

Validez de contenido de una escala para valorar la calidad de los estudios primarios en Meta-análisis

IX Congreso de Metodología de las Ciencias Sociales y de la Salud
Granada, 14 al 16 de septiembre de 2005



Salvador Chacón Moscoso
Susana Sanduvete Chaves
David Alarcón Rubio



INTRODUCCIÓN

- Meta-análisis es un método con el que se trata de concluir acerca de la eficacia de las intervenciones sobre una variable dependiente determinada aunando los resultados encontrados en la literatura.

INTRODUCCIÓN (II)

- ¿Por qué es necesario el análisis de la calidad de los estudios primarios?
 - Hay **sesgos** en los resultados cuando se incluyen estudios de baja calidad:
 - Dependiendo de los **criterios de selección**, se tomarán unos estudios u otros, aspecto que determina los resultados que se hallarán.
 - El **grado de eficacia** tiende a mostrarse mayor de lo que es en realidad.
 - Se pueden incluir estudios de **baja y alta** calidad en un solo meta-análisis y con un **tratamiento diferencial**.
 - Los diseños **experimentales** no siempre son de alta calidad, al igual que los **cuasi-experimentales** no son siempre de calidad baja. Se pueden incluir varios tipos en un meta-análisis:
 - **Separadamente** para ser integrados al final.
 - Estudiando los diferentes **tipos** de diseño como una posible **variable moduladora**.



INTRODUCCIÓN (III)

- La mayoría de los programas de intervención carecen en cierto aspecto de rigor metodológico (Chacón, Sanduvete y Alarcón, enviado).
 - El 17.7% especifica la **orientación teórica**.
 - El 15.7% **asigna** aleatoriamente a las personas a los distintos grupos.
 - El 14.7% utiliza un **diseño experimental**.
 - El 15.3% presenta **seguimiento** un año después.
 - El 31.4% presenta **medidas previas y posteriores** a la intervención.
 - El 5.3% presenta **todas** las medidas antes y después de la intervención.
 - El 6.4% presenta al menos una **medida objetiva**.
 - El 17% presenta el **tamaño de efecto** calculado.



INTRODUCCIÓN (III)

- ¿Por qué una escala como instrumento?
 - Es el método más utilizado (Jüni, Altman & Egger, 2001) .
 - Refleja el grado de calidad de un estudio aportando un dato cuantitativo global.
 - Permite el estudio de índice de fiabilidad y validez.

INTRODUCCIÓN (IV)

- Escalas más frecuentes (Chacón, Sanduvete y Alarcón, 2005):

	CONSORT (Consolidated Standards of Reporting Randomized Trials)	TREND (Transparent Reporting of Evaluations with Nonrandomiz ed Designs)	STRICTA (Standards for Reporting Interventions in Controlled Trials of Acupuncture)	STARD (Standards for Reporting of Diagnostic Accuracy)	Guidelines for evaluating papers on educational interventions
Año de publicación	1996 (revisado en el 2001)	2004	2002	2003	1999 (en fase de prueba)
Número de ítems	22	22	6	25	27
Uso	Estudios aleatorios	Estudios no aleatorios	Acupuntura	Diagnósticos médicos	Intervenciones en educación



INTRODUCCIÓN (IV)

- Si ya existen muchas escalas elaboradas, ¿por qué proponer una nueva?
 - Intentar sistematizar indicadores útiles.
 - Estudiar de manera sistemática y empírica la utilidad del uso de escalas para medir la calidad.



OBJETIVO

- Estudiar la validez de contenido de una escala de calidad previamente elaborada.



