

Acercas de la validación de escalas y cuestionarios

Claudio González¹ y Darío Scublinsky²

¹Bioestadística y Metodología de la Investigación, Universidad Católica Argentina.

²Epidemiología, FEFyM. Farmacología, UBA.

En el terreno de la investigación clínica y aun de la práctica médica actual, es cada día más común el empleo de cuestionarios y escalas de evaluación cuya aplicación merece una serie de consideraciones que conviene tener presente.

En principio, todo cuestionario es una herramienta metodológica y, en consecuencia, caben para su administración juiciosos los mismos principios que rigen a la aplicación de otros instrumentos tales como una balanza, un tensiómetro, un resonador magnético.

No caben dudas acerca de que los cuestionarios propiamente dichos, o los cuestionarios con escalas, constituyen hoy la técnica de obtención de datos más empleada en la investigación clínica. Se trata de herramientas efectivas, poco costosas, que permiten la evaluación a un gran número de participantes con una relativa facilidad y permiten el análisis y el entendimiento del comportamiento de variables que resultaría dificultoso a través de otras vías. Nos permite obtener descripciones apropiadas, datos categóricos y, cuando se llevan a cabo de manera apropiada, universalizarlos.

Tanto las entrevistas como los cuestionarios basan su información en la *validez* de la información verbal de percepciones, sentimientos, actitudes o conductas que transmite el encuestado, información que, en muchos casos, es difícil de contrastar y traducir a un sistema de medida, a una puntuación. Es esta característica lo que hace tan complejo establecer los criterios de calidad de este tipo de instrumentos.

El proceso de construcción y validación de un cuestionario/escala de medida es relativamente complejo y requiere el conocimiento teórico claro del aspecto que queremos medir, así como poseer conocimientos estadísticos y metodológicos adecuados. Lo que se pretende es garantizar que al cuestionario que se diseñe se le puedan aplicar los mismos criterios de validez y fiabilidad que exigimos a otros instrumentos tan precisos como una balanza. Por

tanto, como todo instrumento de medida, debe reunir las siguientes características:

1. Ser adecuado para el problema de salud que se pretende medir (teóricamente justificable -*validez de contenido*) e intuitivamente razonable.
2. Ser válido, en el sentido de ser capaz de medir aquellas características que pretenden medir y no otras.
3. Ser fiable, preciso, es decir, con un mínimo de error en la medida.
4. Ser sensible, que sea capaz de medir cambios tanto en los diferentes individuos como en la respuesta de un mismo individuo a través del tiempo.
5. Delimitar claramente sus componentes (dimensiones), de manera que cada uno contribuya al total de la escala de forma independiente (*validez de constructo*).
6. Estar basado en datos generados por los propios pacientes.
7. Ser aceptable por pacientes, usuarios, profesionales e investigadores (*viable*).

Finalmente, la validez del constructo implicará el grado en que el instrumento de medida refleja las teorías relevantes del fenómeno que mide.

Dado que lo que se está diseñando es una escala de medición que permita tener una puntuación de un aspecto de la salud, y poder comparar la de diferentes individuos o la del mismo individuo en diferentes momentos, se debe asegurar que el instrumento de medida sea fiable y válido.

Para ello recordemos que la *fiabilidad* es el grado en que un instrumento mide con precisión, sin error. Indica la condición del instrumento de ser capaz de ofrecer en su uso repetido resultados veraces y constantes en condiciones similares de medición.

La fiabilidad de un instrumento de medida se valora a través de la consistencia, la estabilidad temporal, la concordancia interobservadores y la repetibilidad.

La *validez* es el grado en que un instrumento mide

Correspondencia

E-mail: dscublinsky@yahoo.com.ar

aquello que realmente pretende medir o sirve para el propósito para el que ha sido construido.

La adaptación o traducción de escalas es una práctica muy útil en investigación, pero no por ello es simple. Resulta de extrema utilidad cuando un instrumento aplicado en otras partes del mundo ha demostrado ser efectivo. El proceso de traducción y adaptación de una escala requiere algo más que la traducción de la lengua origen a la lengua de destino: es necesario asegurar que las puntuaciones obtenidas con el test traducido sean equivalentes a las obtenidas con el test original.

Para alcanzar esa equivalencia, hay que considerar cuatro aspectos del proceso: 1) el contexto cultural donde se va a realizar la adaptación, 2) aspectos técnicos del propio desarrollo y adaptación del test, 3) administración del test, y 4) interpretación de las puntuaciones. En definitiva, es necesario asegurar que el instrumento de medida presenta las mismas propiedades métricas en las dos culturas (origen y destino), y que, por lo tanto, la interpretación de las puntuaciones sea la misma, es decir, que exista una equivalencia métrica.

Hoy en día, la utilidad de estas escalas/cuestionarios no se discute, y su aporte al entendimiento de enfermedades tan complejas como las reumáticas resulta de máxima relevancia. Sin embargo, una adecuada comprensión de sus límites y alcances, así como una clara apreciación de la importancia de su correcta elaboración y aplicación son centrales para su correcta interpretación tanto en el contexto de la investigación como en el de la práctica clínica.

Bibliografía

1. Alonso J, Prieto L, Antó JM. The Spanish version of the Nottingham Health Profile: a review of adaptation and instrument characteristics. *Qual Life Res* 1994; 3(6): 385-393.
2. Alonso J, Prieto L, Antó JM. La versión española del SF-36: un instrumento para la medida de los resultados clínicos. *Med Clin* 1995; 1.104: 771-776.
3. Badía X, Carné X. La evaluación de la calidad de vida en el contexto del ensayo clínico. *Med Clin* 1998; 110: 550-556.
4. Bulloch B, Tenenbein M. Validation of 2 pain scales for use in the pediatric emergency department. *Pediatrics*. 2002 Sep;110(3):e33.
5. Escriba V, Mas R, Cárdenas M, Pérez S. Validity of a job stressors measurement scale in nursing staff: the nursing stress scale. *Gaceta Sanitaria* 1999; 13(3): 191-200.
6. Dawson B, Trapp R. Bioestadística médica. *MM* 2005: 255-272.
7. Martínez-Martín P, Carroza-García E, Frades-Payo B, Rodríguez-Blázquez C, Forjaz MJ, de Pedro-Cuesta J; Grupo ELEP. Psychometric attributes of the Scales for Outcomes in Parkinson's Disease-Psychosocial (SCOPA-PS). *Rev Neurol*. 2009 Jul 1-15;49(1):1-7.
8. Praena Crespo M, Lora Espinosa A, Aquino Llinares N, Sánchez Sánchez AM, Jiménez Cortés A. The Spanish version of the Newcastle Asthma Knowledge Questionnaire for parents of children with asthma (NAKQ). Transcultural adaptation and reliability análisis. *An Pediatr (Barc)*. 2009 Mar;70(3):209-17.
9. Schneider A, Szecsenyi J, Barie S, Joest K, Rosemann T. Validation and cultural adaptation of a German version of the Physicians' Reactions to Uncertainty scales. *BMC Health Serv Res*. 2007 Jun 11;7:81.