relación se confirma de nuevo con los signos de los coeficientes de regresión para todos los índices de TE y la significación de $Q_R$ para $d_p$, $d_{asin}$, $d_{HH}$, $d_{Cox}$ y $d_{probit}$. Solo que para $d_p$ y $d_{asin}$, $Q_E$ también es significativa, lo que indica que queda varianza por explicar (tabla 11).

## CONCLUSIONES

Nuestro objetivo era comparar el comportamiento de los 7 índices explorados por Sánchez-Meca et al. (2003) en su estudio de simulación, con datos reales provenientes del meta-análisis base. Podemos concluir que llegamos a las mismas conclusiones que en ese estudio. Considerando las preguntas iniciales, podemos concluir que las opciones para calcular el tamaño del efecto para variables dicotomizadas no son equivalentes. Asimismo, se pone de manifiesto que el uso de uno u otro índice para calcular el tamaño del efecto puede tener consecuencias importantes, e incluso llevar a conclusiones opuestas en un meta-análisis. Es por esto por lo que con base en los resultados presentados tanto en la simulación realizada por Sánchez-Meca et al. (2003), como en los resultados del presente trabajo, se recomienda que para cálculos del tamaño del efecto para variables dependientes dicotomizadas se utilicen $d_{Cox}$, $d_{probit}$ y $d_{HH}$ para poder comparar sus resultados e incrementar las garantías de un meta-análisis. Esta forma de proceder, que implica un análisis de sensibilidad de las tres opciones (Greenhouse e Iyengar, 1994), es la que maximiza la validez de las conclusiones.

En la comparación de los 7 estimadores de tamaño del efecto, se encontró que $d_{Cox}$, $d_{probit}$ y $d_{HH}$ mostraron ser los más consistentes, con correlaciones más fuertes entre ellos. Por otra parte, se observa que solo $d_{Cox}$ y $d_{HH}$ son los únicos índices que confirman los resultados encontrados en todos los análisis comparados con el meta-análisis base. Es razonable, ya que tanto $d_{Cox}$ como $d_{HH}$ son índices basados en el logaritmo de razón de ventajas y que en el meta-análisis base, para buscar las variables moderadoras, se tomó el tamaño del efecto obtenido por la misma fórmula para calcular $d_{HH}$.

Sin embargo, tomando en cuenta que Haddock et al. (1998) citan a varios autores que consideran que la razón de ventajas y la transformación logarítmica de odds, son las medidas mas adecuadas para calcular tamaño del efecto para tablas 2x2, así como los resultados obtenidos en la comparación de los 7 estimadores, podemos concluir que $d_{Cox}$, $d_{probit}$ y $d_{HH}$ son los índices con mejor ejecución.

Respecto a la estimación de los diferentes índices, no es posible calcular $d_{bis}$ cuando el valor del coeficiente phi biserial ($\phi_{bis}$) es mayor a 1, lo que sucedió cuando había claras diferencias en las frecuencias observadas en las tablas 2x2 (estudios 6, 7 y 11). Se intentaron 3 ajustes a la fórmula, pero con ninguno de ellos se logró una buena ejecución. Ésta conclusión se respalda con lo indicado por Haddock et al. (1998) quienes indican que $\phi$ no provee una medida satisfactoria de tamaño del efecto cuando en una tabla 2x2 las distribuciones marginales no son equivalentes; esto debe ser considerado para $d_{bis}$ ya que implica el calculo de $\phi$.

## ABSTRACT

Comparison among Effect-Size indices for dichotomized outcomes in Meta-analysis. The aim of this study was compare with information of a meta-analysis actually published, the performance of 7 different effect-size indices for estimating the population standardized mean difference from a 2 × 2 table ($d_p$, $d_\phi$, $d_{asine}$, $d_{HH}$, $d_{Cox}$, $d_{probit}$, $d_{bis}$). Those indices were studied by Sánchez-Meca, Marín-Martínez and Chacón-Moscoso (2003) using Monte Carlo simulation. In the same line of that simulation, our results show that the choice of the index may have important consequences in the conclusions of the meta-analysis. We recommended the simultaneous use of $d_{Cox}$, $d_{probit}$ and $d_{HH}$.

## REFERENCIAS

Botella, J. y Gambara, H. (2002). *¿Qué es el meta-análisis?* España: Biblioteca Nueva.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cooper, H. (1998). *Synthesizing research: A guide for literature reviews.* Beverly Hills, CA: Sage. 3ª edición.

Fleiss, J. (1994). Measures of effect size for categorical data. En H. Cooper y L. V. Hedges (Eds.), *The Handbook of research synthesis.* New York: Russell Sage Foundation, 245-260.

Glass, G. (1976). Primary, secondary and meta-analysis. *Educational Researcher, 5* 3-8.

Greenhouse, J. e Iyengar, S. (1994). Sensitivy analysis and diagnostics. En H. Cooper y L. V. Hedges (Eds.), *The Handbook of research synthesis.* New York: Russell Sage Foundation, 383-399.

Haddock, K., Rindskopf, D. y Shadish, W. (1998). Ussing Odds Ratios as Effect Sizes for Meta-Analysis of Dichotomous Data: A Primer on Methods and Issues. *Psychological Methods, 3* (3), 339-353.

Hair, J., Anderson, R., Tatham, R. y Black, W. (1999). *Análisis Multivariante*. España: Prentice Hall. 4ª edición.

Huedo-Medina, T., Sánchez-Meca, J., y Marín-Martínez, F. (2004). La estimación del tamaño del efecto medio en un meta-análisis: Una comparación entre los modelos de efectos fijos y aleatorios. *Metodología de las Ciencias del Comportamiento, Volumen Especial,* 307-315.

Lipsey, M.W. y Wilson, D.B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Longley-Cook, L. (1981). *Problemas de estadística y como resolverlos.* México: C.E.C.S.A.

Manterola, C. Riedemann, P. y Vial, M. (2001). Estrategias de investigación. Un diseño observacional analítico. El meta-análisis *Revista Chilena de Cirugía 53* (6), 615-621.

Sánchez-Meca, J., Marín-Martínez, F., Olivares, J. y Rosa, A. (1999). Variables influyentes en el tratamiento de la adicción al tabaco. Un estudio de las tasas de abstinencia en España. *Psicología Conductual, 7* (2), 301-321.

Sánchez-Meca, J., Marín-Martínez, F. y Chacón-Moscoso, S. (2003). Effect-Size indices for dichotomized outcomes in Meta-analysis. *Psychological Methods, 8* (4), 448-467.

(Manuscrito recibido: 19 Septiembre 2005; aceptado: 7 Abril 2006)