MÉTODO

□ **Muestra:**

- **Primera fase:** todos los artículos disponibles que versaban sobre la medida de la calidad de estudios primarios. Concretamente, Sánchez-Meca & Ato (1990); O'Rourke & Detsky (1989); Weisz et al. (2000); Tritchler (1999); Jüni et al. (2001); Sánchez-Meca (1997); Sutton et al. (2000); Moher et al. (2001); Begg et al. (1996); Moher et al. (1998); Des Jarlais et al. (2004); Bossuyt et al. (2003); Bossuyt et al. (2003); Olivares et al. (2000); Education Group for Guidelines on Evaluation (1999); Campbell et al. (2004); Bosch et al. (2003); Altman et al. (2001); Brown (1991); Jüni et al. (1999); McGuire et al. (1985); Emerson et al. (1990); Moher et al. (1996); McPherson et al. (2002); Moher et al. (2001); and Greenland (1994).
- **Segunda fase:** 10 expertos en meta-análisis y revisiones sistemáticas.



MÉTODO (II)

□ Instrumentos.

- Bases de datos informatizadas para la búsqueda de la documentación referida al estudio de la calidad.
- Cuestionario con 43 ítems (6 referidos a características extrínsecas de los estudios, 14 a características sustantivas -5 referidos a la muestra, 3 al contexto y 6 al tratamiento- y 23 a características metodológicas) (Sánchez-Meca et al, 1998). Escalas con 3 opciones de respuesta. Tres conceptos: representatividad, utilidad y viabilidad de la recogida del dato.
- Distribución de los cuestionarios y recogida de datos a través de internet y personalmente en encuentros profesionales (*V Annual Campbell Collaboration Colloquium*, Lisboa, Febrero 2005).
- El software Microsoft Excel para el análisis de los datos.



MÉTODO (III)

□ Procedimiento.

1. **Búsqueda** de artículos referidos a la calidad de estudios primarios.
2. Recogida de todos los indicadores de calidad y **elaboración de cuestionario** con todos ellos, encuadrándolos en tres dimensiones: características extrínsecas, sustantivas (de la muestra, el contexto y el tratamiento) y metodológicas.
3. **Distribución** de los cuestionarios a expertos y posterior **recogida**.
4. Análisis de los datos: estudio de la **validez de contenido** a través del índice de congruencia de Osterlind (1998):

MÉTODO (IV)

$$I_{ik} = \frac{(N - 1) \sum_{j=1}^n X_{ijk} + N \sum_{j=1}^n X_{ijk} - \sum_{j=1}^n X_{ijk}}{2(N - 1)n}$$

Siendo:

- N = número de dominios
- X_{ijk} = el valor que cada juez otorga a cada ítem
- n = número de jueces

$I_{ik} = 0.5$ es considerado suficiente grado de idoneidad

RESULTADOS

CARÁCTERÍSTICAS EXTRÍNSECAS (N = 10)	R	U	V
1. Tipo de publicación	0.3	0.6	0.7
2. Año de publicación	0	0.2	0.8
3. Índice de impacto de la revista	-0.2	0.1	0.3
4. Base de datos en que se encontró	-0.2	0.3	0.6
5. Entrenamiento de los investigadores	0.2	0.4	0.1
6. Estructura recomendada por la APA	0.1	0.1	0.1

RESULTADOS (II)

CARÁCTERÍSTICAS SUSTANTIVAS (N = 10)	R	U	V
DE LA MUESTRA			
7. Rango de edad	0.6	0.4	0.6
8. Media de edad	0.8	0.8	0.7
9. Desviación típica de la edad	0.4	0.3	0.5
10. Origen cultural	0.2	0.2	0.2
11. Nivel socioeconómico	-0.1	0	-0.3
DEL CONTEXTO			
12. Contexto de intervención	-0.3	0	0.1
13. Campo de intervención	0.3	0.9	0.6
14. País	0.5	0.9	0.4
DEL TRATAMIENTO			
15. Orientación teórica	0.4	0.8	0.1
16. Evidencia empírica previa	0.2	0.2	0.2
17. Periodo de tratamiento	0.7	0.9	0.6
18. Grado de intensidad del tratamiento	0.7	0.9	0.8
19. Unidades (en grupo o individual)	1	0.9	0.9
20. Los puntos fuertes y las debilidades encontrados son discutidos	0.2	-0.2	-0.1

RESULTADOS (III)

CARÁCTERÍSTICAS METODOLÓGICAS (N = 10)	R	U	V
21. Criterios de inclusión y exclusión de las unidades de la muestra especificados	0.6	0.9	0.6
22. Asignación aleatoria de las unidades a los grupos	0.9	1	0.8
23. Tipo de metodología/ diseño	0.9	1	0.8
24. Tamaño de la muestra	0.7	0.8	1
25. Estadístico utilizado para el cálculo del tamaño de la muestra	0.5	0.6	0.3
26. Mortalidad	0.9	0.9	0.2
27. Sin mortalidad	0.7	0.4	0.5
28. Mortalidad entre grupos	0.9	0.9	0.2
29. Exclusiones posteriores a la asignación aleatoria	0.8	0.8	0.4
30. Periodo de línea base	0.2	0.4	0.1
31. Periodo de seguimiento	0.6	0.8	0.3

RESULTADOS (IV)

CARÁCTERÍSTICAS METODOLÓGICAS (II) (N = 10)	R	U	V
32. Momentos de medida	0.9	0.9	1
33. Las medidas del pre-test aparecen en el postest	0.7	0.8	0.4
34. Variables dependientes normalizadas	0.5	0.6	0.3
35. Homogeneidad de la intervención	0.6	0.3	-0.2
36. Técnicas de control	0.8	0.8	0
37. Definición del constructo	0.9	0.6	-0.3
38. Métodos estadísticos para inferir los valores perdidos	0.7	0.6	0.1
39. Especificación de los intervalos de confianza en los estudios estadísticos	0	-0.1	0.3
40. Tamaño de efecto y valor	0.6	0.7	0.6
41. Otros datos aparte de los objetivos	-0.2	-0.1	0.1
42. Interpretación de los resultados	-0.1	-0.2	-0.1
43. Interpretación de los sesgos de los resultados	0.3	0	0



DISCUSION: posible interpretación

- Se obtendrán más o menos ítems idóneos en función del punto de corte que se utilice (de 129 valores, tomando el 0.5 como punto de corte se consideraron 65 idóneos; si se aumentara a 0.7, disminuiría a 43 valores).
- El **estudio empírico** futuro consistirá en aumentar y disminuir la restrictividad en función del **punto de corte** y de los tres **conceptos** estudiados (representatividad, utilidad y viabilidad) .



CONCLUSIONES

- ❑ **Comprobación empírica:**
 - ❑ **Comparación** de los resultados de la calidad de los estudios primarios incluidos en estudios meta-analíticos medidos con las otras escalas encontradas en la literatura y con la aquí presentada, incluyendo en ésta distintas combinaciones:
 - ❑ Todos los ítems con valoraciones positivas en representatividad, utilidad y viabilidad.
 - ❑ Otras combinaciones.



REFERENCIAS BIBLIOGRÁFICAS

- Chacón, S., Sanduvete, S. y Alarcón, D. (enviado). Avances metodológicos para la mejora de la calidad de los programas de intervención dirigidos a personas mayores. *Anales de Psicología*.
- Chacón, S., Sanduvete, S. & Alarcón, D. (February, 2005). Towards the Validation of a Scale to Measure the Quality of Primary Studies for Meta-analysis. Paper presented at the *V Annual Campbell Collaboration Colloquium*. Lisboa.
- Jüni, P., Altman, D. G. & Egger, M. (2001). Assessing the quality of randomized controlled trials. In M. Egger, G. D. Smith & D. G. Altman (Ed.), *Systematic reviews in health care*. London: BMJ.
- Osterlind, S. J. (1998). *Constructing tests items*. Boston: Kluwer Academic Publishers.
- Sánchez-Meca, J., Rosa, A. I. y Olivares, J. (1998). Las técnicas cognitivo-conductuales en problemas clínicos y de salud: meta-análisis de la literatura española. *Psicothema*, *11*(3), 641-654.